Access Grid Performance Exploration: From a Network perspective

Xuehai Zhang
Dept of CS
Univ of Chicago
hai@cs.uchicago.edu

Abstract

We discuss the Access Grid's performance exploration from the consideration about the network requirements and deployment. We focus on two important metrics in Access Grid's network. One is the bandwidth and the other is the multicast protocol. We give detailed study about the requirement of these metrics, how they are deployed in the Access Grid, and we try to offer evaluation of them in Access Grid environment. However, this work should be supported by the real data collected from the network monitoring, but due to specific reasons, we could not do large-scale experiments so far. Therefore, we have given our consideration of future work towards this.

1. Introduction

Through several years of quick development, the technology of Collaborative Virtual Environments (CVEs), including large scale, distributed multi-user virtual reality systems that support collaboration, has been put into practical use in commercial environment from the research domain. Today we have already seen a lot of mature applications of CVEs. Access Grid is a right representation of these CVEs systems.

First, we give a brief description about the Access Grid system. Access Grid is the infrastructure and software technologies enabling linking together distributed Active (Work) Spaces to support highly distributed collaborations in science, engineering and education, integrated with and providing seamless access to the resources of the National Technology Grid. Here is another definition through a deferent viewpoint: Access Grid is an open grid based community, and a platform that supports group-to-group collaboration, and an emerging collection of open source tools, and a way to think about the Grid.

Till now, over 20 Access Grid nodes are deployed all round the world (http://www-fp.mcs.anl.gov/fl/accessgrid/ag-nodes.htm) and they are put into a large-scale usage not only in research area but also in education, commercial, etc fields. It has proved itself a very successful CVE system and should expect a very bright future. As it becomes more widely deployed and gains more attention, so the Access Grid developers and providers would be faced with the problem of how to provide the best and continuous services. This is not a trivial issue, actually it involves an immerse aspects that they should take into consideration. As a matter of fact, there exist some determinative factors that people know little even now. In this paper, we are not supposed to answer such a big question; but instead, we do want to perform some exploration on Access Grid's performance evaluation by keeping the intention in mind to provide a valuable help to answer this question. Since the performance is really a big title, we plan to focus on one aspect in the technical factor of Access Grid – network and communication requirement analysis and evaluation. The reason we want to do such a study is because we have the feeling that in some sense, the quality of service that the Access Grid's network and communication system provide is the decisive factor of the whole system. We all know that Access Grid, as well as other CVE systems, could not become what they are today without the big support of the high performance networking. In the remainder of this paper, we will examine the network and communication requirements of Access Grid System; discussing the current configurations of these two aspects implemented in Access Grid and giving evaluation upon them; at last, we will give our consideration about the improvements.

2. Access Grid Network Topology / Architecture

The Access Grid is an effort to build up a grid of distributed human interactions across the network. Examples of such interaction include videoconference, data sharing, group discussion, etc. The Access Grid design point is group-to-group communication (thus differentiating it from desktop to desktop based tools that are focused on individual communication).

The figure below shows the deployment of Access Nodes in the world. They have constructed a certain kind of grid network.

