Design & Evaluation For Conversion from Dv to Mpeg4 Video in Grid Environment

Er. Meenakshi Gupta
Asst. Professor (CSE)
RIMT Institute of Engineering & Technology
Mandi Gobindgarh, India
mns.meenakshi@gmail.com

Er. Sugandha Sharma
Asst. Professor (CSE)
Chandigarh University
Gharuan, India
sugandha.ss1@gmail.com

Abstract:

The term, grid computing, has become one of the latest buzz words in the IT industry Grid computing is an innovative approach that leverages existing IT infrastructure to optimize computer resources and manage data and computing workloads. The first objective intends to study the grid computing concepts, their types, their relationship with other computing technologies, the open source middleware available for its implementation etc. It also involves the analysis, design and implementation of a grid environment.The second objective intends to study the compute, storage and network intensive problem of DV to MPEG4 video conversion and to implement it on the grid environment designed in above step.Third objective intends to evaluate the performance of grid (designed in step 1) by executing the conversion process (as defined by step 2) with different parameters.

Keywords: Grid Computing, DV, MPEG4, Video Conversion

1. INTRODUCTION

Distributed computing deals with hardware and software systems containing more than one processing element orstorage element, concurrent processes, or multiple programs, running under a loosely or tightly controlled regime. Various hardwareand software architectures are used for distributed computing. At a lower level, it’s necessary to interconnect multiple CPUs with some sort of network, regardless of whether that networkis printed onto acircuit board ormade up of loosely-coupleddevices and cables. At a higher level, it is necessary to interconnect processes running on those CPUs with some sort of communication system.

Distributed computing implements a kind of concurrency. It interrelates tightly with concurrent programming so much that they aresometimes not taught as distinct subjects. [1] The popularization of broadband network and thedevelopment of MPEG-4 compression technologyurge to the multiplexing development of Internet information. Among these techniques, Video-on-Demandis themost popularserviceusedby universitiestodeliver thelecture and research content to remote location on the globe. However, it takes extremely long compression time to convert audio and video data that is the Digital Video also popularity knows as DV standard format into MPEG-4 format which is the new compression standard used on internet and mobile devices. Although MPEG-4 enhances the compression ratio, it still needs massive storage equipment to deposit theaudio and video data. The price of MPEG-4 related hardware equipment still staysat a high level currently.Thus, these problems can easily be solved by using the Grid computing technology, or PC Grid. In this thesis, we use the Linux PC grid made on the middleware of Globus Toolkit4.0.1opensource to achieve the high performance video conversion grid for very large files of camcorder format (DV format) to a very good compression format MPEG-4. Moreover, we use desktop system to make a grid computing technology to make it more convenient and realistic environment for the compression. In video conversion aspect, We uses software tool called “mencoder” to perform the parallel video conversion, with the goal to achieve the best execution time, by enabling that each node to perform in its best processing potency. We want to propose a grid based technique by using some open source based tool which can reduce the conversion time of the compression from DVtoMPEG-4 format, so that more information can reach at remote location in the globe by better conversion and compression infrastructure.

2. BACKGROUND

DV is a format for recording and playing back digital video. It was launched in 1995 with joint efforts of leading producers of video camera recorders. [2] The Moving Picture Experts Group (MPEG) is a working group of experts that was formed by ISO and IEC to set standards for audio and video compression and transmission.[3] It was established in 1988 by the initiative of Hiroshi Yasuda (Nippon Telegraph and Telephone) and Leonardo Chiariglione,[4] group Chair since its inception. The first MPEG meeting was in May 1988 in Ottawa, Canada.

[5][6][7] As of late 2005, MPEG has grown to include approximately 350 members per meeting from various industries, universities, and research institutions. MPEG's official designation is ISO/IEC JTC1/SC29 WG11 – Coding of moving pictures and audio (ISO/IEC Joint Technical Committee 1, Subcommittee 29, Working Group 11)Coding of audio-visual objects. (ISO/IEC 14496) MPEG-4 uses further coding tools with additional complexity to achieve higher compression factors than MPEG-2.[8] In addition to more efficient coding of video, MPEG-4 moves closer to computer graphics applications. In more complex profiles, the MPEG-4 decoder effectively becomes a rendering processor and the compressed bit stream describes three-dimensional shapes and surface texture.[8] MPEG-4 supports Intellectual Property Management and Protection (IPMP), which provides the facility to use proprietary technologies to manage and protect content like digital rights management.[9] It also supports MPEG-J, a fully programmatic solution for creation of custom interactive multimedia applications (Java application environment with a Java API) and many other features.[10][11][12] Several new higher-efficiency video standards (newer than MPEG-2 Video) are included, notably.

