Information and Knowledge Management
ISSN 2224-5758 (Paper) ISSN 2224-896X (Online)
Vol.4, No.12, 2014

Framework for Client-Server Distributed Database System for an Integrated Pay Roll System

Sese Tuperkiye Ebimobowei
Department of Computer Science, Bayelsa State College of Arts and Science P.M.B 168 Yenagoa, Bayelsa State
Email: tupsese@yahoo.com

ABSTRACT
A pay roll is a list of companies' employees and the amount of money that they are to be paid at the end of each work schedule, hourly, weekly or monthly. In a company, pay roll is the sum of all financial records of salaries of an employee, wages, bonuses and deductions.A nominal roll is a list containing employee details in a particular organization. An integrated pay roll system is an automated (computer based) pay roll system where individual autonomous pay roll system of an organizations are networked together to form a distributed pay roll system. This is initiated to ease pay roll transaction processing, improve transparency in pay roll processing and to reduce duplication of pay roll processing functions.The research work is to create a distributed pay roll system for local government areas in Bayelsa Sate. The system consists of a relational database of pay roll variables which could be shared by the various Local Government Areas of Bayelsa State. Each LGA will form a site. The database will be hosted by the server at the local government Service Commission office. All LGAs will access the database via a distributed network. The client/server distributed network architecture is used in the design and implementation of the system. The system is capable of monitoring an employee's nominal roll and pay roll records from the day of resumption of duty to the day of retirement, detection of an employee working in more than one local government, generation of reports concerning an employee or set of employees for promotion, demotion or retirement as the case may be, and automatic monthly pay roll computation.

1.0 INTRODUCTION

The major motivation behind the development of database system is the desire to integrate the operational data of an organization and to provide controlled access to the data.

Although integration and controlled access may imply centralization, this is not the intention. In fact, the development of computer networks promotes a decentralized mode of work. The decentralized approach makes the structure of many companies and organizations which are logically distributed into divisions, departments, projects and so on, and physically distributed into offices, plants, factories were each unit maintains its own operational data.

A distributed computer system consist of a collection of autonomous computers linked by a computer network, and equipped with the necessary computer networking software's such as servers and browsers to effect its communication with other computers (www.answers.com). Navathe (200) defined distributed database (DDB) as a collection of multiple logically integrated databases distributed over a computer network, and a distributed database management system (DDBMS) as a software system that manages a distributed database while making the distribution transparent to the user. The computer systems are connected together for the purpose of sharing data and information via sending message over a communication network to achieve a common objective. In this research, an attempt is made to design a framework for a client/server distributed database system for pay roll transaction in ministries and local government areas of Bayelsa state. WIkipedia (the free Encyclopedia) defined a pay roll as the sum of all financial records of salaries for an employee, wages, bonuses and deduction in a particular company. Also, a nominal roll is a list containing employee details in a particular organization.

The system consists of a relational database of pay roll variables (entities) that could be stored by the local governments of the state. Each Local Government will form a site. The database will be hosted by the Local Government Service Commission. The local government areas will access the database via the distributed network. The client/server distributed network architecture is used for the development of the system. The system is capable of monitoring an employee nominal roll and pay roll records from the day of resumption of duty to the day of retirement, detection of employee working in more than one local government or ministry, generation of reports concerning an employee or set of employees for promotion, demotion and retirement as the case may be, and automatic monthly pay roll computations.

2.0 PAY ROLL TRANSACTIONS IN LOCAL GOVERNMENTS AREAS OF BAYELSA STATE.

On resumption of duty, an employee's personal details are entered into a nominal roll. Credentials are also documented, which forms the data profile (record or file) of an employee on successful completion of the documentation process. A file number is issued to the employee, which is used to prepare his/her identification

card. If required, the employee is posted to the appropriate department. Two years after, a confirmation letter of appointment is issued to the employee. Entrances in a nominal roll include S/N, Name of Staff, Rank, Grade Level, Qualification, Sex, Date of Birth, Date of 1st Appointment, Date of Resent Appointment, Home Town, L G A, and Signature. The record of each employee is found under the department where he/she is posted.

