Chapter 3.7
A Distributed Storage System for Archiving Broadcast Media Content

Dominic Cherry
Technicolor Network Services, UK

Maozhen Li
Brunel University, UK

Man Qi
Canterbury Christ Church University, UK

ABSTRACT

This chapter presents MediaGrid, a distributed storage system for archiving broadcast media contents. MediaGrid utilizes storage resources donated by computing nodes running in a distributed computing environment. A genetic algorithm for resource selection is built in MediaGrid with the aim to optimize the utilization of resources available for archiving media files with various sizes. Evaluation results show the effectiveness of MediaGrid in archiving broadcast media contents, and the performance of the genetic algorithm in resource utilization optimization

INTRODUCTION

It was the mid 1950s when the Ampex Corporation's Quadraplex professional video tape recording system was first widely available. Since then, production, storage and playout of television programs has been centered on magnetic tape technologies. Recently, the advancement of disc based, server technologies has had an impact on the broadcast industry. There is now an increasing trend towards the 'tape-less' production environment (Watkinson, 1990).

It is now possible to shoot an entire programme's material directly to disc and edit it using software tools, such as Avid or Apple's Final-Cut Pro. A completed programme can be delivered to a transmission playout centre using an optical fibre channel network. Here, the content is reviewed and verified from a video server before playing-out as a video stream, at the scheduled time, to

DOI: 10.4018/978-1-4666-0879-5.ch3.7

the transmitter. Recent consumer gadgets like Personal Video Recorder systems allow a viewer to record or time-shift their favourite shows. At no point in this example has the programme content ever touched a piece of magnetic tape.

Server based transmission has been widely used in broadcast infrastructure. In the multi-platform, multi-channel environment, transmission playout from the server is preferred whenever possible. Two of the many advantages are, being able to preview any part of the content close to transmission without having to worry about whether there is enough time to spool the tape back to the start of the programme. Also, it is possible to view the end of the file (on an alternative server output), whilst the beginning is being transmitted. This level of media access is not possible with tape.

Transmission playout servers are specialist items of broadcast equipment which translate content between a file stored on a disc array and a real-time MPEG video stream. When the programme's scheduled air-time occurs, the server plays back the file, under control of an automation system, and outputs an MPEG video stream. Because they are highly specified, these are costly machines with a current price ranging from £40k-£70k each, depending on capacity. Their storage capacity therefore comes at a premium. There is an upper limit on both the number of files each server can hold and the total storage capacity. Due to the cost of these machines, their capability is not a suitable option for longer term storage and archive.

The past few years have witnessed a rapid development of grid computing (Foster and Kesselman, 1998; Li and Baker, 2005; Berman et al., 2003), a computing paradigm to facilitate utilization of resources on the Internet. This article presents MediaGrid, a light-weighed storage system for archiving broadcast media content utilizing resources dispersed in a distributed environment. We applied the concept of grid computing aiming to provide a novel solution to the problem of not having enough storage space to hold the ever-increasing mass of media content. As such, it performs a single function of the material management systems within a broadcast infrastructure.

It is worth noting that MediaGrid is related to GridCast (Harmer et al., 2003). The scope of the GridCast project is much broader, aiming to provide all the basic functions required within a broadcast material or content management system. The extent of this scope includes content sharing, browsing and trafficking, as well as assisting in broadcast scheduling. The GridCast project implements these functions as Web services using open standards and the Globus Toolkit (Sotomayor and Childers, 2005). MediaGrid can be plugged into GridCast as an archive service exploiting desktop PC storage at each site, so could be part of the content storage and retrieval process.

BACKGROUND

Figure 1 shows how and where server based technology is being used within a broadcast infrastructure. This diagram is based on the design model of an existing transmission playout centre.

The pale green blocks represent the areas where the transmission infrastructure is still predominantly broadcast specific, single purpose, proprietary equipment. In the presentation chain, this is equipment such as the video mixer, DVE (Digital Video Effects) processors, and data bridges or inserters. DVEs provide 'real-time' vision processing such as keying and manipulating video layers, providing effects like picture squeezes, captions and picture-in-picture overlays. Data bridges add data, such as teletext subtitling information, or wide-screen-signal coding to the video signal. This data is broadcast with the pictures and decoded at the receiver.

The switching matrix, often called a video router, creates cross-point between video-bandwidth input and output busses. These are used to switch the processed signal onwards to its destination. In the case of this diagram describing a playout