Future Generation Computer Systems 27 (2011) 966–976

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

CineGrid Exchange: A workflow-based peta-scale distributed storage platform on a high-speed network

Shaofeng Liu [a,b,*], Jurgen P. Schulze [b], Laurin Herr [c], Jeffrey D. Weekley [d], Bing Zhu [e], Natalie V. Osdol [c], Dana Plepys [f], Mike Wan [e]

[a] Department of Computer Science and Engineering, University of California San Diego (UCSD), La Jolla, CA, USA
[b] California Institute for Telecommunications and Information Technology (Calit2), UC San Diego, La Jolla, CA, USA
[c] Pacific Interface, Inc., Oakland, CA, USA
[d] MOVES Institute, Naval Postgraduate School, Monterey, CA, USA
[e] Institute for Neural Computation, University of California San Diego, La Jolla, CA, USA
[f] Electronic Visualization Laboratory, Department of Computer Science, University of Illinois at Chicago, Chicago, IL, USA

ARTICLE INFO

Article history:
Received 21 March 2010
Received in revised form
23 July 2010
Accepted 12 November 2010
Available online 21 November 2010

Keywords:
CineGrid Exchange
Digital archiving
4K
Distributed storage
IRODS

ABSTRACT

The Academy of Motion Picture Arts and Sciences (AMPAS) report "The Digital Dilemma" describes the issues caused by the rapid increase of storage requirements for long-term preservation and access of high quality digital media content. As one of the research communities focusing on very high quality digital content, CineGrid addresses these issues by building a global-scale distributed storage platform suitable for handling high quality digital media, which we call CineGrid Exchange (CX). Today, the CX connects seven universities and research laboratories in five countries, managing 400TB of storage, of which 250TB are dedicated to CineGrid. All of these sites are interconnected through a 10 Gbps dedicated optical network. The CX distributed repository holds digital motion pictures at HD, 2K and 4K resolutions, digital still images and digital audio in various formats. The goals of the CX are: (1) providing a 10 Gbps interconnected distributed platform for the CineGrid community to study digital content related issues, e.g., digital archiving, the movie production process, and network transfer/streaming protocols; (2) building a tool with which people can securely store, easily share and transfer very high definition digital content worldwide for exhibition and real-time collaboration; (3) automating digital policies through middleware and metadata management. In this publication, we introduce the architecture of the CX, resources managed by the CX and the implementation of the first series of CX management policies using the iRODS programmable middleware. We evaluate the first phase of CX platform implementation. We show that the CX has the potential to be a reliable and scalable digital management system.

Published by Elsevier B.V.

1. Introduction

"The Digital Dilemma" [1] pointed out that the rapidly increasing use of digital technology in the acquisition, post-production and distribution of media content not only brings significant benefit to the motion picture industry, as well as other stakeholders in the media ecosystems, but also raises serious issues on how to manage a large amount of resulting digital content efficiently over long periods of time. Given that a single version of a movie created in ultra-high quality can fill tens of terabytes [1] of digital storage space, current technologies for preserving large amounts of data seem far behind the industry's needs. A unique feature of digital media content is its mobility. For instance, starting from its creation in the life cycle of a movie, digital media will need to be moved from one system to another many times for post-production, cinema distribution, long-term archiving, and on-demand retrieval. Therefore, how to manage and transfer digital content efficiently becomes increasingly important. The traditional method of delivering film cans by courier was initially adapted to deliver HDD and data tapes. But with the increasing volume of data transfers required for modern media productions which are themselves increasing distributed among team members spread around the world, physical delivery – and physical preservation – of digital media assets no longer satisfy the industry's requirements.

* Corresponding author at: Department of Computer Science and Engineering, University of California San Diego (UCSD), La Jolla, CA, USA.
E-mail addresses: s8liu@cs.ucsd.edu, liushaofeng@gmail.com (S. Liu), jschulze@soe.ucsd.edu (J.P. Schulze), laurin@pacific-interface.com (L. Herr), jdweekley@nps.edu (J.D. Weekley), bizhu@ucsd.edu (B. Zhu), vanosdol@pacific-interface.com (N.V. Osdol), dana@evl.uic.edu (D. Plepys), mwan@sdsc.edu (M. Wan).

0167-739X/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.future.2010.11.017

On the other hand, large amounts of long distance fiber optical cables have been installed during the past decade, which now makes dedicated 1–10 Gbps fiber connections more affordable. This trend is making it increasingly practical to transfer large digital media files between remote sites [2–4]. Furthermore, a distributed storage model has the potential of unifying resources around the world to form a petabyte scale distributed storage platform for media exchange and preservation.

