The GRID as a platform for communication, collaboration and e-science

Stephen Hinde, Lawrence Wilcock
Internet Systems and Storage Laboratory
HP Laboratories Bristol
HPL-2002-125
May 8th , 2002*

E-mail: {Stephen_Hinde, Lawrence_Wilcock} @hpl.hp.com

telecoms, data center, GRID, communications, collaboration

This paper explores the role of communication and collaboration within the GRID. We discuss how a new generation of communications applications are becoming more resource hungry. We propose a platform for communication and collaboration applications that could itself be built on the GRID fabric. We discuss briefly the attributes of such a platform.

* Internal Accession Date Only
Approved for External Publication

The GRID as a platform for communication, collaboration and e-science

Stephen Hinde and Lawrence Wilcock
HP Laboratories
Filton Road, Stoke Gifford, Bristol, BS34 8QZ, UK.
{Stephen_Hinde, Lawrence_Wilcock}@hpl.hp.com

Abstract

This paper explores the role of communication and collaboration within the GRID. We discuss how a new generation of communications applications are becoming more resource hungry. We propose a platform for communication and collaboration applications that could itself be built on the GRID fabric. We discuss briefly the attributes of such a platform.

1. Introduction

Almost everything we do involves communication with other people. Global business, education and scientific research simply would not be possible without electronic communications networks.

Two important trends are changing the way that people communicate electronically. The first is technological, and relates to the increasing use of computation to enhance the way that people find and communicate with others, in areas such as sophisticated routing, persistence, media processing, security and collaborative information presentation. The second is social, and relates to the emergence of new style of communication applications to support online social communities; examples include collaborative science, online gaming, social meeting places, and online learning.

An analysis of both of these trends leads us to believe that new approaches to the construction of communication applications, involving large-scale utilization of computing and storage resources allocated from a global pool, are now possible and necessary to support this new class of application. The emergence of a global computing fabric allows certain assumptions for how to architect a communication infrastructure to be reexamined, in particular how to structure communication sessions and process the real-time media inherent in multiparty communication. The technologies being developed in the GRID community will be critical to the success of a vision of a global computing fabric for communication. A thorough understanding of the requirements and demands placed on such a fabric by communication applications is essential for it's future success.

This is a position paper making the case for large-scale conferencing and collaboration as an important example of a resource hungry application, potentially disruptive in terms of requirements for computation, and an important driver for planetary-scale computing and the GRID. Our central thesis is that the use of a large, distributed computational fabric based on GRID technologies, can be used to produce a scalable communication platform and media processing fabric to support a high-quality communication environment.

We propose that a collaborative platform could be developed as part of the GRID forum. We explore some of the attributes of such a platform.

2. Online social and collaborative applications

Recent trends towards online social communication and large-scale collaborative working in fields such as science and education show that patterns of communication are changing in a number of significant ways that highlight deficiencies in current communication technologies. People now communicate in larger groups. The paradigms for communication are also changing. People meet at virtual places, such as within a game environment or chat room. These places have an internal logical structure, often modeled on real world concepts such as buildings or rooms, and may have media-specific properties such as acoustics. The notion of position within the environment may be significant, both in terms of rendering positional audio, and deciding which region one is in or the group of people to communicate with.

People also use many different kinds of communication devices. The telephone network has been designed on the assumption that everyone has essentially the same device capability, namely the ability to exchange audio at 64kb/s. This assumption is no longer true. A broad spectrum of devices must be supported, from mobile phones, through to powerful computers with spatial audio capabilities connected to multi-megabit data networks.

3. A new generation of collaborative platform

What is required is a scalable communication environment based on an enhanced notion of a session and media processing. The environment should support a notion of space so that people can move around and dynamically form groups just like in the real world, and potentially capable of supporting hundreds of thousands of participants. Without these properties, the usability of an environment with many simultaneous speakers would quickly deteriorate. The audio environment could exhibit a simplified acoustic physics, such as echo, reflections, reverberation etc, in order to generate a convincing sense of presence. In addition, it must be possible to flexibly create complex media interconnection relationships between people, based on the synthetic rules and physics of the environment itself, and logical grouping of individuals. Media processing nodes within the network must be used to adapt to the characteristics of a variety of devices, such that the lowest common denominator is not dominant.

The emergence of a new compute utility, offering large-scale computation resources in the network, offers the possibility of significant advances in the usability, richness, realism, and scale of communication with others in online places. The challenge is to architect a communications infrastructure capable of supporting the audio, video, and other media exchanges in a scalable way, and with the functionality required for the class of new applications characterized as online social communication.

3.1 Related activities

Research platforms within the Collaborative Virtual Environments community share some of the same characteristics as the new class of online social applications, examples being COVEN, DIVE, and MASSIVE. Much of the focus and progress in this community has been to solve problems in the application domain such as consistency of shared world databases. Arguably, less attention has been given to communication architectures for the exchange of real-time media. Despite some promising work within the community, one example being the Locale, aura session model and partial distributed audio mixing techniques [1] developed for MASSIVE [2], little consensus has yet emerged on a common architecture.

The AccessGrid supports group-to-group communication via high-speed networking and state-of-the art display and camera technology over the Internet. While important in terms of providing packaged solutions for some collaborative work, real-time media communication follows the MBONE solution, and consequently shares it's limitations.

3.2. Media processing on a computational GRID

Real-time media processing is a critical part of electronic communication networks. Mixing, transcoding, and efficient distribution of media, is both computationally intensive and places strict quality-of- service demands on the underlying network. However, the approach to the electronic transmission and processing of real-time media for multiple participants, outside of the research community, has not advanced significantly for 10 years.

Within the Internet telephony community, two standards camps have emerged. The H.323 set of standards from the ITU has evolved from earlier standards efforts for the PSTN and shares many of the same limitations as the conventional telephone network. The MBONE tools and SIP conferencing standards from the IETF community are better architected for the Internet, but still share many of the same scaling and functionality limitations as the ITU standards.

In both communities, two principal architectural approaches are used to perform the required mixing and transcoding – endpoint mixing or centralized mixing. In end-point mixing, media from all participants is sent to all end devices and mixed on the device itself. Alternatively in centralized mixing, a single network-based media node, performs the media processing to create a composite view of the conference from the media received from all devices, and returns this to the devices. Neither approach easily supports the structuring mechanisms, scalability, and device adaptation required for online large-scale collaborative applications.

The new opportunity enabled by the large distributed computational fabric based on Grid technologies, is to produce a scalable audio mixing and media processing fabric to support high-quality audio and video communication environments. The elements in the fabric would be connected by a well-provisioned network to create a flexible media-processing island. HP Laboratories is working on a communication architecture and enhancements to the Session model, called *Spaces*, to manage and control hierarchical, scalable, lightweight conferences, running on such a fabric, with the characteristics described to support social applications.

10. References

[1] Radenkovic, M., Greenhalgh C, Benford S, "A Scalable Audio Service for CVEs", *Proceedings of the sixth conference of the UK VRSIG*, Salford, UK, Sept 1999.

[2] Greenhalgh, G, "MASSIVE-3/HIVEK Introduction", http://www.crg.cs.nott.ac.uk/research/systems/MASSIVE-3/docs/massive3-intro-summary.pdf