International Journal of Current Engineering and Technology
E-ISSN 2277 – 4106, P-ISSN 2347 - 5161

Available at http://inpressco.com/category/ijcet

Research Article

Modeling a Conceptual Framework for secure Research Collaboration in Engineering and Technology

Nkolika O. Nwazor[A*] and Hyacinth C. Inyiama[B]

[A]Electronic and Computer Engineering Department, University of Port Harcourt, Choba, Rivers State Nigeria
[B]Electronic and Computer Engineering Department, Nnamdi Azikiwe University, Awka, Anambra State Nigeria

Accepted 05 Sept 2014, Available online 01 Oct 2014, **Vol.4, No.5 (Oct 2014)**

Abstract

This work modelled a Cloud Computing based Research Collaboration Platform (CRCM) for Engineering and Technology. It features automatic real-time network data acquisition from virtual infrastructures. The platform was designed based on Software as a Service cloud computing model. The integration of the process architecture of the CRCM with the backend server was exploited and the feasibility of an adaptive queuing scheme for CRCM service efficiency in a typical cloud scenario using Simevent tool was done. The system has the following features: Content Management System (CMS), Document and File Sharing (DFS), Real Time Chatting (RTC), data acquisition from virtual infrastructures, online research library, research tracking and information management system. The system security is based on modified 2-MD 5 cryptographic algorithm. The performance evaluation of the proposed system was carried out considering other existing collaboration systems in the context of service efficiency and modulated queuing scenario. It competed favourably well with similar systems with a very low page response, very good system security, capability of access to remote equipments, and intellectual property protection. It is a solution to the following engineering research problems: research duplication, poor collaboration, poor funding/ wastage of the available funds, inadequate infrastructure, serious brain drain and poor awareness creation.

Keywords: *Cloud Computing; Research, Collaboration, Digital Encryption, Virtual Infrastructures, Queuing, Modeling*

1. Introduction

Over the years, universities have been saddled with the responsibility of scholarship, innovation, research, knowledge generation and dissemination, and technology adaptation and re-engineering. With the creation of private and government research institutes, universities no longer remain sole citadels of research activities. For example, we have the National Agency for Science and Engineering Infrastructure (NASENI), in Nigeria with about nine research institutes responsible for research in different branches of Engineering and Technology.

Engineering and technology research is saddled with the problem of human limitations that affect our ability to create or design any system. When working at high levels of abstraction, we are slow and error-prone (Whitehead, 2007). Engineering and technology research in developing nations like Nigeria is faced with the following challenges: serious brain drain as the intellectuals in developing nations that travel to developed nations for further studies or greener pastures are retained in those developed nations; poor local and international collaborations; high cost of information and technology infrastructure; difficulty in assessing and managing the large amount of data generated from research; duplication of research work due to poor tracking of the ongoing in the research world; poor interaction between the researchers and the end users/industrialists; poor documentation; inadequate funding; inadequate infrastructures and geographical separation to mention but a few.

In the last decades, research collaboration has become one of the focus themes in engineering and technology studies. With abundant information available, the scientific community (researchers, industrialist, etc) seeks to have effective research collaboration via a virtual platform with robust stability and reliable information provisioning. Collaborative platforms should allow better collaboration and help save precious time that can be put to use in making scientific breakthroughs. This research work goes a long way in solving majority of these problems.

2. Leveraging on Cloud Computing

A lot of software has been developed to aid research work. Previously, this software would typically be a packaged solution installed directly onto each individual personal computer. But over the past few years, with the explosion of cheap bandwidth and server technology as well as the burgeoning open source movement, companies are gravitating towards a database package that is not managed on an individual machine but hosted upon a server and viewed entirely through an internet browser.