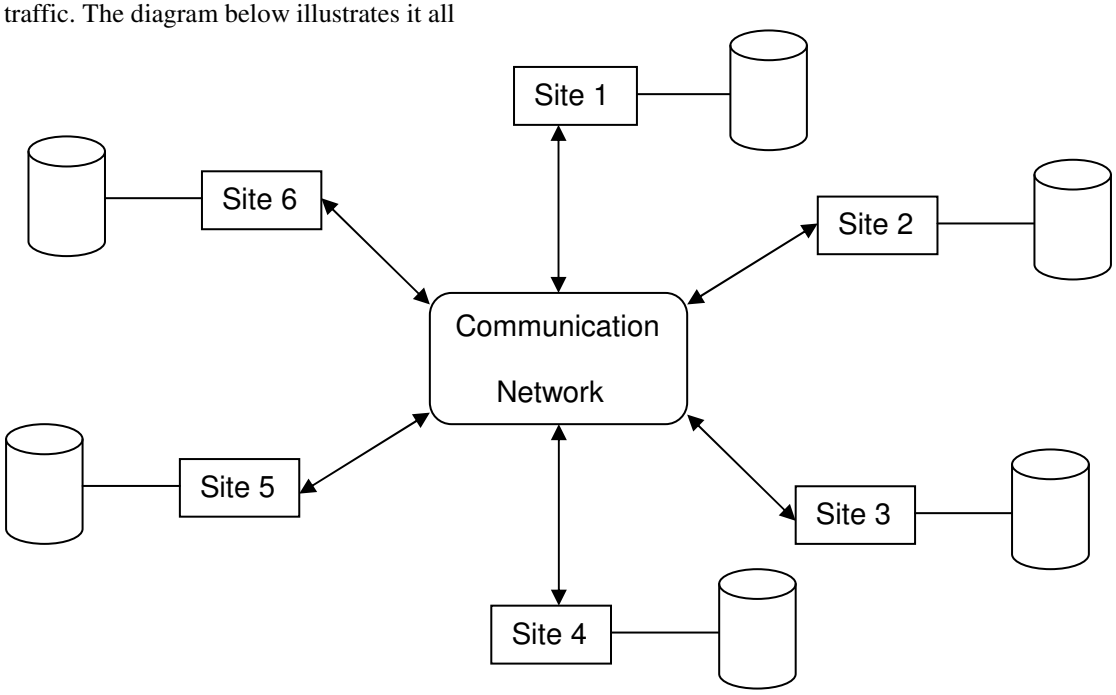
In time of pay roll preparation, each employee grade level, step and basic salary is used to prepare the pay roll. A pay roll consists of Basic Salaries of employees, deduction and allowances. All taxes payable are deducted from the basic salary, and all allowances are added to arrive at the take home pay (net pay) of an employee. The entries in a pay roll include. S/N, Name of Staff, Grade Level, Step, Basic Salary, Taxes, Allowances, Gross Pay, and Net pay.

Nominal rolls are subject to review, new employee details could be entered, promoted, demoted or retired employee details could be updated. Also dead employee details are permanently erased from a nominal roll.

Pay rolls are prepared at the mid week of every month. Each employee is meant to sign the nominal roll before his/her payment for the month is prepared. Failure to sign a nominal roll may warrant some sanctions, such as non-payment of salary for that particular month.

3.0 THEORY OF DISTRIBUTED DATABASE AND DISTRIBUTED DATABASE MANAGEMENT SYSTEM

A distributed system is defined as an information processing system that contains a number of independent computers that cooperate with one another over a communication network in order to achieve a specific objective. It can also be defined as a database in which portions of the database are stored on a multiple computers within a network. The definitions above means that a distributed system is not stored in its entirety at a single location. Instead, it is spread across a network of computers that are geographically dispersed and connected through communication network in which each site is a database in it's own right. The sites have agreed to work together so that data user can access data anywhere in the network exactly as if the data where all stored at the user's own site. Users have access to the portion of the database at their location so that they can access the data relevant to their tasks without interfering with work of others. Distributed databases system use a client/server architecture to process information requests. It allows faster local queries and can reduce network traffic. The diagram below illustrates it all



A Distributed Database Architecture

3.1 Advantages of distributed database system

1. Management of distributed data with different levels of transparency (Replication transparency, Fragmentation transparency Location transparency, Performance transparency, Transaction transparency Catalog transparency).
2. Capacity and growth: An advantage of distributed databases is that as the organization grows, new sites can be added with little or no upheaval to the DBMS. Compare this to the situation in a centralized

system, where growth entails upgrading with changes in hardware and software that affect the entire database.