CineGrid [5] is a research community with the mission "To build an interdisciplinary community that is focused on the research, development, and demonstration of networked collaborative tools to enable the production, use, preservation and exchange of very high quality digital media over photonic networks". Members of CineGrid are a mix of post-production facilities, media arts schools, research universities, scientific laboratories, and hardware/software developers around the world connected by up to 10 Gbps networks. Since 2005, CineGrid members have conducted pioneering experiments in digital media production and post-production, network streaming delivery, exhibition, and remote collaboration. These experiments created media assets that CineGrid members wanted to access over time, which meant they had to be stored somewhere, managed to ensure access and preservation, and transferred upon request among members scattered around the world. The first CineGrid Exchange (CX) nodes at UCSD/Calit2, UvA and Keio/DMC were established on an ad hoc basis to fulfill this requirement. But over time, as the number and size of CX nodes has increased, and the number and variety of digital media assets in the CX have grown, the CineGrid community has seen a need to systematically combine and integrate existing CX resources in order to provide a scalable solution for the storage of digital assets.

Just as distributed rendering architectures have been adopted by cinema post-production facilities to deliver visual effects shots on tight schedules, the trends in cloud computing (private and public) for virtualization of servers, storage and high-speed network infrastructure can be adopted for distributed digital content creation, distribution, library and archiving services. The CX is a pioneering effort to establish a global-scale networked test bed that can be used to experiment with pre-commercial technologies and prototype collaborative workflows. At the same time, the CX fulfills CineGrid's own requirements for secure access to the organization's terabytes of digital media content by replicating assets in geographically distributed repositories connected by persistent 10 Gbps networks.

With support from AMPAS and contributions from other CineGrid members, the CX development project was started in 2009 to design and implement a multi-layer open source digital media asset management system. This is the first large scale, distributed global storage implementation designed to handle digital motion picture materials at the highest quality and for the investigation of issues related to digital preservation storage networks.

The three major goals of the CX are: (1) providing a dedicated 1–10 Gbps interconnected distributed platform for the CineGrid community to study digital content related issues, e.g., digital archiving, the movie production process, and network transfer/streaming protocols; (2) building a tool with which people can securely store, seamlessly stream, share and transfer very high quality digital content across the world; (3) automating digital policies through middleware and metadata management, so people can define flexible content ingestion, digestion, replication, and deletion processes.

In Section 2 of this publication, we will discuss previous related work; in Section 3, we will describe the three-layer architecture of the CineGrid Exchange, in Section 4, we will talk about the CX policies and workflows. In Section 5, we show the user interfaces in the CX, and some experimental results and experiences. Finally, we will conclude and suggest future work in Section 6.

Fig. 1a. HIPerSpace, one of the world's largest display walls in Calit2, UC San Diego.

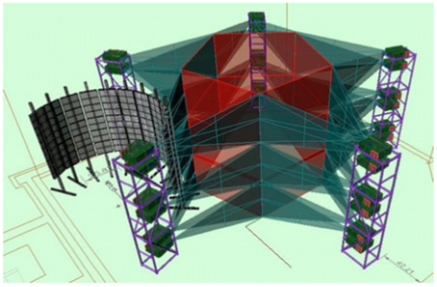

Fig. 1b. The STARCAVE, a third-generation CAVE and virtual reality OptIPortal in Calit2, UC San Diego.

2. Related works

For a few years, people have been experimenting with super-high-definition digital content, such as 4K, and transfer these content using high-speed lambda networks such as streaming 4K video clips. "4K" is a super-high-definition motion picture format. As defined by the Digital Cinema Initiative (DCI) consortium of Hollywood studios in 2003, 4K has up to 4096 pixels per line and up to 2160 lines per frame. In our paper, 4K refers to a frame size of 3840 × 2160 pixels, which is exactly four times the resolution of 1080p HDTV. 4K digital motion pictures can be captured by 4K cameras and displayed on 4K projectors, or 4K digital motion pictures can be created synthetically, without a camera, using the existing computer animation and scientific visualization tools. In addition to 4K projectors, the Optiportal [6] and the StarCave [7] are other examples of display devices for 4K and higher resolution digital content, shown in Figs. 1a and 1b.

The world's first (compressed) 4K tele-presence was demonstrated at the iGrid Conference in 2005 [8] as shown in Fig. 2. Both live and pre-recorded 4K content was compressed at 500 Mbps and streamed in real time using NTT Network Innovations Laboratories JPEG 2000 hardware codec via 1 Gb IP networks, from Keio University in Tokyo to iGrid 2005 in San Diego. Uncompressed 4K streaming typically consumes a bandwidth of 6–8 Gbps, and requires high-end hardware components throughout the data path; compressed streaming media may have data rates under 1 Gbps.

Those work showed the tremendous potential of using high-speed optical network connections to exchange large amounts of data between remote storage systems, which is superior to physically moving storage devices around. However, these experiments exemplify temporary installations that require configuration, tuning, testing, and in turn a great deal of effort to be repeated.

Data management software is a key component in large-scale digital content preservation. To gain flexibilities of managing digital content dynamically, we investigated iRODS, a piece of evolving software that provides certain programming capabilities allowing us to integrate digital managing policies into the software seamlessly, through the API provided, called "*rules*". **iRODS** the **Integrated Rule-Oriented Data System**, is "a data grid software