*Corresponding author: **Nkolika O. Nwazor**

Cloud computing has this capability. Cloud computing is a model for enabling network users' on-demand access to a shared pool of configurable computing resources that can be rapidly provisioned and released to the client without direct service provider interaction (Slaheddine, 2012). It removes the cost of infrastructure and maintenance from the end user and places it on the service providers. Instead of having the software installed in individual computers with the high licensing fee involved, it is installed in only one computer and made available to others through the cloud. Cloud computing will help in accessing and manipulating more data, providing security and applying more complex calculations. Generally speaking, the architecture of a cloud computing environment can be divided into five layers: facility layer, infrastructure layer, hardware layer, platform layer and application layer. It can also be in the following models: Software as a Service (SaaS), Platform as a Service (PaaS) and Infrastructure as a Service (IaaS). The major components of cloud computing includes: application, client, infrastructure, platform, service and storage. Cloud computing can be deployed as a private, community, public or hybrid cloud. A data center, which is home to the computation power and storage, is central to cloud computing and contains thousands of devices like servers, switches and routers (Zhang *et al*, 2010)

3. Objectives of the Research Work

This work aims at modeling a research collaboration platform for engineering and technology. In this regard, the work develops a conceptual framework for secure research collaboration in engineering and technology with automatic real-time network data generation from virtual infrastructures. In our approach, a conceptual framework is developed, the platform integration is developed, and the complete system specification is presented while the data center simulation for the collaboration framework is achieved.

4. Review of Related Works on Research Collaboration Platforms

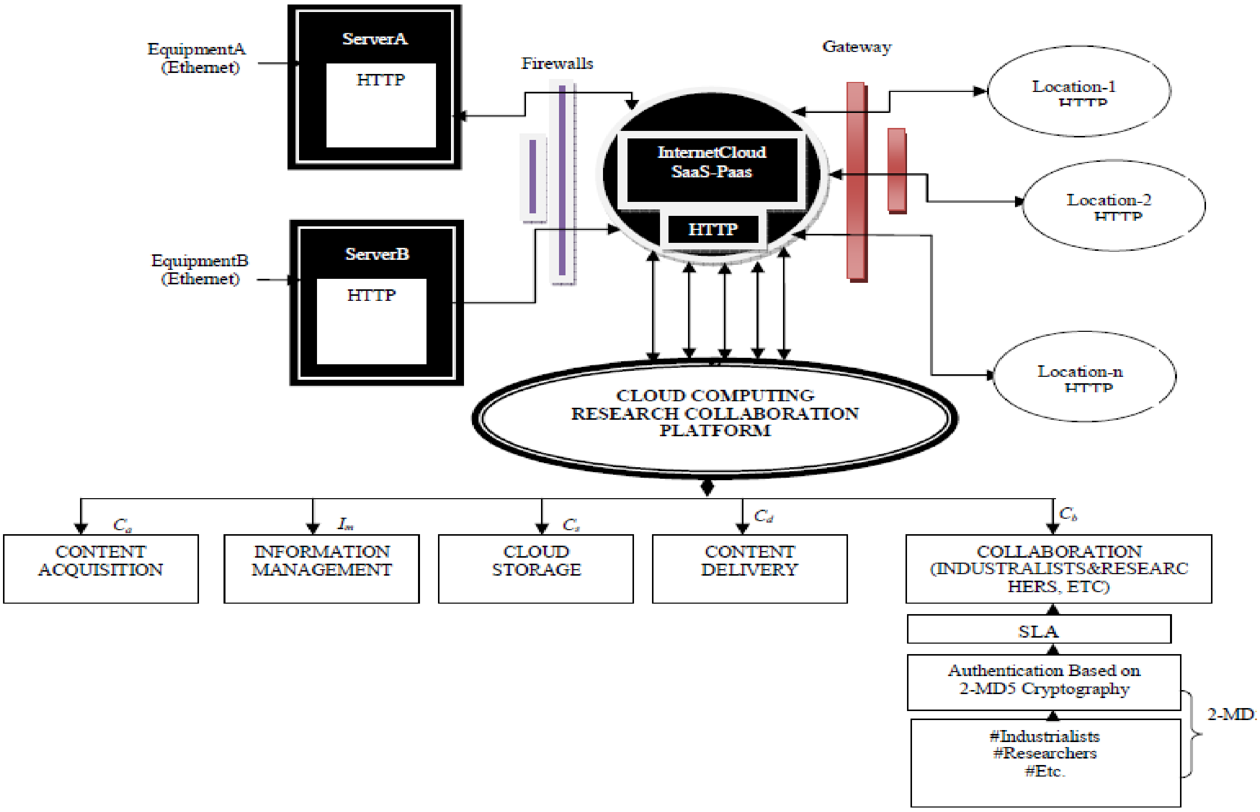
A research collaboration platform is a category of business software that adds broad social networking capabilities to work processes (Search content management, 2012). The goal of a research collaboration platform is to foster innovation by incorporating knowledge management into research processes so researchers can share information and solve research problems more efficiently. This work studied various collaboration research and collaboration systems. Lavina *et al* in their work presented a possible solution for the development of a virtual team environment that fosters better learning and innovation and ensures a good collaboration between individuals working in an innovation project (Lavinia *et al*, 2006). Some of the existing research collaboration platforms include: MyNetResearch, Research gate, Laboratree, HubZero, Alfresco etc. From the analysis of these and other systems (Berman *et al*, 2012), we noted the following limitations: High Density Web Queuing, Real time Traffic limitations, Application Domain Restrictions, Poor Tracking, Security Vulnerability Challenge, Intellectual Property challenges, High Infrastructure Cost Index (High Capex), and no real time Data Generation for traffic analysis to mention but a few.