3. PROBLEM FORMULATION & OBJECTIVE

It is clear from literature that computing environments have evolved from single-user environments to Massively Parallel Processors (MPPs), clusters of workstations and distributed systems to (most recently) grid computing systems. Every transition has been a revolution, allowing scientists and engineers to solve complex problems and sophisticated applications previously in capable of solving. However every transition has brought new challenges and problems in its wake, as well as the need for technical innovation. The evolution of computing systems has led to the current situation in which millions of machines are interconnected via the Internet with various hardware and software configurations, capabilities, connection topologies, accesspolicies and so forth. The formidable mix of hardware and software resources on the Internet has fuelled researchers' interest in investigating novel ways to exploit this abundant pool of resources in an economical and efficient manner, as well as in aggregating these distributed resources so as to benefit a single application.

4. OBJECTIVES

The major objectives are:

1. To study and implement the Grid Environment.
2. To design and implement DV to MPEG 4 conversion process on Grid Environment.
3. To evaluate the performance of Grid Environment for DV to MPEG 4 conversion with respect to time.

The first objective intends to study the grid computing concepts, their types, their relationship with other computing technologies, the open source middleware available for its implementation etc. It also involves the analysis, design and implementation of a grid environment.

The second objective intends to study the compute, storage and network intensive problem of DV to MPEG4 video conversion and to implement it on the grid environment designed in above step.

Third objective intends to evaluate the performance of grid (designed in step 1) by executing the conversion process (as defined by step 2) with different parameters.

5. IMPLEMENTATION AND CONFIGURATION SETUP

It this Grid Environment, Grid client (NodeA) connecting to a grid that appears as a "cluster" of Torque (PBS) job managed machines represented by NodeB and a "cluster" of Sun Grid Engine (SGE) job managed machines represented by NodeC to NodeH. The Globus toolkit is the middleware of the Grid environment.In the Grid Environment we make sure that installation instructions get us to the starting Design assumptions (Figure 1):

1. 12 machines running Fedora Core 4 i386 (FC4) with Globus toolkit 4.0.1 installed to form a grid environment.
2. A working CA Authentication Server to authenticate the grid hosts and user to use the grid resources.
3. A NTP (Network time Protocol) server to synchronize the grid node to a uniform time standard. Each machine should be running NTP client service to synchronization grid machine
4. To establish a network on an existing network, each machine should be configured with network address and Domain name and resolution method. This was accomplished by configuring the IP address, Domain Name and editing the /etc/hosts file.
5. RPM (RedHat Package Manager) configured so that administrators are able to install, update, and remove packages using RPM.
6. Root access to all the machines is necessary in order to complete some of the installation and for other administrative tasks
7. Make sure the proper user accounts are set up.
    a. Root Account : Administration ,installation ,updating and removing to software packages
    b. Globus account : Administration ,installation ,updating and removing to Globus toolkit 4.0.1 source package with CA updates
    c. Jane account :User of grid system

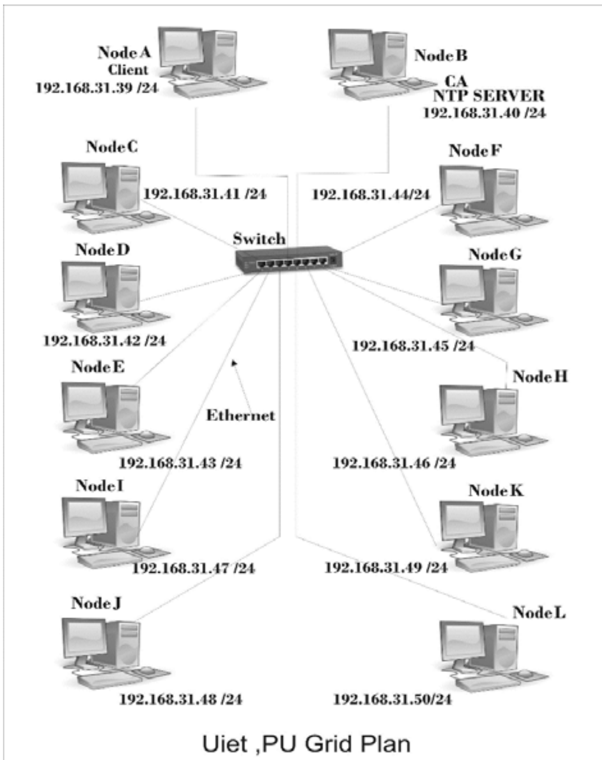

Figure 1 Grid Plan Used

The installation has been divided into 9 sections which are mentioned below .We are only giving brief on 3 nodes nodeA(client node), nodeB (CA and NTP server), nodeC (Grid Processing node) , other nodes are replication of nodeC with some configuration differences of Domain name and IP address in the installation and configuration scripts shown in Table1.

