Applications of Grid Computing in Agriculture: An Indian Scenario

Xitij U. Shukla[1], P.S. Parsania[2] and Krunal C. Kamani[3]

1 Assistant Professor, B.A. College of Agriculture, AAU, Anand – 388 110
2 Assistant Professor, College of Food Processing Technology and Bio Energy, AAU, Anand – 388 110
3 Assistant Professor, S.M.C. College of Dairy Science, AAU, Anand - 388110
Email: xus@aau.in

ABSTRACT

Grid computing is a recent phenomenon associated with coordination and federation of computing resources distributed across organizations. The architectural style, components, present status and potential applications of Grid Computing in agricultural research in India is surveyed here.

Keywords: *Grid Computing*

INTRODUCTION

Networked computers demonstrated a novel approach towards communication and promised an unfathomed horizon of endless applications in the late 1960's following the cut-throat competition to achieve excellence in satellite and space technologies in the either super-power countries. Supported by United States Department of Defense's ARPA (Advance Research Project Agency), ARPANET became the first operational infrastructure for networked computers. (J.C.R. Licklider And The Universal Network, 2000)

Dr. Leonard Kleinrock – one of the major contributors to the research of this network had envisioned the utility of networked computers in 1969, as "As of now, computer networks are still in their infancy … but as they grow more sophisticated, we will probably see the spread of 'computer utilities', which, like present electric and telephone utilities will service individual homes and offices across the country" (UCLA to be first station in nationwide computer network, 1969). This might have seemed as a mere dream at those times though it has become a solid reality in present era. ARPANET eventually lead the scientific community towards the Internet, upon development of the TCP/IP protocol suite (Transmission Control Protocol / Internet Protocol) – a universal specification and means for assimilation of data across the heterogeneous network.

Figure 1 : Grid Computing

Image Courtesy (http://www.adarshpatil.com/)

Grid Computing

CPU (Central Processing Unit), RAM, storage space, network bandwidth, software etc. are classified as computational resources. Tasks solved by computers require some or all of these computational resources depending upon the complexity of the tasks. If the task is more complex it requires higher computational resources and so on.

Many research oriented activities such as weather forecast, climate change modeling, earthquake simulation, aircraft designing, protein mapping, data analysis etc. require

higher computational power. If they are solved on a single computer, the computation may take more time. If they are distributed on multiple computers and computed collectively, the time taken will be relatively less in comparison to the former approach.

Supercomputers consisting of thousands of CPUs, higher RAM and massive storage space are used to undertake such tasks.

However, it is not always possible to utilize such dedicated supercomputers on account of very high initial cost, complexity in operation and recurring maintenance for organizations specifically academic institutes involved in such research. Often multiple organizations are working on a common research goal. In such scenario some organizations have one type of computational resources and some have the other.

Grid Computing bridges this gap of resource scarcity among these organizations, through federation and sharing of these resources across. Grid Computing is defined as a phenomenon "That coordinates resources that are not subjected to central control, built using standard, open and general purpose protocols and interfaces and that delivers non-trivial quality of service". (Figure 1) (Foster I. , 2002)

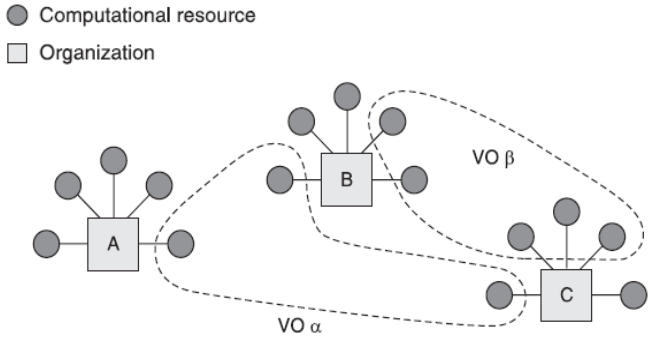

Figure 2 : Virtual Organizations

Image Courtesy Borja Sotomayor

Grid computing introduces the concept of virtual organizations, a boundary transparent architecture for flexible, dynamic, secure, coordinated sharing of resources. (Figure 2) (Foster I. K., 2001).