5. Proposed System: Cloud-Based Research Collaboration Model

This system is made up of five major sub-systems carefully integrated into the internet cloud for remote access. They include: content acquisition (for collection of research data and information), information management (for content search, data analysis and organisation), storage infrastructure (for storage of data and information generated), content delivery (for display and exchange of the acquired data/information), and collaboration (consists of the scientific community- researchers and industrialists) subsystems. The system architecture presented in Fig.1 is the envisaged system which will not only solve the problem of research collaboration philosophies but will address the challenging problems of how to effectively allocate system resources to meet Service Level Objectives (SLOs) for users and cloud services providers. From this framework, we propose a scheme for autonomous performance control of our cloud collaboration application platform which uses a queuing model predictor with an online adaptive modulation loop that enforces moderate queue and admission control of the incoming requests to ensure the desired response time target is met. The proposed Queuing-Model-Based Adaptive Control approach combines both the modelling power of queuing theory and self-tuning power of adaptive control. Therefore, it can handle both modelling inaccuracies and load disturbances in a better way. To evaluate the proposed approach, the queuing model for our case study cloud application was built and simulated with MATLAB simevent components widely used in industry.

Also, Fig.1 captures the cryptographic security algorithm implemented in the CRCP where a double Message Digest5 scheme is applied in our collaboration module of the CRCP. Such intelligent web integration with the CRCP/platform of Fig.1 demonstrates high efficiency and low cost deployment which is fundamentally lacking in the existing systems. From the block diagram of Fig.1, the layer-1 authentication utilizes the connection framework of Fig.2 for security validation. The users are required to supply their login details which are encrypted with a cryptographic algorithm which is a modified form of Double Message Digest 5 (2-MD5). Afterwards, the grid control or the logic instance then connects the user process to the server process if all the layer security requirements are fulfilled. This work uses an advanced encryption algorithm to encrypt user login details from the server process.

Firstly, the user process is initialized before the validation by the modified 2-MD5 from the server process at the backend. After the validation, to access the collaboration module of the CRCP will require a basic SLA endorsement. All the 2-MD5 passkeys are securely stored in the server connecting to the server follows the framework shown in Fig.2.



. **Fig.1** Block diagram of the proposed system

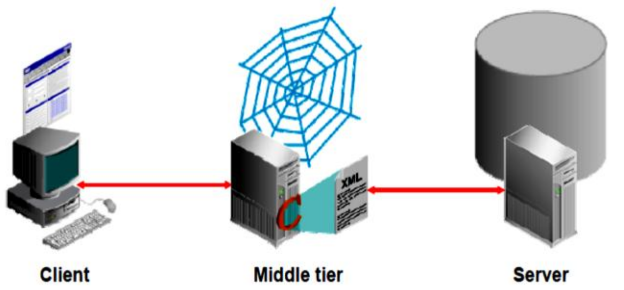


Fig.2 User instance server connection

From Fig.3, a logic instance is the connection bridge between the user process and the server process down to the backend server. In the design phase, the connection manager is the link between the users and the backend server as depicted in Fig.3.