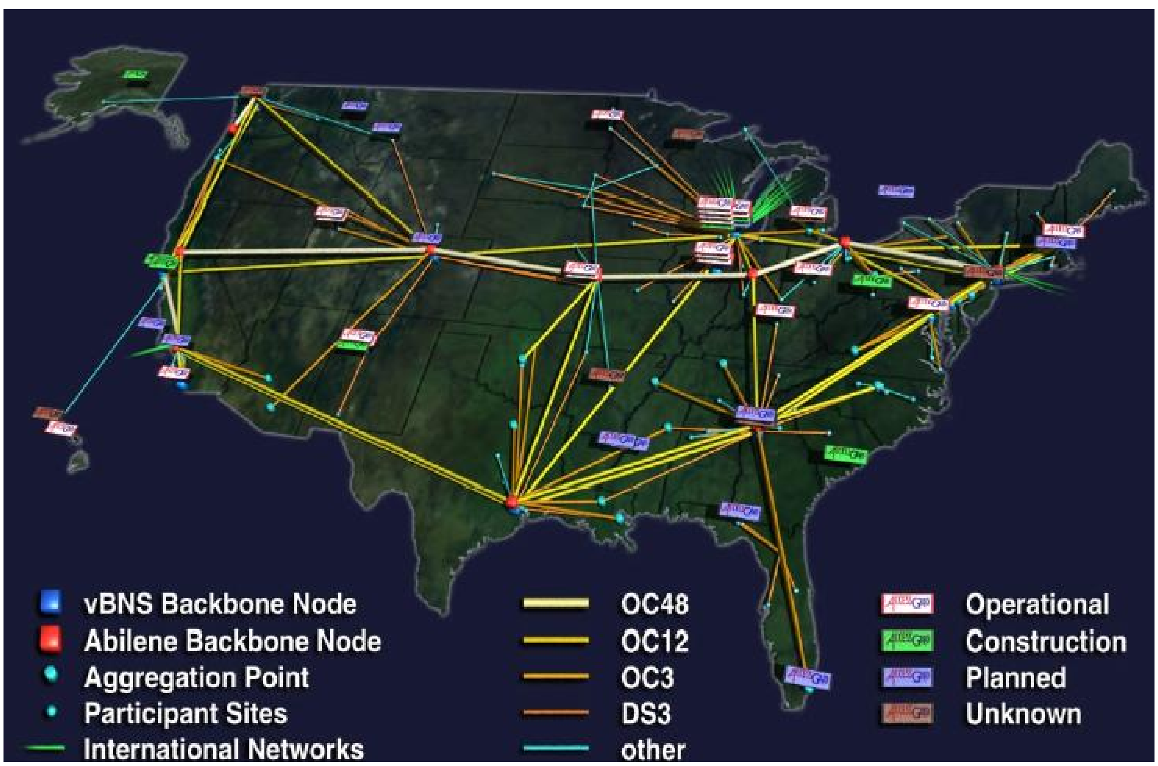

Figure 1: Access Grid Nodes in the World

The above picture gives us an overall impression about Access Grid network. Then, we turn to one Access Grid node for more technical details. Typically, one Access Grid node consists four main components: Display Computer, Video Capture Computer, Audio Capture Computer, and Control computer. Figures 2 describes such a structure and provides the relationship among these components. We do not discuss the functionalities of these components here in details; you could find more specific descriptions at main Access Grid web site (http://www-fp.mcs.anl.gov/fl/accessgrid/).

