Enhanced Accounting Scheme for Grid Computing Architecture

Akinwunmi A.O.
Department of Computer Science and Information Technology Bowen University Iwo Nigeria

Aderounmu G.A.
Department of Computer Science and Engineering ObafemiAwolowoUniversity Ile-Ife Nigeria

Onifade O.F.W.
Department of Computer Science University of Ibadan, Nigeria

ABSTRACT
This research developed a modification to the Grid Accounting Scheme (GridBank) by formulating a model to enhance the scheme. The enhanced scheme was simulated and its performance was evaluated. Thiswas done with a view to eliminating the manual mode of processing as well as speed up transactions and reduce time delay. The PayPal layer was added to the existing three layers which enhanced the scheme to allow for the automation of the GridBank administration module. The enhanced scheme was formulated using the web service approach that allowed cross platform interoperability. The results of the simulation showed that as the number of users increased, the processing time gradually reduced for the enhanced scheme which made its processing delay to be reduced. Also as the number of available resources increased the enhanced scheme scaled the load properly. These showed improvements over the existing scheme. It was concluded that the enhanced accounting scheme provided the required automation for efficient and secure grid accounting operations.

General Terms
GridAccounting, Resources, Web Service, GridBank

Keywords
Grid Accounting Scheme, Processing, Delay, Load, Scalability

1. INTRODUCTION
The recent trend of growing computational power requirements of various applications has promoted the need to synergistically couple high performance computing resources distributed across different organizations together [1]. This implies that there is a need for proper management and accountability of resources usage [3]. It is essential for resource owners to account for the usage of their various resources shared among different organizations. Therefore, an accounting system that is reliable is required for proper management of the operations.

Manual mode of operations has not been adequate in any worthwhile information processing system and grid accounting system was not an exception. The result from such system tends to be inconsistent, inaccurate, associated with poor quality of service and an attendant time delay because of human factor.

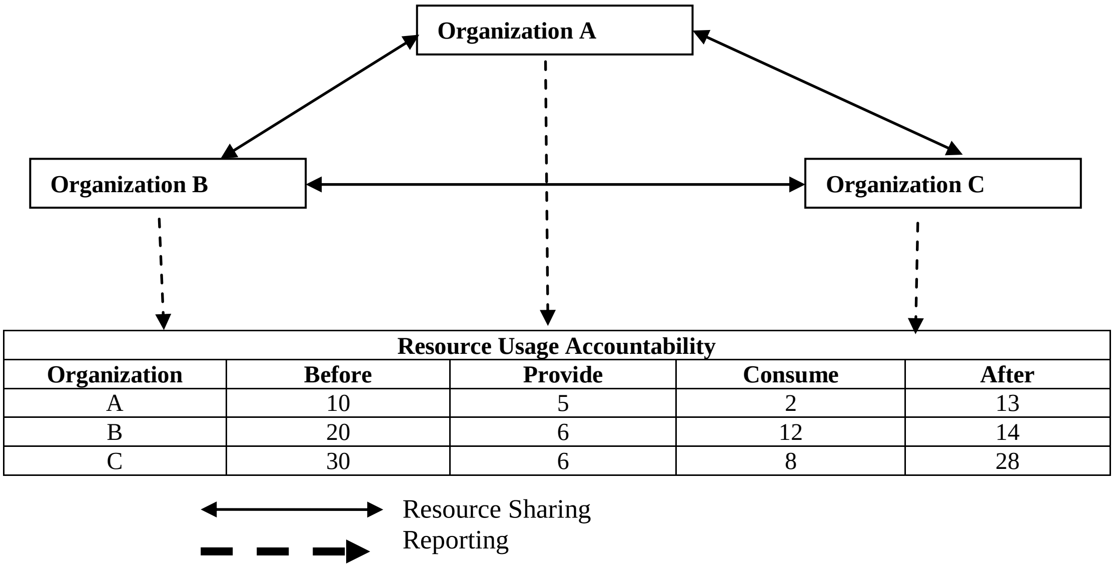
Hence there is the need for such manual mode of operations to be automated [4] so that performance degradation of such system could be reduced and result generated would be more reliable and timely.

In the Grid Accounting System GridBank (GB), account management functions are carried out by administrator as well as the transfer of real money to and from client manually. Hence the system performance is inefficient and insecure.