3. Reliability and availability: An advantage of distributed databases is that even when a portion of a system (i.e. a local site) is down, the overall system remains available. With replicated data, the failure of one site still allows access to the replicated copy of the data from another site. The remaining sites continue to function. The greater accessibility enhances the reliability of the system.
4. Efficiency and flexibility: An advantage of distributed databases is that data is physically stored close to the anticipated point of use. Hence if usage patterns change then data can be dynamically moved or replicated to where it is most needed.
5. Distributed database sharing: An advantage of distributed databases is that users at a given site are able to access data stored at other sites and at the same time retain control over the data at their own site.
6. Protection of valuable data: If there were ever a catastrophic event such as a fire, all of the data would not be in one place, but distributed in multiple locations.
7. Improved performance: Data is located near the site of greatest demand, and the database systems themselves are parallelized, allowing load on the databases to be balanced among servers. (A high load on one module of the database won't affect other modules of the database in a distributed database.)

3.2 Disadvantages of distributed database system

1. Increased storage and infrastructure requirements because multiple copies of data are required at various separate locations which would require more disk space.
2. Security lapses have increased since data are in multiple locations.
3. Integrity control becomes more difficult.
4. Database design becomes more complex.

A key objective for a distributed system is that it looks like a centralized system to the user. The user does not need to know where a piece of data is stored physically.

3.3 Categories of Distributed Data

There are five categories of distributed data. They include:

- Replicated data
- Horizontally fragmented data
- Vertically fragmented data
- Reorganized data
- Separate-schema data

A distributed database management system (DDBMS) is a software system that permits the management of a distributed database and makes the distribution transparent to the users. A centralized distributed database management system (DDBMS) manages the database as if it were all stored on the same computer. The DDBMS synchronizes all the data periodically and, in cases where multiple users must access the same data, ensures that updates and deletes performed on the data at one location will be automatically reflected in the data stored elsewhere. The users and administrators of a distributed system should with proper implementation, interact with the system as if the system was centralized. Query optimization is essential if a DBMS is to achieve acceptable performance and efficiency. Distributed Database Management System is required to maintain distributed database and make it transparent to clients. Sometimes distributed database is used to refer jointly to the distributed database management system. Consequently, an application can simultaneously access and modify the data in several databases in a network. The main thing that all such systems have in common is the fact that data and software are distributed over multiple sites connected by some form of communication network. The Distributed Database Management System basically addresses the following technical processes:

1. Replica synchronization: This is about synchronizing data based on relatively smaller transactions where the said transactions may consist of several read and write operations on the server. But some applications can take relatively bigger data production jobs can write a whole file which can be a relatively large transactional file.
2. Synchronous and asynchronous replication: Replication may be done through synchronous or asynchronous or batch replication method which makes replicas be in synch (synchronous) or out of sync (asynchronous) for a certain period of time. Update reconciliation may be done at certain intervals such as every hour or every night.
3. Network servers and load: This refers to the management of computer on the network networks nodes which can act either as server or client or both server and client at certain circumstances. Under this area, other important considerations include traffic management and security aspects.

4. Heterogeneous data stores management: Different computer servers may be implemented on different platforms so support for heterogeneous data store should be greatly considered. Different kinds of data may be stored in different formats by different vendors. Even in the case of two different database paradigms namely relational and object oriented, a DDBMS needs to consider this aspect. The standard protocol used for directory information such as Lightweight Directory Access Protocol (LDAP) falls under this consideration.