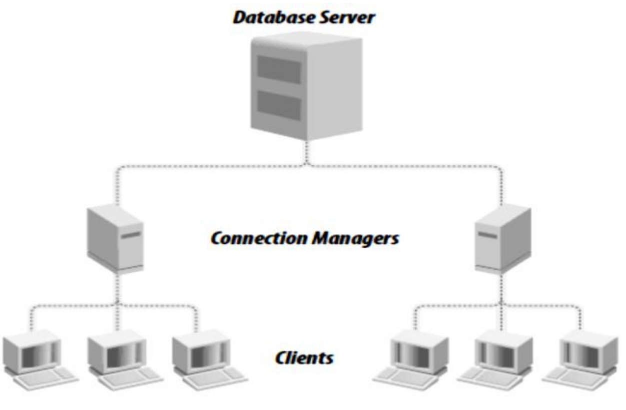


Fig. 3 Logic instance connection middleware

When a failure occurs in the connection manager, then the CRCP will experience a momentary outage until the background processes restores such crash. As shown in Fig.4, the CRCP security block is designed to encapsulate the presentation layer which is the graphical user interface, the application layer which is the business logic or the logic instance and the exchange infrastructure layer which communicate with the server for service provisioning. From Fig.4, our main integration framework uses CRCP grid control to address the issues of software provisioning, manageability, High availability, performance, Security, Information integration and application service monitoring.

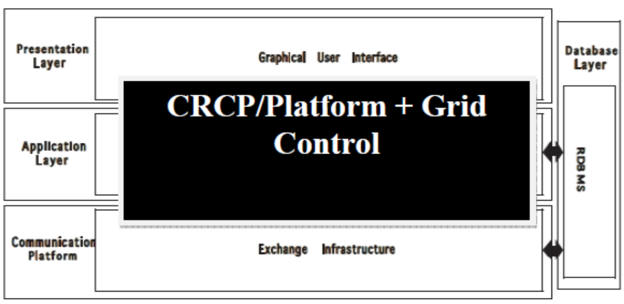


Fig.4 Block diagram of CRCP Interfaces with grid control

6. Cloud Queuing Model using MATLAB Simevent Tool

MATLAB Simevent tool 2009b was used to develop the cloud queuing model. Fig.5, shows the process model of the implemented model showing the Researchers/Industrialist background, the Internet

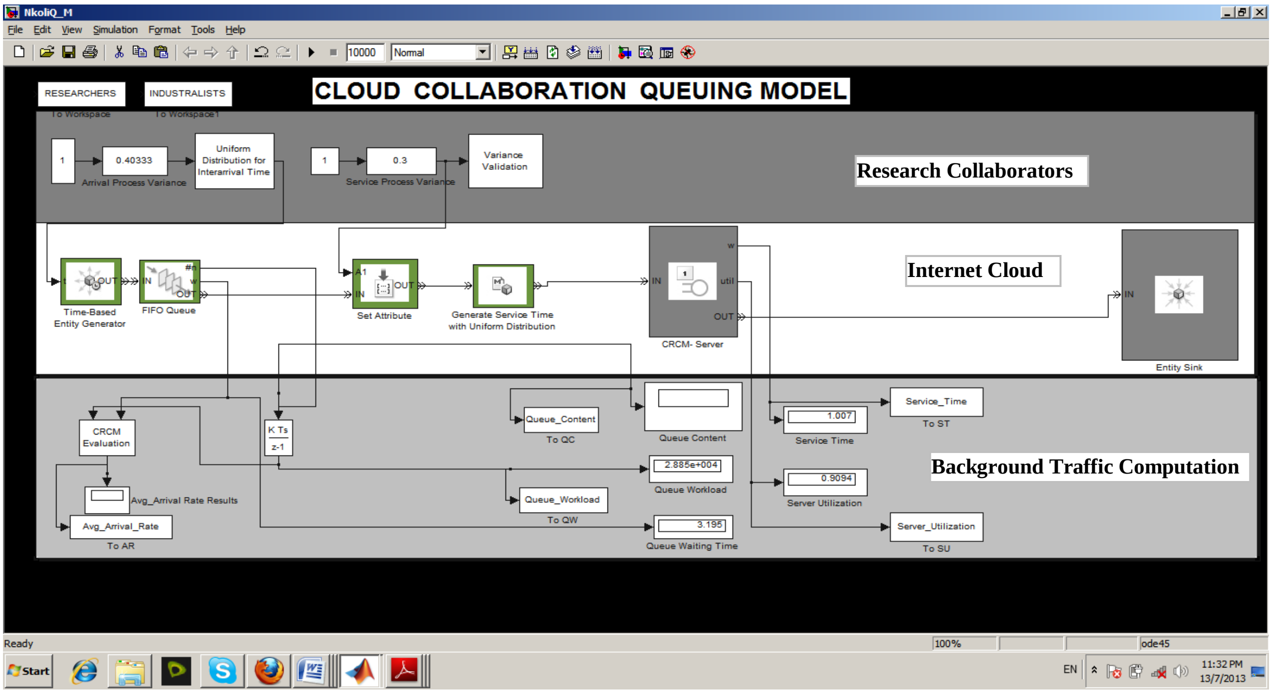


Fig.5 Cloud collaboration queuing model for CRCM

background and the traffic background as well. From Fig.5, the system response was ascertained by moving the arrival process variance slider or the Service Process Variance slider during the simulation while observing how the queue content changes. When traffic intensity is high, the average waiting time in the queue is approximately linear in the variances of the inter-arrival time and service time. The larger the variances are, the longer an entity has to wait, and the more entities are waiting in the system.

6.1 Cloud Queuing Model Deductions

From the deductions made from the Cloud collaboration queuing model, it was observed that the Queuing workload *QW* which shows a nominal throughput gave 2.88×10^4, Queuing waiting time was observed to be 3.195Sec, Server Utilization was shown to be 0.9094, the service time (TS) was observed to be 1.007sec. These were obtained by ensuring that the optimal values of the arrival process variance (0.4033) and service process variance (0.3) were kept constant. In the integration phase of Fig.1, this work now proposes these values for such integration to facilitate service efficiency and optimal performance at large

6.2 Grid Control Management Framework (GCMF)

From Fig.1, this work introduced a GCMF whose role is to coordinate all the activities of the CRCP concurrently. The starting point in GCMF is the process architecture that will streamline the interaction between the user and the server backend. Fig.6 shows the developed block diagram for the process architecture. The metrics of Fig.5 will be considered in GCMF particularly at the server end where resource utilization and background process management is vital.

The CRCP client server architecture shown in Fig.6 serves as the engine block of figure1. The architecture is divided into the user process, instance, server process, background processes and the database viz-a-viz the server.

In its operational status, the CRCP server creates server processes that handle the requests of users or client processes that connect to the logic instance. The logic instance is a collection of the background processes and the shared memory allocated by the server during the normal operation. It is also the connection bridge as shown in Fig.6. The user process is the connection to the CRCP server having passed the authentication and instance connection In CRCP, a set of background processes for the logic instance that interacts with each other and with the server to manage memory structure, and asynchronously perform I/O to write data to the server are enabled. They provide parallelism for better performance and reliability.

They include: the System Monitor which performs crash recovery when the logic instance fails in the CloudMesh ERP; the Process Monitor which performs process recovery when a user process fails; Server database (DBW*n)* which writes to table blocks in the buffer cache of MySQL server on server disk; Checkpoint (CKPT) process which is responsible for signalling DBWn for recent updates in the CloudMesh and Log writer (LGWR) which writes buffer log entries to the server while the archive stores log entries of the XAMP database as used in the design. Essentially, these background processes are taken care of by the server during the installation process. From Fig.6 also, the server process uses the Program Global Area (PGA) privately to communicate to the background process logically.

The Data mining capability of the CRCP is realized in the storage structure of the CRCP server. The storage structure consist of the Logical storage which has the database, schema, and table space for logical data storage and the Physical storage which has the data files, control files, and redo log files for multiplexing of data, security reintegration, and automatic updates on the server.

Every transaction in the CRCP of Fig.1 forms a session based on structured query language patterning for user connection into the server via the logic instance as shown in Fig.7

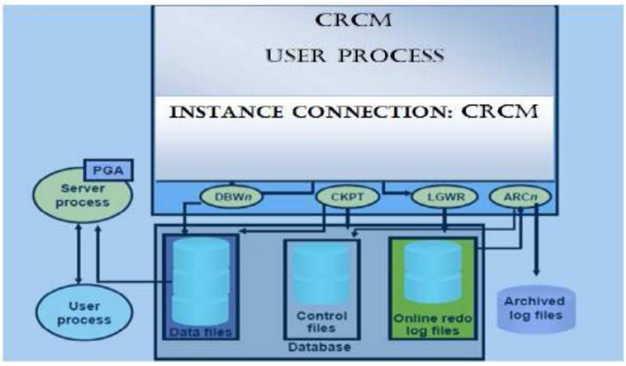


Fig.6 CRCP Process architecture block

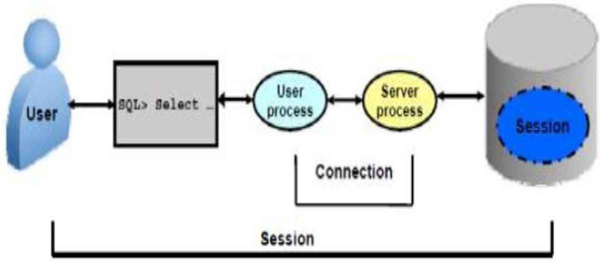


Fig.7 Interfacing with the CRCM Server

7. CRCP System Design Methodology

Program design involves solution statement and coding. The composites design approach adopted in the design of CRCP involves the waterfall (Software Development Life Cycle) SDLC approach. The design objective in this case is to develop the system flowcharts encompassing (CRCP Global Module, CRCP Content Acquisition Module, CRCP Information Management Module, CRCP Cloud Storage Module, CRCP Content Delivery Module, CRCP Collaboration/Industrialization Module, CRCP Service Level Agreement Module, etc). This was achieved using OODA with XAMP Php, MySQL server database, so as to realize Fig.1 while taking cognizance of the performance of the queuing metrics discussed in Fig.5.

7.1 System Specification Requirements

The software requirements include: Microsoft windows vista or windows7 while using Linux Redhat for production deployment, Apache server version 1.3.14 & above, PHP version 4.1.0 & above, MYSQL version 4.1.0 & above, adequate temporary space for paginations to virtual memory, 64-bit and 32-bit OS, OS patch level 1, System and kernel parameters must be enabled, Sufficient swapping, nonempty XAMP htdocs_HOME and MySQL database

The hardware requirements include: Monitor, 4GHZ or faster processor, 4GB of RAM, 1TB of available hard-disk space, 1280 X 800 display with 64-bit video card, 1 GB for the logic instance (grid control), 1.5 GB of swap space, 400 MB of disk space in the /tmp directory, Between 1.5 GB and 3.5 GB for the CRCP, 1.2 GB for the preconfigured database (optional), and 2.4 GB for the flash recovery area (optional) temporary space for paginations to virtual memory, 64-bit and 32-bit OS, OS patch level 1, System and kernel parameters must be enabled, Sufficient swapping, nonempty XAMP htdocs_HOME and MySQL database.

7.2 Programming and Database Technologies

Hypertext pre-processor (PHP) was chosen for the implementation of the system owing to its light weighted nature which is acceptable in cloud computing context. Also it is widely used for general purpose scripting and can be embedded into HTML. Owing to the presence of the apache server, the CRCP generally runs on a web browser which needs to be configured to process the PHP code. External modules (CMS, Opendoc, Real time chats, backup and recovery for the cloud compute storage, etc) are all unified using the PHP script. Fig.8 is a sample web page of the CRCP platform. The database was realised using MySQL server database for all the subsystem modules. The CRCP Data mining capability involves the analysis and implementation steps of the MySQL XAMPP control panel Knowledge Discovery process, (or KDD). This work implemented this capability as a computational process of discovering patterns in large data sets that connects the logic instances to the DBMS server. The overall goal of the data mining process is to extract information from a data set and transform it into an understandable structure for the application use. It basically involves database and data management aspects, data pre-processing, and inference considerations for post-processing of discovered structures, visualization, and online updating. The actual data mining task is based on the entity relationship model for the chat system, content management, messages and other data storage schemes. These are made to be the automatic analysis of large quantities of data to extract previously unknown interesting patterns such as groups of data records (cluster analysis), and dependencies (association rule mining). In MySql database, this usually involves using database techniques such as spatial indices and complex predictive analytics to organize the data sets as they are inputted into the database.

7.3 CRCP Data Center Design Model

Fig.9 shows the data center implementation carried out using OPNET simulator. The design specifications of the data centre model for the developed CRCP comprise a multi-tier backend integration involving external equipment connections to CRCP servers which delivers web services to user locations and the devices. The model has database layered design supporting various interfaces such as enterprise business ERP and CRM solutions from Siebel and Oracle, etc. It relies on security and application optimization services to be provided in the network. Also, it facilitates high-performance computing (HPC), parallel computing, and high-throughput computing (HTC) environments, but can also be associated with grid/utility computing. This design is typically based on customized or proprietary application architecture for CRCP. Since, the multi-tier data centre model is dominated by HTTP-based applications, the simulation design runs separate processes on the same machine as a server comprising of the following three tiers: Web-server, Application and Database

Fig.8 CRCP Data centre Simulation Testbed for CRCP HTTP Service

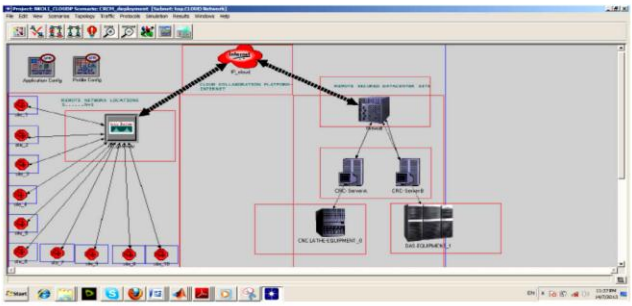


Fig.9 CRCP Data centre Simulation Testbed for CRCP HTTP Service

Conclusion

A Cloud computing platform for secure remote research collaboration in engineering and technology with automatic real-time network data generation from virtual infrastructures particularly for developing nations was designed in this work. This is geared towards limiting some of the challenges facing engineering and technology research in developing countries like Nigeria due to lack of infrastructure, geographical barriers, and access to required expertise to mention but a few

References

Jim Whitehead, (2007), Collaboration in Software Engineering: A Roadmap, http:/ /citeseerx.ist.psu. edu/viewdoc/download?doi=10.1.1.79.7594 &rep=rep1&type=pdf.

Slaheddine MAAREF, (april 2012), Cloud computing in Africa Situation and perspectives. *REPORT by the telecommunication development sector of ITU*., http://www.itu.int/ITUD/treg/publications/Cloud_Computing_Afrique-e.pdf.

Qi Zhang, Lu Cheng and Raouf Boutaba, (2010), Cloud computing: State-of-the-Art and Research Challenges, *Journal of Internet Services and Applications (JISA),* http://cloud.pubs.dbs.unileipzig.de/sites/cloud.pubs.dbs.ui-leipzig.de/files/fulltext.pdf

Search content management, (2012), Collaboration platform. http:// searchcontentmanagement. techtarget. com/definition/collaboration-platform.

Lavinia Precup *et al*, (2006), Virtual team environment for collaborative research projects. *International Journal of Innovation and Learning. Inderscience Publishers.* Volume 3, Number 1. pp 77-94.

Berman *et al*, 2012 Collaborative software for traditional and translational research. *Human Genomics.* http://www.humgenomics. com/ content/6/1/21. Retrieved 23rd February 2013