The focus of this study is to ensure good quality of service and system performance improvement through automation of the account management functions.

This work was organized into sections. Section one gave the background to the problem. Section two focused on the related work. Section three discussed the model architecture while modeling and simulation of the proposed model was dealt with in section four. In section five results of the simulation was discussed. Section six focused on the model validation while section seven concluded the work.

2. RELATED WORK
In brief, Grid accounting could be defined as a mechanism to maintain accounts or report on the Grid resource usages between different parties involved in the resource sharing [3]. Hence, the final outcome of a Grid accounting system would be a report of accountability of each party in terms of their consumption and provision of resources as illustrated clearly in Figure 1 below.



Resource Usage Accountability				
Organization	**Before**	**Provide**	**Consume**	**After**
A	10	5	2	13
B	20	6	12	14
C	30	6	8	28

Figure 1: A high level overview of a grid accounting process and the eventual outcomeexpected[2]

There are had been several attempts in providing a reliable accounting framework for the grid architecture some of which are shown in Table 1 below:

Table 1. Various grid accounting systems

Name	Feature	Developed By	Used By
GSAX	Extensible OGSA accounting and logging framework which provided functional modular accounting framework that can be expanded [5]	GGF	None
NGS	NGS User accounting system which handled the automation of user registration and account approval [7]	National Grid Service	National Grid Service
SGAS	SweGrid accounting solution was also another which made provision for accounting solution for clusters within SweGrid Grid [8]	SweGrid	SweGrid, Globus4
APEL	Another attempt was the Accounting Processor for Event Logs (APEL) which dealt with job accounting tool using relational Grid Monitoring Architecture [9]	LCG	LCG/EGEE
GASA	GridBank (Grid Accounting Services Architecture) which gave Grid accounting requirements and economic models that can be enhanced [4]	Gridbus	Gridbus
MOGAS	MOGAS (Multi-organizational Grid Accounting System) provided a general architecture for accounting in Multi-organizational Grid environment [3]	Nanyang Technological University	Nanyang Campus Grid
DGAS	DGAS(Distributed Grid Accounting System) provided accounting services in distributed grid environment [10]	EGEE	EGEE
GRATIA	The Gratia system was designed to be a robust, scalable, trustable, dependable grid accounting service. [2]	Fermilab	Open Science Grid
EEGAS	A Java Enterprise based grid accounting system for Distributed Multi-Domain Resource Accounting[6]	Brenner et al	European Grid

For the purpose of this work GridBank was chosen for the following identified reasons:

- It consists of several modules that can be enhanced or replaced without affecting other modules.
- It is inherently highly secured.
- It allows for proper interaction among various components of the grid
- It is a web service itself
- Access scalability is provided for.
- There are different payment strategies to cater for different choice of service

2.1 GridBank

The existing GridBank's server architecture (see Figure 2) consisted of several modules that could be enhanced or replaced without affecting other modules [4]. These modules were organized into three layers. Accounts Layer deals with database and account operations. Payment Protocol Layer defines payment schemes, message formats and communication protocols. Security Layer ensures that any messages passed to Payment Protocol and consequently invoking operations in Accounts Layers are authentic.

GridBank(GB) database module is a relational database that stores account and transaction information.

GB Accounts is the core module interacting with the GB database. It provides functions for basic account operations such as creation of accounts, requesting and updating account details, transfer of funds from one account to another, locking funds and transfer from locked funds. This module is independent of payment scheme, protocols used and underlying security model. Its purpose is to perform database operations that deal with manipulating and managing GridBank's database.

GB Admin module provides account management such as deposit, withdrawal, change credit limit, cancel transfers and close account functions. These functions are performed by GridBank's administrators who are responsible for transferring real money to and from clients. In the future, this process can be automated by using other payment systems such as PayPal or credit card transactions by importing data from those systems [4].

Administrators had a privileged access and the credentials for such access are checked by GB Administration module.

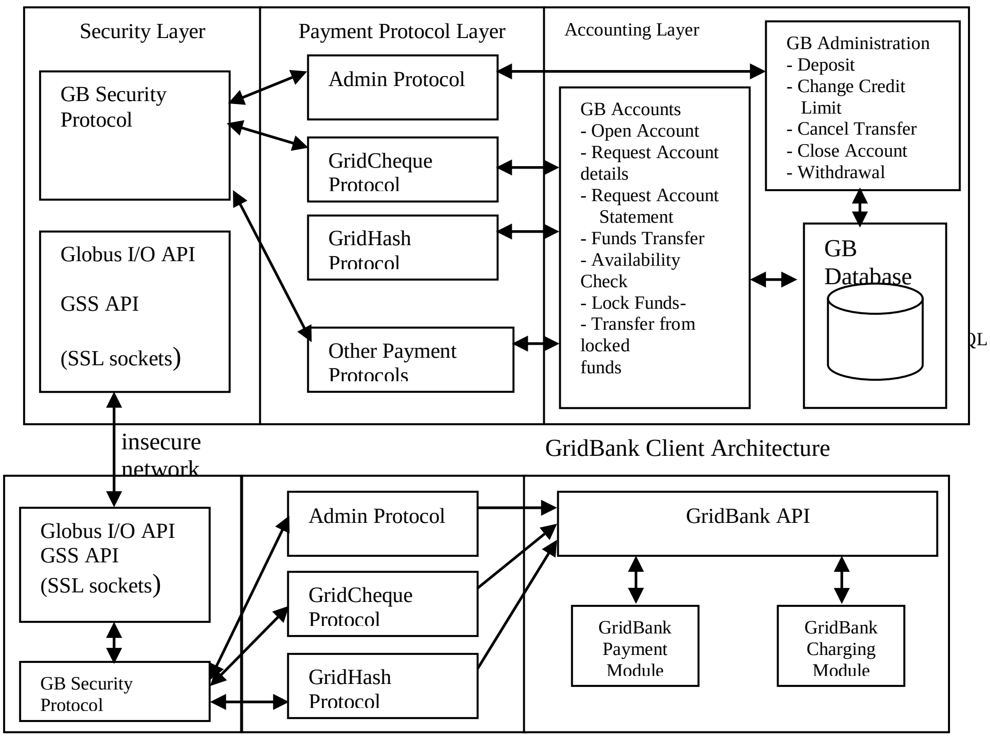


Figure 2: GridBank server architecture [3]

2.2 Web Service Approach

Since GridBank administrators carried out the functions of GridBank Administration module in the accounting layer manually and they were also responsible for transferring real money to and from client.

An approach based on web services is proposed in this research work in order to solve these problems by automating the functions as well as importing data from the PayPal system in order to improve the efficiency of the system.

Web services help solve the interoperability problem by giving different applications a way to link their data. Using Web services approach, data can be exchanged between different applications and different platforms [10].

3. THE MODEL ARCHITECTURE

The web service created automated the GridBank Administration module of the GridBank Server Architecture in the Accounting Layer. It provided account management functions such asdeposit, withdrawal, change credit limit, cancel transfers and close accounts. It also established link with PayPal system to ensure the transfer of real money to and from clients by importing data from PayPal system (Figure 3). The enhanced administration module was organized into four layers namely:

i. Accounting Layer which deals with database and account operations.

ii. Payment Protocol Layer defines payment schemes, message formats and communication protocols.

iii. Security Layer ensures that any messages passed to Payment Protocol and consequently invoking operations in Accounts Layers are authentic.

iv.PayPal Layer (System) which interfaced with the existing module handles the transfer of real money to and from clients by importing data from PayPal system. PayPal is an online payment service that allows individuals and businesses to transfer funds electronically [12].

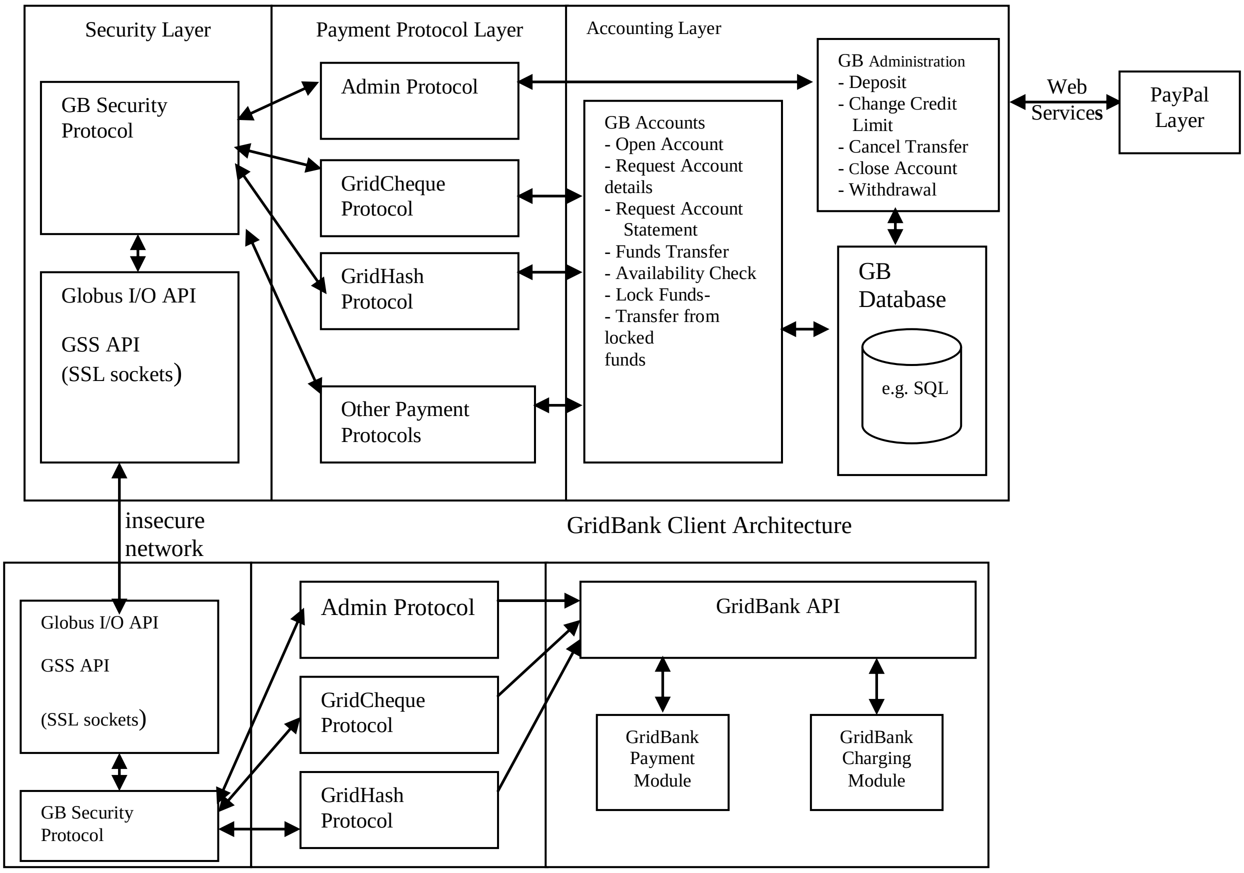


Figure 3: Enhanced scheme architecture

The enhanced scheme architecture class diagram is shown below in figure 4 in which there are two classes namely GridBank and PayPal. Both classes are associated to one or more of each other.

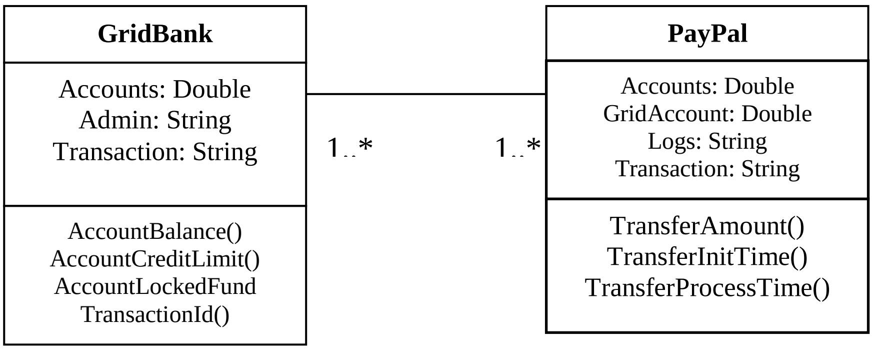


Figure 4: Class diagram of the enhanced scheme

3.1 Model Assumptions

The following assumptions were made while developing the model.

i. It was assumed that the required network bandwidth will be available for the scheme to work effectively.

ii. That the PayPal system would be robust enough to accommodate more users.

iii. That the GridBank server would be available all the time for transactions.

3.2 Mathematical and Performance Parameter Specifications

This section discussed the mathematical and performance parameter specifications for the enhanced accounting scheme.

3.2.1 Mathematical analysis

The following factors were identified as necessary in the overall analysis of the scheme

(i.) Processing Time (P_T): This is the time taken for a task to be performed. It depends on the following:
a. the available number of resources denoted

by α

b. the scheduling algorithm used denoted by β

c. the quality of service required denoted by χ

d. job priority denoted by δ

$P_T= f(\alpha, \beta, \chi, \delta)$(3.1)

Processing Time (P_T) was determined by these constituent attributes which then determine whether it would be high or low.

(ii) Users (U_N): This is the total pool of users. It depends on the following:

a. platform offered by the grid denoted by ε

b. services offered by the grid denoted by ϕ

c. billing technique denoted by φ

$U_N = f(\varepsilon, \phi, \varphi)$ (3.2)

Users (U_N) was determined by these constituent attributes which then determine the total pool of users.

(iii) Resources (R_N): These are the available computing entities. These depend on the following:

a. the size of the grid denoted by γ

b. the number of clients willing to share resources denoted by η

c. policy of the virtual organization denoted by ι

d. platform interoperability of the grid denoted by κ

e. the economy model used by the grid service provider denoted by λ

f. cost of the resource denoted by μ

g. quality of the resource denoted by ν

$R_N = f(\gamma, \eta, \iota, \kappa, \lambda, \mu, \nu)$(3.3)

Resources (R_N) were determined by these constituent attributes which then determine the available computing entities for grid users.

(iv) Load (L_N): This is the sum of all constituent loads of the users. This depends on the following:

a. job complexity denoted by ϑ

b. capability of the available resourcesdenoted by ρ

c. waiting time denoted by σ

$L_N= f(\vartheta, \rho, \sigma)$ (3.4)

Load (L_N) was determined by these constituent attributes which then determines whether the load would be heavier or lighter.

3.2.2 Performance parameter specification

The following parameters: processing delay and load scalability were used for the performance evaluation.

Processing delay

Processing delay (P_D): The processing delay in this regard referred to time delay between the instant of service request by the users to the instant of having the result of the service request. This depends on processing time (P_T)and users (U_N) which were also dependent on some factors identified in 3.3.1 above. The concern here was to ascertain how the two schemes would respond to service requirement of the users. It was expected that the processing delay should reduce with inclusion of more serviceable resources.

$P_D= f(P_T, U_N)$ (3.5)

Load scalability

A scalable system is one that can easily be altered to accommodate changes in the number of users, resources and computing entities attached to it. Load scalability (L_S) depends on Load(L_N)and Resources (R_N) which were also dependent on some factors identified in 3.3.1 above. It was expected that the enhanced scheme would be able to expand and contract its resource pool to accommodate heavier or lighter loads. This would be considered during three different periods; peak, off-peak and holiday.

$L_S= f(L_N, R_N)$ (3.6)

4. MODELING AND SIMULATING THE SCHEMES

The previous GridBank Accounting System and the enhanced one which incorporated PayPal system were modeled and simulated using a visual modeler[13] for analyzing variation in load and resources available on them in order to carry out performance evaluation of the schemes.

The previous GridBank Accounting System, which had three layers, was modeled as having three resources while the enhanced one was modeled as having four resources because of the additional PayPal layer.

Supplying these values of resources for the existing and the enhanced systems into the modeler, resultswere generated for the two simulated schemes. The outputs generated were the parameters used for the performance evaluation in the next section.

5. SIMULATIONRESULTS AND DISCUSSION

The outputs generated by the modeler were used to analyze the performance of the schemes considering processing delay and load scalability.

5.1 Processing Delay

The analysis for the Processing delay was carried out using 5 (five) different users on the visual modeler in order to ascertain how the existing and enhanced GridBank would respond.

The existing scheme had fewer resources (3) as compared to the enhanced scheme (4). As a result, the processing delay in the existing increases and drops from time to time because of less number of resources available and also because of the variations in demand of resources. Some resources were oversubscribed while some were undersubscribed. The gap between the existing and the enhanced schemes at the initial stage was as a result of oversubscription for some resources.The addition of more resources in the enhanced scheme resulted in high demand for service but the scheme gracefully sustained the demand and this was evident from the gradual drop in the processing time(See Figure 5).

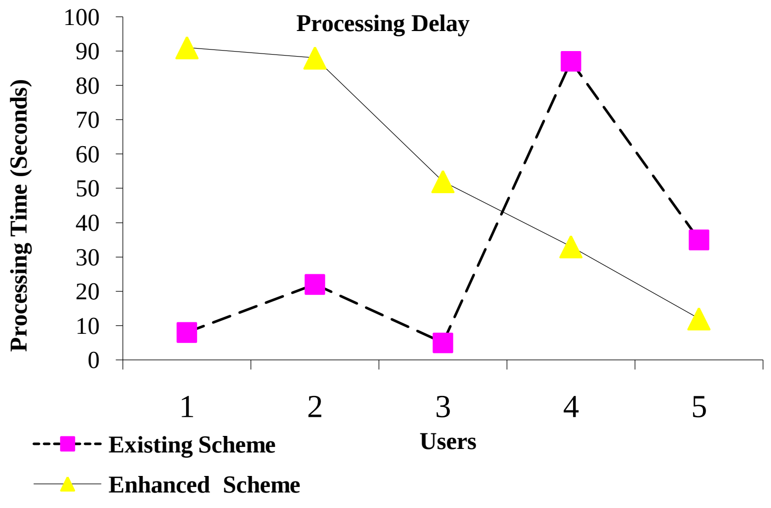


Figure5: Processing delay for the existing and enhanced schemes

5.2 Load Scalability

The load scalability was considered during 3 different periods namely: peak period, off peak period and holiday period.

5.2.1 Load scalability: peak period

The load was getting reduced and there was an increase in load diversity factorwith addition of more resources. The maximum load in the existing scheme was 0.8 and minimum was 0.4 with less resources. While the maximum in the enhanced scheme was 0.9 and minimum was 0.2. All these were noted during the peak period analysis.

Figure 6on the next page showed the relationship between the load and number of resources available on both the existing and enhanced schemes.

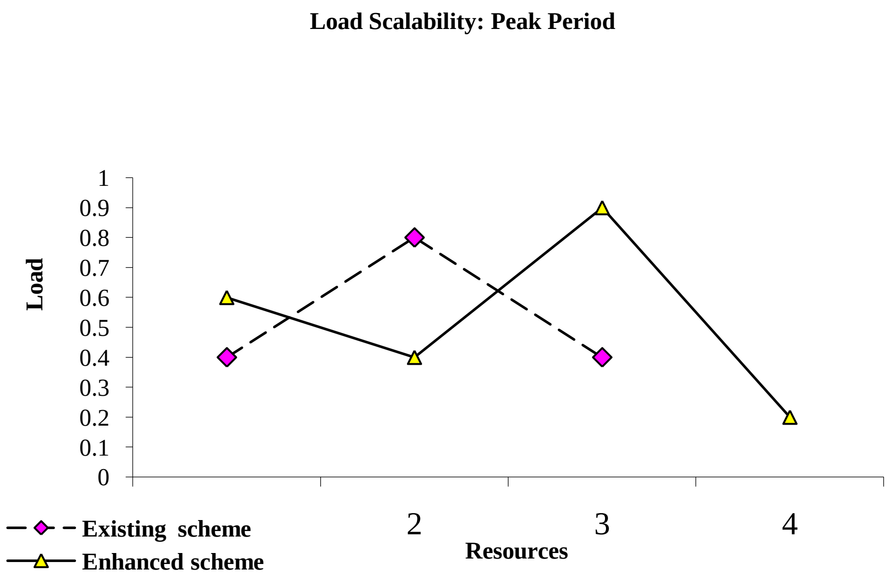


Figure 6: Load scalability for the existing and enhanced schemes during peak period

5.2.2 Load scalability: off peak period

During the off peak period, the existing and enhanced schemes had the same maximum load of 0.9. This resulted from high demand for a particular resource during this period. But the minimum for existing scheme was 0.5 with lessresources while the minimum was 0.0 for enhanced scheme. In the enhanced scheme the load demand for resources was graciously reduced because of more resources on it.This also increased load diversity factor of the system and hence improved its efficiency.

Figure 7below showed the relationship between the load and number of resources available on both the existing and enhanced schemes during this period.

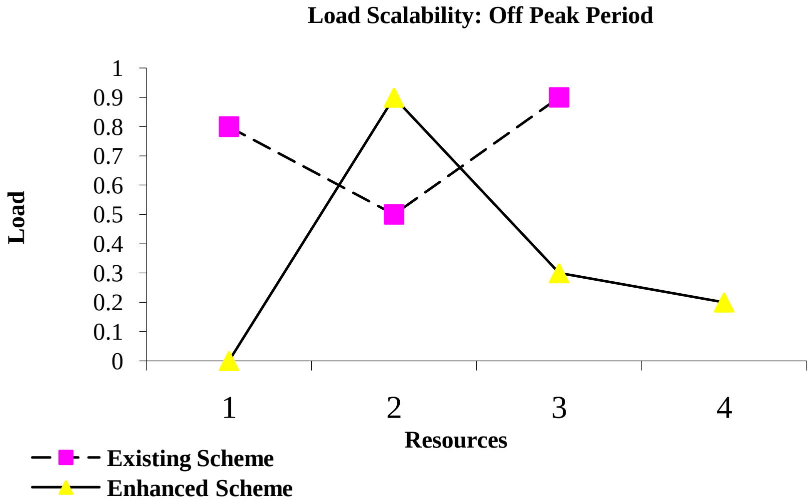


Figure 7: Load scalability for the existing and enhancedscheme during off peak period

5.2.3 Load scalability: holiday period

During the holiday period, the maximum load for the existing scheme was 1 and enhanced scheme had the same value. This resulted from high demand for a particular resource during this period. But the minimum for existing scheme was 0.4 with fewer resources while the minimum was 0.3 for the enhanced scheme. The inclusion of more resource in the enhanced scheme brought about the difference in load capacity of the two schemes. The point at which maximum load was achieved for the existing scheme was the same point at which minimum load was achieved for the enhanced scheme. This also increased the load diversity factor of the system and hence improved its efficiency.
Figure 8 showed the relationship between the load and number of resources available on both the existing and enhanced schemes.

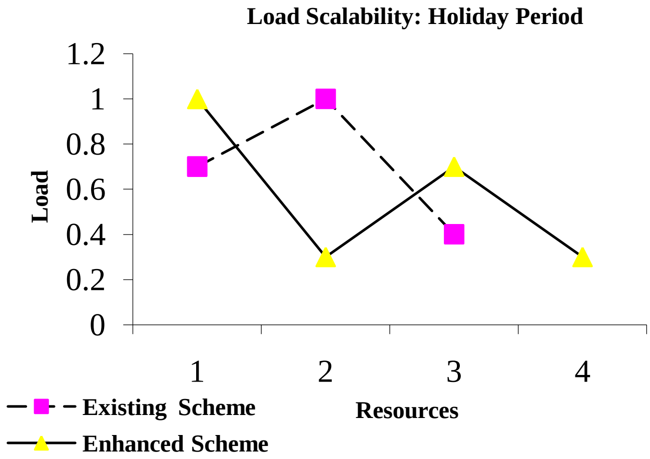


Figure 8: Load scalability for the existing and enhanced schemes during holiday period

6. THE SCHEME VALIDATION

Regression analysis tool called R-squared value was used to confirm the goodness of fit of the schemes. The existing GridBank was the existing scheme while the enhanced GridBank was referred to as enhanced scheme. The following were the observations:

6.1 Processing Delay

The fitness value generated from R squared for the curve representing the existing scheme was 0.3191while that of the enhanced scheme was 0.9598, which was very close to 1 (Table 2). The closer the value to 1 the better the curve generated[14]. This actually showed that the enhanced scheme had greatly improved on the existing scheme.

6.2 Load Scalability: Peak Period

The fitness value generated from R squared for the curve representing the existing scheme was 0.0203. The enhanced scheme had a fitness value of 0.0541, which when considered was a value closer to 1 (Table 2). The closer the value to 1 the better the curve generated. This actually showed that the enhanced scheme had improved on the existing scheme.

6.3 Load Scalability: Off-Peak Period

The fitness value generated from R squared for the curve representing the existing scheme was 0.0085. The enhanced scheme had a fitness value 0.0287, which when considered was a value closer to 1 (Table 2). The closer the value to 1 the better the curve generated. This actually showed that the enhanced scheme had improved on the existing scheme also during the off-peak period.

6.4 Load Scalability: Holiday Period

The fitness value generated from R squared for the curve representing the existing scheme was 0.02500. The enhanced scheme had a fitness value 0.4158, which when considered was a value closer to 1 (Table 2). The closer the value to 1 the better the curve generated. This actually showed that the enhanced scheme had improved on the existing scheme also during the off-peak period.

Table 2.R-Squared values for the existing scheme and enhanced scheme

Parameters	Existing Scheme	Enhanced Scheme
Processing Delay	0.3191	0.9598
Load Scalability(Peak)	0.0203	0.0541
Load Scalability(Peak)	0.0085	0.0287
Load Scalability(Holiday)	0.2500	0.4158

7. CONCLUSIONS AND RECOMMENDATIONS

7.1 Conclusions

The objectives of the research were achieved. The enhanced administration module of the GridBank was automated as well as transferring of real money to and from clients by the additional Paypal layer. This became achievable by using an approach based on web services technology.

The enhanced scheme was modeled and simulated with the simulation results showing a proof of superiority over the existing scheme. Therefore it can be concluded that the enhanced scheme would perform better than the existing scheme. Hence, the enhanced scheme is recommended for use.

7.2 Suggestions for Future Work

GridBank could be expanded in a bid to take care of multiple virtual organizations as well as cloud computing platfromswhich are just evolving so that resource or service usage can be properly accounted for. Hence, this would also create room for more research work in the areas of resource and service management. The option of having multiple GridBank could also be looked into such that service providers and users can have more available choices.

8. REFERENCES

[1] Browne, J.C. 2004. Grid Computing as applied Distributed Computation. IEEE International Symposium on Cluster Computing and the Grid.

[2] Canal, P.2011 Gratia: New Challenges in Grid Accounting. Journal of Physics: Conference Series 331doi:10.1088/1742-6596/331/6/062028

[3] Dudy, L., Quoc, T.H., Unwei, Z., Bu, S. L. and Yew S.O.2005 MOGAS, A Multi-Organizational Grid Accounting System. International Journal of Information Technology, 11: 84-103.

[4] Barmouta, A., and Buyya, R. 2003. 'GridBank: A Grid Accounting Service Architecture (GASA) for Distributed Systems Sharing and Integration'. Proceedings of the 17th Annual International Parallel and Distributed Processing Symposium (IPDPS 2003), IEEE Computer Society Press, USA.

[5] Beardsmore, A., Hartley, K., Hawkins, S., Laws, S., Magowan, J. and Twigg, A. 2002. 'GSAX Grid Service Accounting Extensions', Global Grid Forum Technical Report.

[6] Brenner, M., Wiebelitz, J. and Smith, M.2010 A JEE-based Architecture for Distributed Multi-Domain Resource Accounting Proceedings of the Sixth IEEE International Conference on e-Science Workshops IEEE Computer Society Washington, DC, USA

[7] Ford, M. 2006. NGS Resource Usage Service Manchester Computing. The University Manchester USA.

[8] Elmroth, E., Gardfjall, P., Mulmo, O.,Sandgren, A. and Sandholm, T. 2003. "A Coordinated Accounting Solution for SweGrid". Swedish National Allocation Committee (SNAC) Technical paper

[9] Byrom, R., Cordenonsi, R., Cornwall, L., Craig, M., Djaoui, A., Duncan,A., Fisher,S., Gordon, J., Hicks, S., Kant, D., Leake, J., Middleton, R., Thorpe, M., Walk, J., and Wilson, A. 2005. 'APEL: An implementation of Grid accounting using R-GMA'. Proceedings of the 2005 UK e-Science AllHands Meeting.

[10] Piro, R. 2005. DGAS: DataGrid Accounting System. Enabling Grids for E-science in Europe Workshop on e-Infrastructures May 13 2005 Amsterdam

[11] Refsnes Data (2008). Introduction to Web Services.An online Tutorial. [online] http://www.w3schools.com/ webservices/default.asp. Visited on January 22, 2008

[12] Grabianowski, E.andCrwaford, S.2012. How PayPal Works. How Stuff Work Article.HowStuffWorks Inc.[online] http://money.howstuffworks.com/paypal7.htm Visited January 17, 2012

[13] Sulistio, A. Yeo, C. S. and Buyya, R. 2003. VisualModeler for Grid Modellingand Simulation (GridSim) Toolkit.InInnovative Grid Computing Workshop, Proceedings ofthe 3rd International Conference on Computational Science (ICCS 2003).

[14] Weiss, N.A.2002.Introductory Statistics, 6th ed. Addison-Wesley, USA.