Computational Resource Optimization in Ugandan Tertiary Institutions
Richard Ssekibuule, Joel Lakuma and John Ngubiri
Faculty of Computing and Information Technology
Makerere University
Email: {rkayondo, jlakuma, ngubiri}@cit.mak.ac.ug

Abstract. Insufficient computational power, caused by limited budgets is a big challenge to high quality research and education in developing countries. Some institutions suffer acute shortages while others enjoy comparatively more resources. Many institutions therefore source for funds to procure more resources while paying less attention to optimally utilizing what they have or creating facilities for sharing with other institutions for mutual benefit. Current developments in Grid computing technology provide for mechanisms of consolidating isolated computing resources to provide a high power computing environment. In this project, we aim at studying ways (limited) computer resources can be optimally used in a financially constrained setting as well as ways for building an environment for providing high computing power for learning and research in a cost effective way.

Keywords: Computational-power, grid computing, resources

I. INTRODUCTION

Information and Communication Technology (ICT) has revolutionalized ways in which businesses and day to day activities are run. This has been both at the internal unit/business operation as well as business internetworking to cater for resource sharing, communication, cooperation and other business needs. Graduates from tertiary institutions are expected to have adequate ICT knowledge related to the field of specialization and are as well expected to be adequately trained in ICT facilitated institutions. ICT is therefore both a medium and goal of tertiary training.

In financially constrained countries like Uganda (and other developing countries), there is hardly a tertiary institution with adequate ICT infrastructure to meet its demands throughout the year. Some institutions and individuals have idle resources during off peak seasons for their operations in the year, and also experience bottlenecks during peak periods, yet other institutions experience shortages throughout the year. This is caused by lack of money to procure adequate ICT infrastructure (and train the human resource). This has a high potential of creating a vicious circle of inadequate training of the students and hence not exploring their full potential at the workplace. This leads to lower business productivity, low income, and recurrence of poverty. This vicious circle needs to be broken at the earliest opportunity to stimulate development and transform developing countries into developed countries.

To address the scarcity of ICT infrastructure, tertiary institutions in Uganda have mainly put emphasis on sourcing for funds to procure more infrastructures to reduce the deficit. Less emphasis however has been put at optimal utilization of the existing ones, identifying specifications that can offer higher value for money in a certain situation or exploring inter-institutional resource sharing for mutual benefit.

In this project, we aim at:- (i) Investigating ways ICT facilities can be optimally utilized both within and among tertiary institutions in Uganda; (ii) Investigating metrics institutions can use to procure ICT resources in a cost effective way in specific work and institutional environments and (iii) investigating ways different organizations can share ICT resources in a developing country environment. Overall, the research aims at improving the overall benefit per unit cost as means of increasing the relative ICT prevalence in tertiary institutions.

The rest of the paper is organized as follows; we present the motivation behind the inception of this project in Section II and work done related to it in Section III. We discuss the developing country specific challenges this project is likely to encounter in Section IV. These challenges to a large extent make this project unique from related projects in the developing countries. We justify the project in Section VII and describe the overview of the research blocks in this project and make our conclusions in Section VIII.

II. PROJECT MOTIVATION
Despite the general fall in prices of ICT equipment and resources, many tertiary institutions in developing countries cannot afford them in adequate amounts. In fact, some resources like bandwidth and computer hardware are more expensive in developing countries than developed countries. While sourcing for extra funds can reduce the scarcity, we believe devising ways of using new and existing technologies that put the resources to optimal utilization can also create a substantial impact. It also generates more utility from the funds mobilized by the institutions. The following are the main factors that motivated the inception of this project.

A. Low utilization levels:

The individual utilization of a micro computer is very low with resources like CPU usage hardly reaching 15% for average computer users. We need to investigate ways in which we can get optimal utilization of computer systems both inside individual tertiary institutions and across institutions and individuals in the region.