Table 1 Network and Host configuration details

Node	**IPaddress**	**Hostname**
nodeA	192.168.31.39	nodeA.ps.univa.com
nodeB	192.168.31.40	nodeB.ps.univa.com
nodeC	192.168.31.41	nodeC.ps.univa.com
nodeD	192.168.31.42	nodeD.ps.univa.com
nodeE	192.168.31.43	nodeE.ps.univa.com
nodeF	192.168.31.44	nodeF.ps.univa.com
nodeG	192.168.31.45	nodeG.ps.univa.com
nodeH	192.168.31.46	nodeH.ps.univa.com
nodeI	192.168.31.47	nodeI.ps.univa.com
nodeJ	192.168.31.48	nodeJ.ps.univa.com
nodeK	192.168.31.49	nodeK.ps.univa.com
nodeL	192.168.31.50	nodeL.ps.univa.com

6. RESULTS

The Results of evaluations have been divided into 3 parts according to the length of input video. Following are the 3 cases that have been taken during evaluation.

1. Evaluation of grid with 6 minute job video
2. Evaluation of grid with 8 minute job video
3. Evaluation of grid with 10 minute job video

Figure 2 Physical Layout of Grid Environment in Lab.

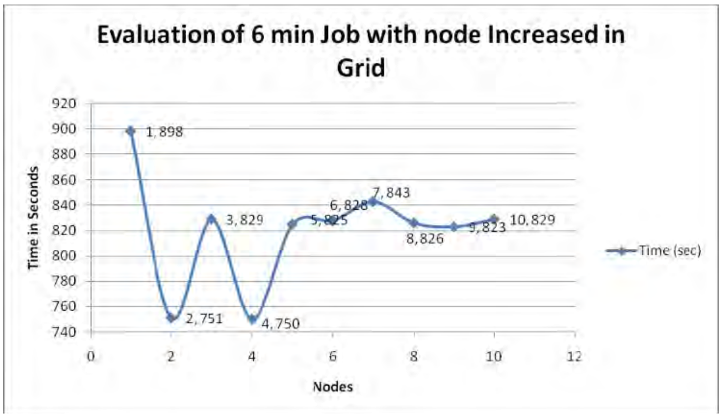

Figure 3 Evaluation of 8min video on grid nodes

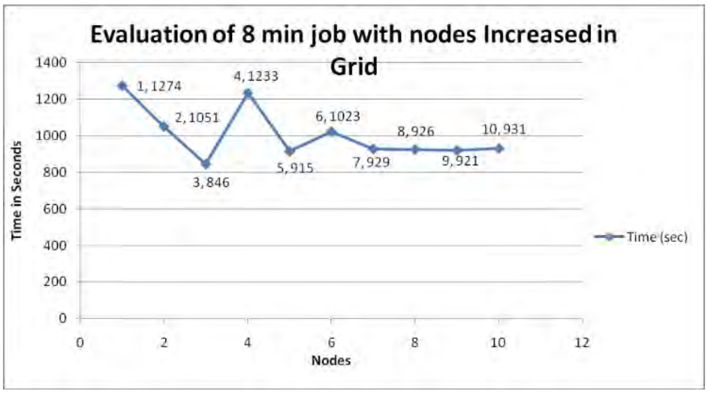

Figure 4 Evaluation of 8min video on grid nodes

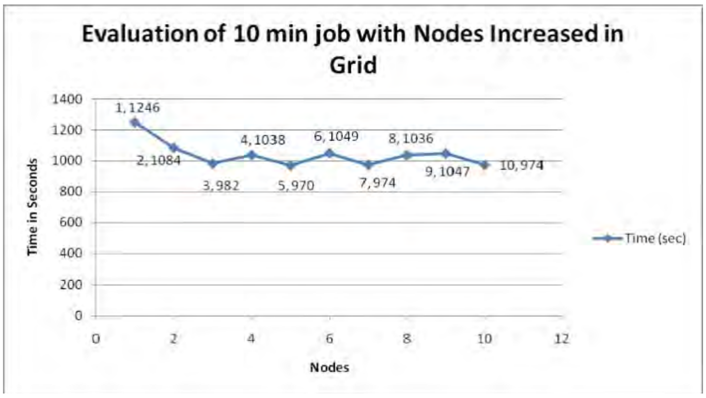

Figure 5 Evaluation of 8min video on grid nodes

It is clear from the graph in figure 3,4,5 that grid shows benefits over single processing systems but it does not show linear increase in performance by increasing the number of grid processing nodes.