Grid computing infrastructure is built over *Web Services* – an operating system and programming language neutral distributed computing framework over XML (eXtensible Markup Language) based technologies. The commonly used technologies are WSDL (Web Services Description Language), SOAP (Simple Object Access Protocol) and UDDI (Universal Description, Discovery and Integration). (Ian Foster, 2002) (Borja Sotomayor, 2006)

XML provides a framework skeleton for exchange of messages by web services. SOAP lets the hosts communicate via XML across the network by transporting messages over HTTP. Web service's interface e.g. types of operations, message format etc. are defined by WSDL (Web Service Description Language). UDDI provides a registry service to announce the availability of a web service.

Role of NABG

Under ICAR's (Indian Council of Agricultural Research) NAIP (National Agricultural Innovation Project), component 1 NABG (National Agricultural Bioinformatics Grid) Grid has been set up within organizations such as National Bureau of Plant Genetic Resources (NBPGR)– New Delhi, National Bureau of Animal Genetic Resources (NBAGR) – Karnal, National Bureau of Fish Genetic Resources (NBFGR) – Lucknow, National Bureau of Agriculturally Important Microorganisms – Kusmaur, National Bureau of Agriculturally Important Insectsc(NBAII) – Bangluru. (NABG)

These organizations are acting as partner organizations, under the lead centre Indian Agricultural Statistics Research Institute – New Delhi. The major objectives of this initiative are to create local databases and Bioinformatics Data Warehouse (BinDW) for genomic resources across species and to promote inter-disciplinary research groups with focus on agricultural bioinformatics.

The partner organizations are involved in collection, compilation, validation and storage of the genomic data of their respective domain, followed by identification of bioinformatics related issues in them.

The lead centre facilitates a bioinformatics data warehouse (BinDW) through the data provided by the partner organizations and a mechanism for effective storage and retrieval of the same. Further, it also provides a web based information system on biological resources in agriculture – a framework for analysis of genomic data to the partner organizations requiring higher amount of computational power. (Figure 3)

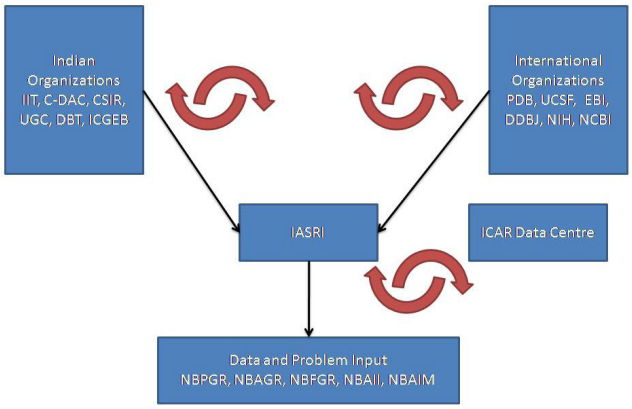

Figure 3

Through Grid computing, NABG shall not only provide effective platform for research but also an opportunity to interact with different people associated with different domain and a platform for holistic capacity building of the human resource.

CONCLUSION

Grid computing can provide a collaborative approach to problem solving by letting associated organizations utilize each other's resources. Resources are scarce at sometimes and costly at some other and are not always possible to acquire, install, own and maintain for every other individual/organization. Through Grid computing, associated organizations can collectively mitigate this issue, work in a more collaborative way and at the same time reduce their total cost of ownership through sharing their resources.

REFERENCES

Borja Sotomayor, (2006) *Globus Toolkit 4, Programming Java Web Services.* Morgan Kaufmann Publishers.

Foster, I. K. (2001). The anatomy of the Grid: Enabling scalable. *International Journal of Supercomputer Applications , 15* (3).

Foster, I. (2002). What is Grid? : A Three Point Checklist.

How does the Internet work? (n.d.). Retrieved from World Wide Web Consortium: http://www.w3.org/wiki/How_does_the_Internet_work

Ian Foster, C. K. (2002). *The Physiology of Grid.*

J.C.R. Licklider And The Universal Network. (2000).

NABG. (n.d.). Retrieved from http://nabg.iasri.res.in

Requirements for the Internet Hosts (1989)- *Communication Layers.* IETF.

Roy Fielding, J. M.-L. (1996). *HTTP/1.1 (RFC2616).* Network Working Group, Internet Engineering Task Force.

Tanenbaum, A. S. (2003). *Computer Networks.* Prentice Hall.

UCLA to be first station in nationwide computer network. (1969).

Received : May 2014 : Accepted : October 2014