B. Cost and utility mismatch:

Application and operating system software is so dynamic (with windows changing nearly every two years and linux far less). Within a few years, the software is considered outdated. In many cases, a software upgrade is not effective since the hardware may not fit the minimum requirements of the upgraded software version. In most cases an operating system or application software upgrade might as well necessitate a hardware upgrade or a total replacement. A long lasting and hence expensive computer may not be the best option in a financially constrained environment if the organization is to keep abreast with the software regime.

C. Software appropriateness:

In many cases, a computer that cannot work with specific software can work with another (especially less graphical software). This is because of the different resource requirements by different platforms. The cost of replacing the computers may be higher than the cost of acquiring a resource conserving software. Likewise, the software may have a lower rate of change in resource requirements that it can be used for a longer time on the same hardware system than another system whose resource requirement change more often. Additionally, computer hardware that could be considered ‘old’ for some platforms could be competent enough to support other software systems with lower system requirements. Old computers can therefore be reused and hence increase their utility.

D. Resource sharing:

With the existence of the Internet, Intranets and some commercial and private optical fiber networks, remote/ grid computing can be a possible cost saving option. Different institutions can share or combine otherwise expensive resources like software and processors to achieve high computational power. Technological advancements in Grid computing [1] around the world provide a platform for a probable solution to large-scale resource sharing [2]. The Grid architecture can provide a means for combining low capacity inexpensive processors into a distributed mass of high capacity computational infrastructure.

E. Network for collaboration:

Institutions participating in resource sharing can easily enhance collaboration in their production or research areas that are within or outside the scope of this project for their mutual benefit.

III. RELATED WORK

The work that we wish to carry out in this project builds on a foundation that has been created else where most especially in Europe. Several projects in Europe have been successfully implemented to facilitate optimal computation power utilization and research. In this section we present projects whose achievements and ambitions are similar to the ones in this project proposal:
A. European Grid Computing

Through a grant from the European Union, an application test-bed for Grid computing [3] was implemented to provide a network of high performance computing within European counties. Additionally, the European Grid is used to provide support and a framework for developing Grid software among many other software and related scientific research problems.

B. Albertross

The Wide Area Cluster Computing project Albatross [4] was implemented in The Netherlands among four collaborating universities to help them investigate and understand application behavior on wide-area networks. A host of cluster computing research problems is being investigated in this project and a host of research publications have been realized as a result of the presence of this collaborative facility.

C. Particle physics data grid

The particle physics collaborative project [5] is used by physicists and grid computing scientists for investigate high energy problems and network infrastructures. This project has served needs for experimental physics and research requirements for computer scientists.

IV. IT INFRASTRUCTURAL CHALLENGES

Unlike most of the related work carried out, this project is taking place in a technologically and financially different setting and hence faces somewhat unique challenges in addition to those faced by related projects. These challenges may therefore call for different approaches and also render results from some related research inapplicable. While some make what would other wise be obvious assumptions stringent constraints, they highly contribute to the originality and vitality of the research project. Below, we briefly explain the nature and severity of the (potentially unique) challenges on this research.

A. Low Inter-network speeds:

Like in other developing countries, Ugandan institutions of learning and research are mostly connected to low speed Internet links. This may be a hindrance to remote processing especially when large amounts of data are to be transferred. It is also expensive for developing countries to upgrade speeds in network links.

B. Low availability of computer hardware and software:

Though the world price of computers is falling, it is still out of reach for many organizations and persons in Uganda. Considerations of the scarcity of money to procure computers and the scarcity of computers themselves in the organizations have to be put into consideration. Additionally, the total cost of ownership (TCO) remains high for Ugandan institutions, yet it is rarely taken into consideration by these institutions when procuring systems.

C. Quest for high computational power in new applications:

High education demand in Uganda is increasing; this is coupled with research some of which needs high computational power (despite fewer computers). The need for high computational power in the day today running of the institutions will be highlighted in the research.

V. OPPORTUNITIES FOR RESEARCH

We now discuss the items that are to make the units of this project. Just like in other research projects, new research ideas can come up leading to the deepening of the individual unit or creating another unit of research. Therefore we cannot guarantee the conclusiveness of the items just like we cannot guarantee the conclusiveness of the depth of each.
A. Utility maximization

Here, we shall seek to investigate how resource on a single computer for example can be put to maximum utilization. This can involve development of multi-user workstations in a local setting. It also involves investigating the optimal number of users that can run on a computer (as a function of time, application, user characteristics, etc).