After conducting thirty runs for three experiments by taking videos of different lengths on one to ten processing nodes, it is clear that grid shows benefits over single/centralised systems but it does not show up linear increase in performance by increasing the number of grid processing nodes. Theoretically, the experiment was expected to give linear increase in performance by increasing the grid processing nodes but this unexpected behaviour is due to following reasons:

1. Network Bandwidth
   As the network plays a substantial role in the grid processing and performance .So faster network (Gigabit) is proposed to be implemented in grid system implementations.

2. Scheduling Type
   The default scheduling strategy used by implementation is random but better results are expected if other scheduling strategies (e.g. priority based) are used.
3. Switching Fabric
   Network cannot be made efficient without making the switching fabric more efficient because it's the backbone of any network standard. So efficient switch fabric is a necessary requirement in grid systems.
4. Node's Local Processing and Availability
   Nodes in grid should have high availability of its resources.

As the research is going on better scheduling and better bandwidthnetworks for grid environment, the grid systems will one day prove to be equivalent to Cluster Computing Environments.

7. CONCLUSION

Throughout the experimental evaluation, we find that a grid computing can utilize the unused resources of your computer infrastructure (utilize them for processor intensive job) but its performance is far beyond the Cluster Computing Infrastructure.

After conducting thirty runs for three experiments by taking videos of different lengths on one to ten processing nodes it has been observed that grid shows benefits over single/centralised systems but it does not show up linear increase in performance by increasing the number of grid processing nodes.Theoretically, the experiment

was expected to give linear increase in performance by increasing the grid processing nodes but this unexpected behaviour is due to following reasons.
Network Bandwidth: As the network plays a substantial role in the grid processing and performance. So faster network (Gigabit) is proposed to be implemented in grid system implementations.
Scheduling Type: The default scheduling strategy used by implementation is random but better results are expected if other scheduling strategies (e.g. priority based) are used.
Switching Fabric: Network cannot be made efficient without making the switching fabric more efficient because it's the backbone of any network standard. So efficient switch fabric is a necessary requirement in grid systems.
Node's Local Processing and Availability: Nodes in grid should have high availability of its resources. As the research is going on better scheduling and better bandwidth networks for grid environment, the grid systems will one day prove to be equivalent to Cluster Computing Environments.

8. FUTURE SCOPE

In future, work can be initiated to add a better fault-detection policy to make it more robust. Python script designed and implemented during the evaluation are very static in job assignment. So in the future work we will try to make our script more dynamic for dynamic grid computing environments which will be able to detect and re-assign the same job to other computing node that is listed by MDS service. In future we will try to use better infrastructure for networking, so the drawbacks can be minimized at the level.

REFERENCES

[1] J Chem., "A database-centric virtual chemistry system", Journal of chemical information and modelling, Vol 46, Pg. 3, Jun 2006.
[2] Puhovski, Nenad (April 2000). "DV – A SUCCESS STORY". Www.stanford.edu. Retrieved 26 August 2013.
[3] John Watkinson, The MPEG Handbook, p.1
[4] Jump upHans Geog Musmann, Genesis of the MP3 Audio Coding Standard (PDF), retrieved 2011-07-26
[5] "About MPEG". Chiariglione.org. Retrieved 2009-12-13.
[6] Jump up "MPEG Meetings". Chiariglione.org. Retrieved 2009-12-13.
[7] Jump up chiariglione.org (2009-09-06). "Riding the Media Bits, The Fault line". Retrieved 2010-02-09.
[8] The MPEG Handbook, pp. 4–5
[9] Understanding MPEG-4, p.83
[10] "MPEG-J White Paper". July 2005. Retrieved 2010-04-11.
[11] "MPEG-J GFX white paper". July 2005. Retrieved 2010-04-11.
[12] ISO. "ISO/IEC 14496-21:2006 – Information technology – Coding of audio-visual objects – Part 21: MPEG-J Graphics Framework eXtensions (GFX)". ISO. Retrieved 2009-10-30.