3.4 Client/Server: Client/server systems are constructed so that the database can reside on a central computer, known as a server, and be shared among several users. Users access the server through a client or server application. In large client/server systems, thousands of users may be connected to a SQL Server installation at the same time. SQL Server has full protection for these environments, with safeguards that prevent problems such as having multiple users trying to update the same piece of data at the same time. SQL Server also allocates the available resources effectively, such as memory, network bandwidth, and disk input/output, among the multiple users.
Advantages of client/server:

- More efficient division of labour
- Horizontal and vertical scaling of resources
- Better price/performance on client machines
- Ability to use familiar tools on client machines
- Client access to remote data (through standards)
- Full DBMS functionality provided to client workstations

3.5 Collaborating Server: Herewe can have a collection of database servers, each capable of running transactions against local data, which cooperatively execute transactions spanning multiple servers. When a server receives a query that requires access to data at other servers, it generates appropriate sub-queries to be executed by other servers and puts the results together to compute answers to the original query. Ideally, the decomposition of the query should be done using cost-based optimization, taking into account the costs of network communication as well as local processing costs.

3.6 Middleware: Middleware is a crucial component of modern IT infrastructure. It is a set of common business-unaware services that enable applications and end users to interact with each other across a network. In essence, middleware is the software that resides above the network and below the business-aware application software. The services provided by these routines are available to the applications through application programming interfaces (APIs) and to the human users through commands and/or graphical user interfaces (GUIs).

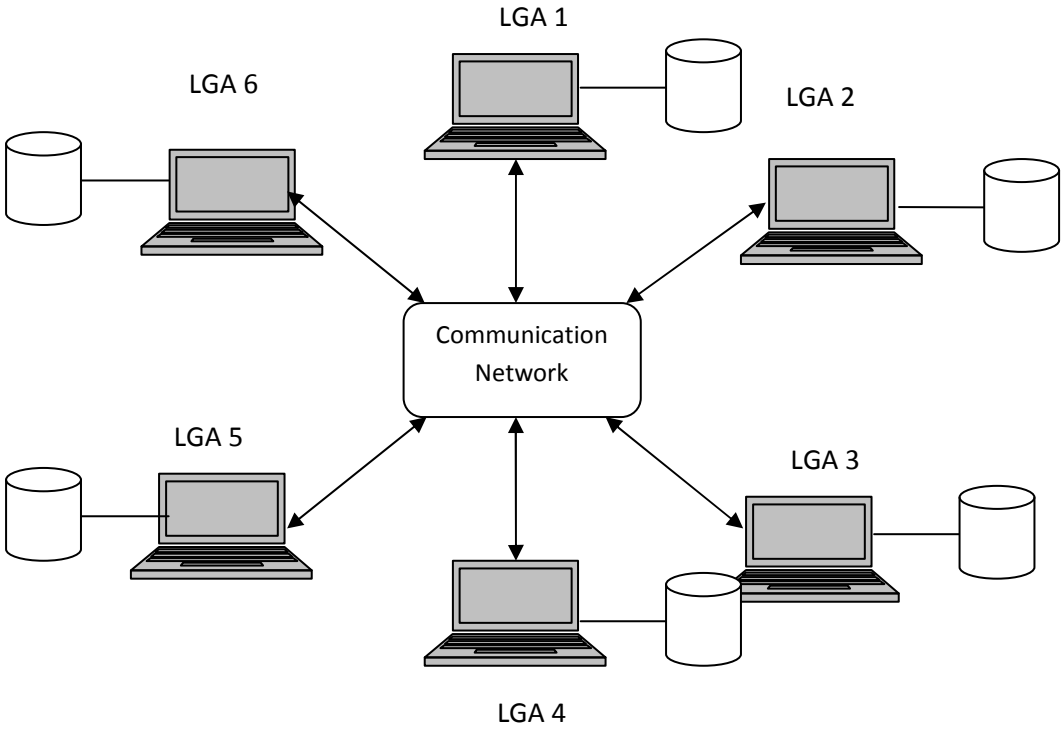
4.0 FRAMEWORK FOR THE CLIENT/SERVER ARCHITECTURE OF A DITRIBUTED SYSTEM FOR AN INTEGRATED PAY ROLL PROCESSING SYSTEM FOR LOCAL GOVERNMENTS OF BAYELSA STATE