B. Resource Management and brokerage

This is to be in inter-unit (with in an organization) and inter-institutional sharing of resources. This is to ensure optimal resource usage, incorporation of owner policies in a shared environment, investigating ways multi agent can be used in a (micro/ mini) grid setting with in the constraints of a slow communication network, heterogeneous environment and relatively slower machines. We are also to investigate (multi) programming issues.

C. Security

Sharing and cooperative processing extends the computers to more threats to attacks and violation of privacy. Sharing and coordinating use of computational recourses opens up new security challenges that can be grouped into three categories: integration with existing systems and technologies in distribute virtual organizations [2], interoperability with different hosting environments (e.g., J2EE servers, .NET servers, Linux systems), and trust relationships among interacting hosting environments also identified by Nagaratnam et al. [6]. This project will give researchers in Uganda and others involved and opportunity to have hands-on research for the security challenges involved in coordination with experts in the same field.

D. Robustness

Factors like network failure, server failure or node failure cannot be ruled out. This may render the reliability of such a cooperation set up poor. To achieve user satisfaction, the system has to be reliable. Studying and developing techniques to make the system more robust/reliable can be issues for upcoming research.

E. Additional system support

This is to involve development of middleware that can help in management of such cooperation network in the developing country set up.

VI. JUSTIFICATION

Other than the increased availability of computational resources for training and research, there are other fields/ organizations that can tap from this project to improve their efficiency. Other areas of research in Uganda that could tap into the computational abilities of this research work are:-

A. Support for Epidemic Research

Uganda is one of the key participants in Virus research and infectious diseases. Institutions that require this computation power already exist (like the Uganda Virus Research Institute, Join Clinical Research Centre) and their results and Mean Time to Deliver Results (MTDR) will be reduced by existence and utilization of a high computational infrastructure.

B. Support for other High Performance Applications

Current and future research centers in Uganda (will) need high but cheap computational power. These include weather prediction centers, disaster preparedness, and molecular biology research among others. Foster et al. [7] and Czajkowski et al. [8] provide examples to some of the services that can be developed to benefit from distributed resource sharing.
C. Distributed data management

In Uganda, like in many countries where there is decentralized governance, there is distributed data on a similar subject with remote data often needed for supervision and comparative purposes. This research will as well look at distributed data management which can be beneficial to such structures.

VII. PROJECT STRUCTURE

In this section, we present in chronological order the general steps that will be taken to realize successful implementation of this project.

A. Feasibility Study

In this phase of the project, we shall investigate the current state of computer availability, utilization trends, and connectivity resources and systems versioning among others. The researchers will establish the computational resources on the ground, numbers, specifications, institutional needs, gaps, bridging strategies, strengths, weaknesses and other related factors. By the end of this Phase, the researchers are expected to have information on:-

i. The current state of computational resources;
ii. The requirements of different institutions that will be expected to participate in the high power computing facility;
iii. The relationship between the ideal, reasonable and existing IT infrastructural levels;
iv. The extent to which IT has been incorporated in the management, administration and carrying out of business in tertiary institutions in Uganda;
v. The existing opportunities in IT resource optimization in Ugandan tertiary institutions and
vi. The possible threats that may be created as a result of consolidating computing resources will be investigated.

B. Main Research

The ability for this project to support and trigger a spin-off of research problems and experiences brings us to the discussion the key research components. The main research is divided into high level and low level research. High level research is expected to deal with the organizational, policy and administrational issues while low level research will deal with the software development, performance and resource optimization techniques.

1) High level research: The themes to be considered in high level research include:
   i. Policy and Regulation issues;
      Policies that will be used to regulate usage and distribution of computing resources will be done as high level research.
   ii. Specifications for middlewares;
      This activity involves specification of the types of middleware and needs that are to be served.
   iii. Administrative & support frameworks;
      This level of research in the project will define how administrative work for technological facilities will be supported.

2) Low level research: The themes to be included in low level research include:
   i. Middleware development / Deployment; Middleware in a grid environment is supposed to provide abstraction of the underlying technical details of the hardware and operating systems from grid applications. Middleware development and deployment is key to successfully realization of grid services. InteGrade [9] is a good example of middleware that leverages idle computing power for desktop machines. Grid middleware has also been used to provide transparent migration of non-grid applications to grid environments with applaudable success [10].
ii. Resource scheduling;

Resources in a Grid system are distributed geographically in small components, which are consolidated into a single pool of large consolidated base. This pool of resources is then shared among formal or informal groups of individuals and organisations that are commonly known as virtual organisations [8], [2]. Resources in such distributed network systems need to be well managed in order to have efficient and effective support for demand and resource utilization with economies of scale being put into consideration as illustrated in the research work done by Buyya et al. [11]. The high power computing facility will provide and opportunity for researching scheduling challenges on low speed networks. Little work has been done on low speed networks in a high computing facility, particularly because most high power computing systems are in industrialized countries where bandwidth is not a big problem like in developing countries, especially in Africa.

iii. Transaction & deadlock management;

In a high power computing facility where jobs are being submitted by several users and organisations that may want results in the shortest time possible, deadlocks and transaction bottle necks are most likely to occur. Transaction and deadlock management will definitely be an interesting research problem for academicians who will be involved in the project. Transaction and deadlock management has been investigated by several researchers [12], [13]. We need to investigate and evaluate these solutions for wide area cluster of heterogeneous systems and probably improve on them to suite any unique challenges that we shall be facing.

iv. N-tier service brokerage;

Most distributed applications have been built on multi-tier technology in order to support scalability in service resources and provision of simplified administrative tasks. Research work by Shan et al. [14] presents recent technologies like logical solution architecture, process choreography, business rule engine, enterprise service bus and service composition among others in service oriented solution architectures for mainly web applications. These technologies can also be investigated for application in the grid systems. Research in N-tier service brokerage will be useful for leveraging high power computing resources for grid applications.

v. Quality of service on communication channels;

For the kind of networks that exist in Uganda and the East and Central African region, quality of service research really needs to be done whilst paying attention to the low speed networks. Consolidation of idle CPU over low speed networks and probably accessing these resources from a few high speed fiber links that exist in the country is expected to give us interesting challenges for research.

VIII. CONCLUSION

The wide attention that has been given to Grid computing by researchers around the world provides a stimulating opportunity for scientists in Uganda to actively participate and contribute to the growing technology. We hope that by bringing the high power computing facilities near researchers in the country at relatively affordable financial and administrative costs, research in the region will be enhanced. This project will be a brave step towards bridging the technological gap between a developing country Uganda and the industrialized economies. The successful implementation of an inter-institutional Grid will provide computer scientist with an opportunity for hands on research. Other scientists outside computer science will have an opportunity to research in fields that were previously impossible without a high power computing resource.

REFERENCES


AUTHORS

Richard Ssekibuule
Richard is an Assistant Lecturer in the Department of Computer Science, Faculty of Computing and IT, Makerere University. He holds a BSc (Computer Science, Mathematics) from Makerere University Kampala, and an MSc (Computer Science) from Radboud University Nijmegen. His research interests are in the area of security in open distributed computing.

Joel Lakuma
Joel is an Assistant Lecturer in the Department of Computer Science, Faculty of Computing and Information Technology, Makerere University. He holds a BSc (Telecom Engineering) from Trinity University, Terra Haute, IA, USA; and an MSc (Computer Science) from Makerere University, Kampala, Uganda. His research interests are in Algorithms, and Computing Complexity.

John Ngubiri
John is an Assistant Lecturer in the Department of Computer Science, Faculty of Computing and Information Technology, Makerere University. He holds a BSc (Maths, Physics) with Education, a Postgraduate Diploma in Computer Science and a Master of Science in Computer Science (All from Makerere University). He is currently undertaking doctoral studies at Radboud University in The Netherlands. His current research interest is in Resource management and scheduling in supercomputers.