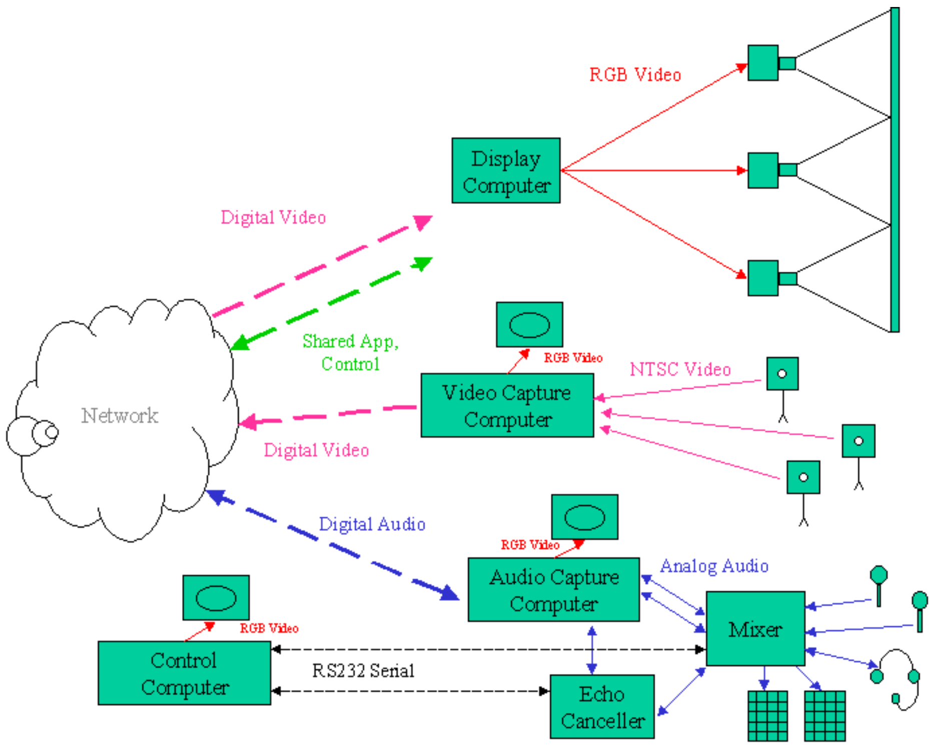

Figure 2: Component of an Access Grid Node

3. Network Analysis and Evaluation

Collaborative Virtual Environments, like almost all distributed systems, are complex and encompass a broad range of interrelated concerns. Among these concerns, network requirements are particularly important; in some sense, it is a decisive factor of the final quality of service the CVEs might provide. It arouses us a lot of interests and we decide to focus our research upon this issue. Since there are many sub fields in network requirements, we narrow our research work to two metrics: one is bandwidth, the other is multicast protocol. We will provide the reason why we select these two topics in the following chapters.

3.1 Bandwidth

3.1.1 Requirements For Access Grid

Although the final network requirements will depend on exactly what the exact Collaborative Virtual Environment is, the bandwidth requirement (understanding both the volume of network traffic that will produced and also its general characteristics) is such an important issue that every CVE should take into serious consideration.

The underlying reason is quite simple: CVEs are built on the top of networks (Internet or native multicast infrastructure) and since most of them are large-scale, distributed, many-to-many collaborative systems, they have a critical requirement about bandwidth allocation cause several joining parts in the CVE have to share the network bandwidth simultaneously; and in more specific technical details, there are multiple streams that coexist in the network flow at the same time.

It is very difficult to predict the network traffic that will be generated by CVEs. Many factors contribute to this. CVEs are an inherently complex class of application. Multiple users share a computer generated virtual world through local- or large-scalable network; can communicate with others using various media, including audio, video, text, graphics, etc. Consequently, CVEs involve a variety of traffic types from dynamic movement updates, through continuous audio and video to potentially large state-transfers as users join new worlds. Further more, network traffic is likely to be highly dependent on patterns of user behavior.

The same situation happens in Access Grid. In Access Grid, bandwidth requirements are variable based on the number of separate video and audio feeds per virtual venue (multicast group address). The provider gives their estimation. In practice, a configuration of 6-10 persistent AG nodes generates approximately 10 Mbps of traffic, while running idle with no user interaction. Bandwidth climbs when nodes are in use due to increased motion in video streams, increased use of audio and additional site joining. The minimum available bandwidth to each site should be greater than 10 Mbps for less than 5 sites in an access-grid network. Each additional node increases the minimum requirement by approximately 1.5 Mbit/s per site.

3.1.2 Network Traffic Predictive Model

In the above chapter, we describe the approximate bandwidth requirements for Access Grid provided by the Access Grid provider. But unfortunately, they do not give any explanation about how they obtain such a conclusion and what experiments they have performed for this targets.

Although it is very difficult to build a precise network traffic predictive model to estimate the exact bandwidth requirements that the CVEs require, scientists have done several in-depth researches to try to provide reliable solution. The model invented by Prof. Greenhalgh is such a widely accepted one.

3.1.3 Greenhalgh's Predictive Model

Greehalgh's predictive model of network traffic for a given CVE application takes account of the number of simultaneous participants, the pattern of their activity and the behavior of the application itself. It conducts extensive trials with the application in order to capture and then analyses two kind of data. The first kind of data is detailed event logs that record and time-stamp every action of every participant, where the set of possible actions is determined by the particular application. The second kind is measurements of network traffic on different network links, captured using the tcpdump utility.

This data is then analyzed according to three steps:

- User behavior model – statistical analysis of the event logs results in a user behavior model that describes typical patterns of user activity with this application. This model summarizes the average frequencies of actions; the extent to which patterns of activity vary according to participant and phase of the applications; and the extent to which actions are correlated among users. Correlation is concerned with whether participants tent to perform specific actions at the same time, whether they avoid overlaps or whether they are acting independently.
- Application behavior model – this step involves developing an application behavior model that identifies the traffic that will appear on a specific network link when a given application event occurs. There are two aspects to this model. First, through a process of controlled traffic measurement, it determines the size of the messages that are sent between different application components whenever a given action occurs. Secondly, by considering how the logical communication architecture of the application maps down onto the underlying physical network topology it considers how these messages become concentrated on different network links. Combining these two aspects allows us to predict the network traffic that should appear on a given link for a specified sequence of action. The application model is validated by comparing the traffic that is predicted for an example session with the traffic that was actually observed.
- Network Traffic Model – this step combines steps 1 and 2 to create a model of network traffic that predicts the traffic that might be generated by this particular application. It is particularly interested in the mean and peak level s f traffic that might be expected on a given link. The mean traffic on a link is predicted from the average frequencies of events; the peak traffic is given by the worst-case situation, where users act as much as is possible.

3.1.4 Implement Greenhalgh's Model in Access Grid

Greenhalgh's model described in the above chapter provides us a predictive mechanism for general CVEs. It could be used to Access Grid, but we should have some specific consideration about the implementation this model to this specified environment. Here we give some tentative thinking about it.

- It is almost true that in CVEs, all of those network requirements derive directly from what the users chose to do and when they choose to do it. In Access Grid world, the actions available to a user include connecting to the Venues and entering the Access Grid Lobby; logging off from the Venues; communicating (speaking, conveying graphics information about their identity, physical presence, movement, etc) and other actions. We should learn every significant action that a user performs has an impact on the system's network communication.

- The Access Grid supports a range of possible user actions. It also communicates each of theses actions to the rest of the system using slightly different information in a slightly different form. We think the Access Grid has three key actions: user audio, user video and user awareness about the other participants. Each of these component activities has a characteristic communication requirement. These intrinsic system values must be combined with information about actual user activity in order to predict what will happen in practice.

3.2 Multicast Routing Protocols

3.2.1 Requirements For Access Grid

Collaborative Virtual Environments such as the Access Grid system require transfer of high bandwidth media stream s to a large number of users. Traditional l Internet protocols are mostly unsuitable in handling such high bandwidth traffic flows. Several revolutionary developments have taken place. Among these, the solution is multicast communication. In these CVEs, multicast is a mechanism for one-to-many or many-to-many delivery of data over the Internet in a bandwidth efficient manner.

During the first five to six years of tits deployment, multicast is development in the type as a virtual and experimental network constructed on top of the Internet. We called this topology the Multicast Backbone (MBone) which used the distance Vector Multicast Routing Protocol (DVMRP). However, in the later, this architecture proved to have some inherited disadvantages because it’s building solely on top of existing network. People began to think to evolve it into a native multicast infrastructure that will buy them more benefit about deploying multicast. Several new protocols are invented and put into practical use. For examples, Protocol Independent Multicast – Dense Mode (PIM-DM), protocol Independent Multicast – Sparse Mode (PIM-SM), the Multicast Border Gateway Protocol (MBGP) and the Multicast Source Discovery Protocol (MSDP).

3.2.2 Access Grid Solution

Since Access Grid is aimed at providing large-scale group-to-group collaborative activity and has several real-time media streams involved, it required running a robust, stable and scalable multicast routing infrastructure. The Access Grid providers suggest such a combination solution: to use PIM-SM in the LAN and PIM-SM / MBGP / MSDP for peering between Autonomous Systems. The following is their original thinking and arrangement:

- This combination will provide Native IP Multicast instead of the old-style multicast architecture based upon the existing Internet.
- The multicast service uses state of the art multicast protocols such as PIM sparse mode, MBGP, and MSDP. PIM sparse-mode offers a scalable solution to distribute multicast traffic based on explicit joins there by reducing the network load and optimizing the multicast routes across different domains with all the routing management features that are inherited from BGP. MSDP allows multicast sources to be announced across multiple peering domains. The advantages of such

arrangement are that information sent based on receiver joining group (explicit join). In the other hand, the flood and prune nature of DVMRP does not scale. MBONE tunnels are typically constructed using DVMRP. This topology is not compatible with the goals of the AG deployment.

3.2.3 Evaluation Analysis (Multicast Monitoring)

CVEs like Access Grid's large-scale and distributed nature bring a big challenge to perform evaluation of its multicast protocol's functionalities. As to Access Grid system, the provider does not offer the evaluation means by his own. They suggest we could do the evaluation through multicast monitoring by using the tools provided by the third parties.

But the rapid pace of developments in multicast has been accomplished with only a few monitoring and debugging tools. Tool development and monitoring efforts have not kept pace with the rapid deployment. Some monitoring tools are insufficient for use in the current infrastructure or for use with current protocols. They also provide only limited data collection, analysis and display functionality. Also most of the existing tools were developed in the context of MBone, and its flat routing topology. Here, we will describe two multicast monitoring tools that suffice such requirements and could be used to measure the Access Grid's performance.

3.2.3.1 Multicast Beacon

The multicast monitoring and debugging tool used and recommended by Access Grid is NLANR's Multicast Beacon. It is the key component used to generate very low bandwidth multicast sender/receiver to enable debugging of multicast topology. It is an active measurement software to monitor the performance of a multicast session. It has four components:

- **Beacon** -- an active probing program running at each machine. A set of Beacons sends packets continuously to each other through a multicast session, and measures the performance of the transmission. It then reports to the BeaconServer periodically.
- **BeaconServer** -- a central server collecting the performance information from the Beacons. A BeaconServer is capable of handling multiple multicast sessions.
- **BeaconViewer** -- a GUI program to query the BeaconServer and display the performance information for a multicast session.
- **WebView** -- a web interface to the BeaconServer without running the BeaconViewer.

Teamed up with the Access Grid, the Multicast Beacon will give measurements to the current multicast traffic on the network. The following is the current measurement metrics for a multicast session:

- **Loss** -- percentage of packet loss from a host to another
- **Delay** -- one-way delay from a host to another

- **Jitter** -- jitter of the one-way delay
- **Order** -- percentage of packets which arrived out-of-order
- **Duplicate** -- percentage of duplicated packets

The Access Grid provides the current state of multicast beacons at http://beaconserver.accessgrid.org:9999/. The figure below is a snapshot of the output.

3.2.3.2 Mantra

Mantra is a Java-based tool for monitoring multicast at the network layer and for presenting the results interactively. It is developed at University of California, Santa Barbara. The primary function of Mantra is to collect data by periodically capturing multicast router state. Mantra's ability to collect data at the router level enables us to monitor activity based on what the router sees instead of what users see.

The design of Mantra is based on four major modules each of which is responsible for performing a major monitoring task. The figure below displays these modules and the data flow in Mantra.

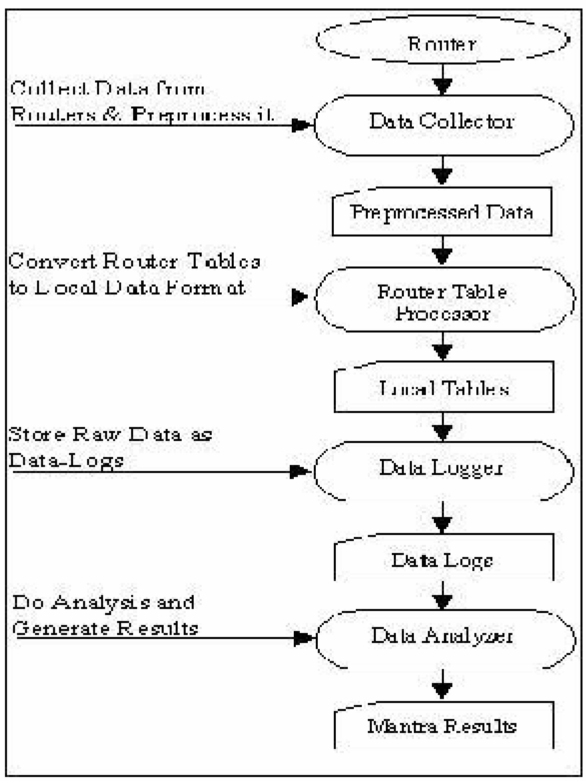

Figure 3: Mantra modules and information flow

We briefly describe the design and functionalities of there modules.

- Data Collector: Through this module, Mantra collects monitoring data from multicast routers and prepares the collected data fro further processing. The main functions include: acquiring data from the router; assuring the accuracy of data transfer; and preprocessing the raw dart.

- Router-Table Processor: This module converted collected data to Mantra's local data format. The main task of it is to map each of the pre-processed data tables to the corresponding local tables. It also involves estimating the missing data sets, estimating durations of various entries, in the table, and removing noise from the data sets.
- Data Logger: The data logger stores the processed data. This data can be later used for detailed off-line analysis as well as for long-term trend analysis.
- Data Analyzer: the data analyzer is a collection of modules that allow Mantra to analyze monitoring data and generate online results. Every time a new set s of data tables is processed and logged, analysis modules are invoked.

The Access Grid has not deployed Mantra to monitor its multicast yet. So, it is an interesting topic that some day we deployed Mantra as the alternative monitoring tool for Access Grid, thus we could compare its result with those came from Multicast Beacon. We are expected to obtain more positive, reasonable evaluation on Access Grid's multicast performance at that time.

4. Conclusion and Future Work

In this paper, we describe our research work upon the network requirements for Access Grid. Our research work is carried out in two metrics in networks, the bandwidth and the multicast protocol. We discuss in details about the requirements of these two metrics, and the current solution in the Access Grid. Also, we provide some alternative consideration about the models and tools in correspondence with the existing ones.

Due to some constraints, we could not perform large scale real testing experiments to provide even stronger support in our research work. In the future work, we will try to make these experiments into practice.

Network requirements are only one major aspect to evaluate the performance of the Access Grid. It is such a complicated CVE system that a lot of aspects not only technical aspects but also human factors involve in it. So, it is not a trivial work to do the performance evaluation upon it.

5. Reference:

Dommel P., Aceves JJ (1995) Floor Control for Activity coordination in Networked Multimedia Application - Proc. 2nd Asian-Pacific Conference on Communications (APCC)'95, Osaka, Japan, June 12-16, 1995.

Dommel P., Aceves JJ (1998) Efficacy of floor control protocols in Distributed multimedia collaboration – Cluster computing journal, Nov 1998

Floyd S., V. Jacobson, S. McCanne, C. G. Liu, and L. Zhang (1995) A Reliable Framework for Light-Weight Sessions and Application Level Framing - ACM SIGCOMM '95. Boston. August 30-September 1, 1995.

Floyd S., Varadhan K., Estrin D (998) Impact of Network dynamics on End-to-End protocols: Case studies in TCP and Reliable Multicast.- research draft "http://www.isi.edu/~kawnan/VINT/ic98.ps"

Ott J., Bormann C (1997) MTP/SO - Self organising Multicast Internet draft 1997, draft-bormann-mtp-so-01.txt