Among the types of system architectures for informationprocessing, the client-server architecture is chosen for this framework. A Client is a computer or device requesting data from the Server, which is a computer hosting the data. It works through a network protocol, more than likely TCP/IP.Servers are computers that hold the actual databases and run only the database management system and related software. They are usually multiprocessor computers, with generous memory and raid disk arrays used for stable storage. Hardware database accelerators, connected to one or more servers via a high-speed channel, are also used in large volume transaction processing environments. Clients rely on servers for resources, such as files, devices, and even processing power. Interaction between client and server might proceed as follows during the processing of an SQL query:

- The client parses a user query and decomposes it into a number of site queries. Each site query is sent to the appropriate server site.
- Each server processes the local query and sends the resulting relation to the client site.
- The client site combines the results of the sub-queries to produce the result of the originally submitted query.

In this approach, the SQL server has also been called a transaction server (or a database processor (DP) or a back-end machine), whereas the client has been called an application processor (AP). In large client/server systems, thousands of users may be connected to a SQL Server at the same time. SQL Server has full protection for these environments, with safeguards that prevent problems such as having multiple users trying to update the

same piece of data at the same time. SQL Server also effectively allocates the available resources, such as memory, network bandwidth, and disk I/O, among the multiple users. While SQL Server works effectively as a server, it can also be used in applications that need stand-alone databases stored locally on the client. SQL Server can configure itself dynamically to run efficiently with the resources available on a client, without the need to dedicate a database administrator to each client. The interaction between client and server can be specified by the user at the client level or through a specialized DBMS client module that is part of the DBMS package. For example, the user may know what data is stored in each server, break down a query request into site sub-queries manually, and submit individual sub-queries to the various sites.



Software modules in a typical DDBMS are divided into three levels:

1. The server software is responsible for local data management at a site, much like centralized DBMS software.
2. The client software is responsible for most of the distribution functions; it accesses data distribution information from the DDBMS catalog and processes all requests that require access to more than one site. It also handles all user interfaces.
3. The communications software (sometimes in conjunction with a distributed operating system) provides the communication primitives that are used by the client to transmit commands and data among the various sites as needed. This is not strictly part of the DDBMS, but it provides essential communication primitives and services.

The client is responsible for generating a distributed execution plan for a multi-site query or transaction and for supervising distributed execution by sending commands to servers. These commands include local queries and transactions to be executed, as well as commands to transmit data to other clients or servers. Hence, client software should be included at any site where multi-site queries are submitted. Another function controlled by the client (or coordinator) is that of ensuring consistency of replicated copies of a data item by employing distributed concurrency control techniques.

5.0 SYSTEM DESIGN

The integrated pay roll system has the following design components.

DATABASE DESIGN

A database is a collection of logically related files (tables). A database management system is a software used in the creation and management of database. A file (table) is a collection of logically related records. Each record is uniquely identified by a primary key. A record is a collection of logically related fields (columns) that uniquely identifies a person. The database design is subdivided into conceptual and logical database design.

CONCEPTUAL SCHEMA

A conceptual schema describes data in terms of entities, attributes, and relationships among entities. An entity is an object or a concept under study. Entities of the system are: Nominal Roll and Pay Roll. Attributes of the two entities are:

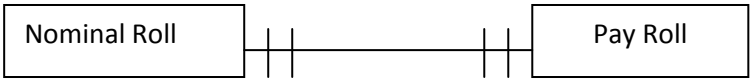
Nominal Roll:

S/N	Staff Name	Rank,	G/L	Qualification	Sex	Date of Birth	First Appointment Date	Recent Appointment Date	Home Town	LGA	Signature Passport

Pay Roll:

S/N	ID	Staff Name	Basic Salary	Tax	Medical Allowance	Transport Allowance	Total Deductions	Total Allowances	Gross Pay	Net Pay

There exist a one to one relationship among the two entities.



LOGICAL SCHEMA

The logical schema describes data in terms of a table. It maps the entities and their attributes into a table of rows representing individual records and columns representing the fields (attributes) describing the records. The structure of the two tables (nominal roll and pay roll) is presented below.

