Chapter II

HYDRA: High-performance Data Recording Architecture for Streaming Media\(^1\)

Roger Zimmermann, University of Southern California, USA

Kun Fu, University of Southern California, USA

Dwipal A. Desai, University of Southern California, USA

ABSTRACT

This chapter describes the design for High-performance Data Recording Architecture (HYDRA). Presently, digital continuous media (CM) are well established as an integral part of many applications. In recent years, a considerable amount of research has focused on the efficient retrieval of such media for many concurrent users. The authors argue that scant attention has been paid to large-scale servers that can record such streams in real time. However, more and more devices produce direct digital output streams, either over wired or wireless networks, and various applications are emerging to make use of them. For example, cameras now provide the means in many industrial applications to monitor, visualize, and diagnose events. Hence, the need arises to capture and store these streams with an efficient data stream recorder that can handle both recording and playback of many streams simultaneously and provide a central repository for all data. With this chapter, the authors present the design of the HYDRA system, which uses a unified architecture that integrates multi-stream recording and retrieval in a coherent paradigm, and hence provides support for these emerging applications.

This chapter appears in the book Video Data Management and Information Retrieval by Sagarmay Deb. Copyright © 2005, IRM Press, an imprint of Idea Group Inc. Copying or distributing in print or electronic forms without written permission of Idea Group Inc. is prohibited.
INTRODUCTION

Presently, digital continuous media (CM) are well established as an integral part of many applications. Two of the main characteristics of such media are that (1) they require real-time storage and retrieval, and (2) they require high bandwidths and space. Over the last decade, a considerable amount of research has focused on the efficient retrieval of such media for many concurrent users. Algorithms to optimize such fundamental issues as data placement, disk scheduling, admission control, transmission smoothing, etc., have been reported in the literature.

Almost without exception, these prior research efforts assumed that the CM streams were readily available as files and could be loaded onto the servers offline without the real-time constraints that the complementary stream retrieval required. This is certainly a reasonable assumption for many applications where the multimedia streams are produced offline (e.g., movies, commercials, educational lectures, etc.). In such an environment, streams may originally be captured onto tape or film. Sometimes the tapes store analog data (e.g., VHS video) and sometimes they store digital data (e.g., DV camcorders). However, the current technological trends are such that more and more sensor devices (e.g., cameras) can directly produce digital data streams. Furthermore, some of these new devices are network-capable, either via wired (SDI, Firewire) or wireless (Bluetooth, IEEE 802.11x) connections. Hence, the need arises to capture and store these streams with an efficient data stream recorder that can handle both recording and playback of many streams simultaneously and provide a central repository for all data.

The applications for such a recorder start at the low end with small, personal systems. For example, the “digital hub” in the living room envisioned by several companies will, in the future, go beyond recording and playing back a single stream as is currently done by TiVo and ReplayTV units (Wallich, 2002). Multiple camcorders, receivers, televisions, and audio amplifiers will all connect to the digital hub to either store or retrieve data streams. At the higher end, movie production will move to digital cameras and storage devices. For example, George Lucas’ “Star Wars: Episode II, Attack of the Clones” was shot entirely with high-definition digital cameras (Huffstutter & Healey, 2002). Additionally, there are many sensor networks that produce continuous streams of data. For example, NASA continuously receives data from space probes. Earthquake and weather sensors produce data streams as do Web sites and telephone systems. Table 1 illustrates a sampling of continuous media types with their respective bandwidth requirements.

In this chapter, we outline the design issues that need to be considered for large-scale data stream recorders. Our goal was to produce a unified architecture that integrates multi-stream recording and retrieval in a coherent paradigm by adapting and extending proven algorithms where applicable and introducing new concepts where necessary. We term this architecture HYDRA: High-performance Data Recording Architecture.

Multi-disk continuous media server designs can largely be classified into two different paradigms: (1) Data blocks are striped in a round-robin manner across the disks and blocks are retrieved in cycles or rounds on behalf of all streams; and (2) Data blocks are placed randomly across all disks and the data retrieval is based on a deadline for each block. The first paradigm attempts to guarantee the retrieval or storage of all data. It is often referred to as deterministic. With the second paradigm, by its very nature of
Domain Specific Custom Search for Quicker Information Retrieval

[www.igi-global.com/article/domain-specific-custom-search-for-quicker-information-retrieval/101358?camid=4v1a](www.igi-global.com/article/domain-specific-custom-search-for-quicker-information-retrieval/101358?camid=4v1a)