Norminal Roll					
S/N	Staff Name	Sex	G/L	Step	Birth Date
1	John Ebi	M	09	13	12/98/1978
2	Emeka Edith	F	08	6	09/11.1983

Pay Roll								
ID	S/N	Staff Name	G/L	B. Salary	Medical Allws	Tax	G Pay	N Pay
1	2	Emeka Edith	08	123000	5%	3%	…….	…….
2	1	John Ebi	09	231000	5%	3%	……	……..

5.1 COMPONENTS OF THE DISTRIBUTED SYSTEM

The distributed system has Internet facing Web-enabled applications that can be accessed remotely by the users either within the organization or remotely. The following is list of the information technology (IT) infrastructure components of the system

Firewall: A system designed to prevent unauthorized access to or from a private network. Firewalls can be implemented in both hardware and software, or a combination of both. Firewalls are frequently used to prevent unauthorized Internet users from accessing private networks connected to the Internet, especially intranets. All messages entering or leaving the intranet pass through the firewall, which examines each message and blocks those that do not meet the specified security criteria. There are several types of firewall techniques:

- *Packet Filter*: This looks at each packet entering or leaving the network and accepts or rejects it based on user-defined rules. Packet filtering is fairly effective and transparent to users, but it is difficult to configure. In addition, it is susceptible to Internet Protocol (IP) spoofing.
- *Application Gateway*: This applies security mechanisms to specific applications, such as File Transfer Protocol (FTP) and Telnet servers. This is very effective but can impose performance degradation.
- *Circuit-Level Gateway*: This applies security mechanisms when a Transmission Control Protocol (TCP) or User Datagram Protocol (UDP) connection is established. Once the connection has been made, packets can flow between the hosts without further checking.
- *Proxy Server*: Intercepts all messages entering and leaving the network. The proxy server effectively hides the true network addresses.

Router: A router is a special purpose computer or software device that enables two or more dissimilar networks to communicate. Routers route traffic, which consists of Transmission Control Protocol/Internet Protocol (TCP/IP) packets.

Host: A computer that is connected to a TCP/IP network, including the Internet.

Serve: A server is a dedicated computer that allows other computers to connect to it. The following are the various types of sever available:

- Domain Name System
- Web servers

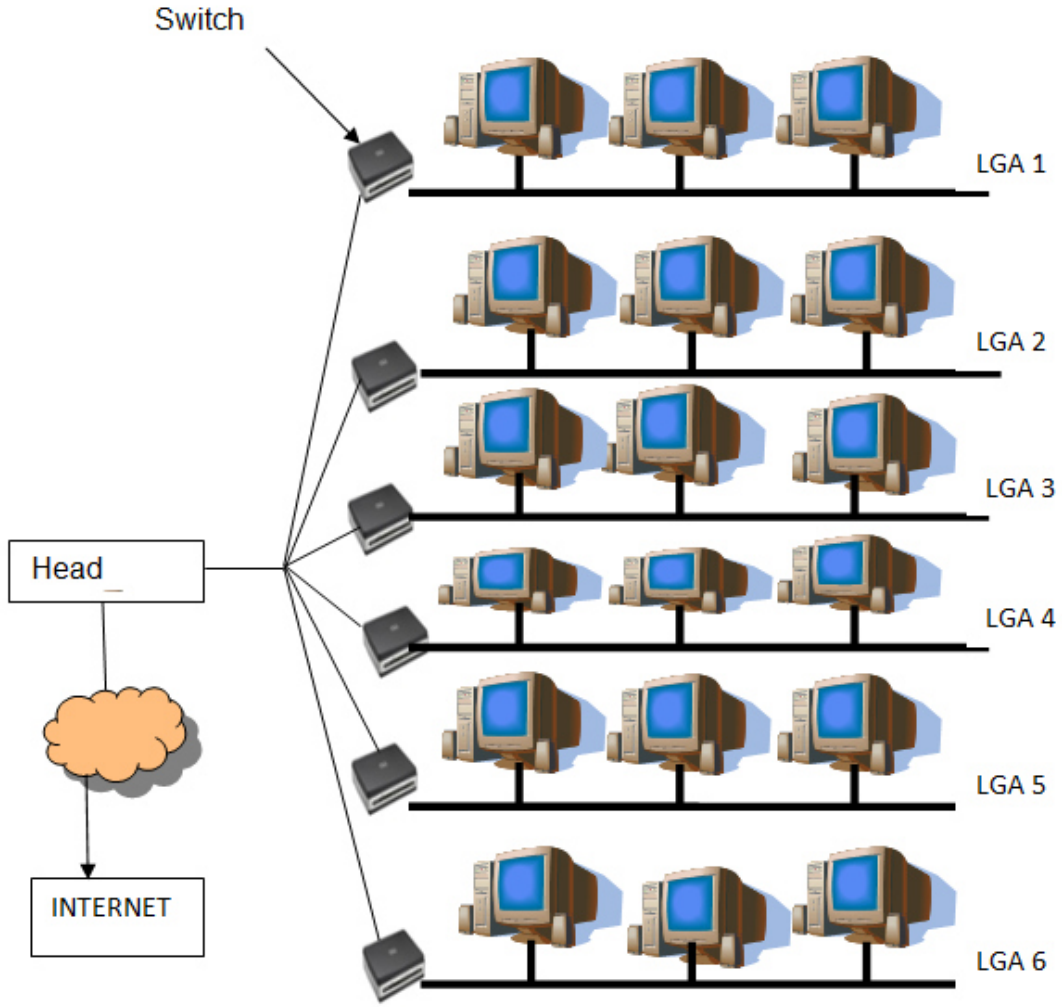
- Internet banking servers
- E-mail servers
- Proxy servers

Workstations: In networking, a workstation refers to any computer connected to a local area network. It could be a workstation or a personal computer.

Intrusion Detection Systems: Intrusion detection is fundamentally the process of monitoring computer networks and systems for violations of computer policy.

5.2 NETWORK STRUCTURE

A metropolitan area network (MAN), provides long distance transmission of data, image, audio, video information over large geographic area that may comprise a country, a continent or even the whole world. MAN is best for this kind of system. This system grew from earlier community antenna systems used in areas with poor over-the-air television reception. In these early systems, a large antenna was placed on top of a nearby hill and the signal was then piped to the subscriber's homes. The diagram below describes the structure of this framework.



Metropolitan Area Network (MAN) based system

6.0 CONCLUSION

The research work is to create a distributed pay roll system for local government areas in Bayelsa Sate. The system consists of a relational database of pay roll variables which could be shared by the various Local Government Areas of Bayelsa State. Each LGA will form a site. The database will be hosted by the server at the local government Service Commission office. All LGAs will access the database via a distributed network. The client/server distributed network architecture is used in the design and implementation of the system. The system is capable of monitoring an employee's nominal roll and pay roll records from the day of resumption of duty to the day of retirement, detection of an employee working in more than one local government, generation of

reports concerning an employee or set of employees for promotion, demotion or retirement as the case may be, and automatic monthly pay roll computation.

REFERENCES

Bayelsa state local government commission

Candace .C. Fleming and Barbara von Halle (1989): *Handbook of Relational Database Design.* Addison Wesley Longman.

Charles P. Pfleeger and Shari Lawrence Pfleeger (2003): *Security in Computing,*

Elmasri and Navathe, *Fundamentals of database systems*, 3rd Edition. Addison-Wesley Longman.

Fred R. McFadden and Jeffrey A. Hoffer (1994): Modern *Database management.* England. Addison Wesley Longman

http://www.cs.rpi.edu/~noel/distr_scaleup/distriuted.html, May 17, 2004.

Navathe, E. (2000): *Fundamental of Database Systems,* 3rd Edition. Teturo Sawada Exclusive Publisher and Distributor

Ozsu, M. T. and Valduriez, P. (2011): *Principles of Distributed Database System,* 3rd Edition. Springer.

Prentice Hall Professional Technical Reference, Upper Saddle River,New Jersey.

Rick Noel, "*Scale Up in Distributed Databases: A Key Design Goal for Distributed System,"*

Thomas M. Connolly and Carolyn E. Begg (2002): Database Systems (A Practical Approach to Design, Implementation, and Management.Third Edition. Pearson Education Limited. Addison Wesley.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar