Future Generation Computer Systems 27 (2011) 886–892

Contents lists available at ScienceDirect

Future Generation Computer Systems

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

Real-time long-distance transfer of uncompressed 4K video for remote collaboration

Jiří Halák [a,*], Michal Krsek [a], Sven Ubik [a], Petr Žejdl [a], Felix Nevřela [b]

[a] CESNET, Zikova 4, Prague 6, Czech Republic

[b] CINEPOST, Kříženeckého Nám, 5/322, Prague 5, Czech Republic

ARTICLE INFO

Article history:
Received 26 February 2010
Received in revised form
16 June 2010
Accepted 12 November 2010
Available online 5 December 2010

Keywords:
Video
Color
Collaborative computing
Remote systems
Network communication
High-speed

ABSTRACT

Better-than-high-definition-resolution video content (such as 4K) is already being used in some areas, such as scientific visualization and film post-production. Effective collaboration in these areas requires real-time transfers of such video content. Two of the main technical issues are high-data volume and time synchronization when transferring over an asynchronous network such as the current Internet.

In this article, we discuss design options for a real-time long-distance uncompressed 4K video transfer system. We present our practical experience with such transfers and show how they can be used to increase productivity in film post-production, as an application example.

1. Introduction

Video transfers are an expected driver application area of the future Internet. Picture resolution has been increasing over time. Better-than-high-definition-resolution video (such as 4K) is already used in some areas, such as scientific visualization and the film industry. For the ultimate quality, required for instance in the color-grading process in film post-production, working with a signal that has not been compressed is preferable.

Presentation of high resolution video is now well possible using rendering devices with the corresponding resolution, using tiled displays (e.g., SAGE [1,2]) or multi-dimensional systems (CAVE [3,4]). The productivity of a distributed team can be significantly increased when the video signal can be transferred over the network in real time, to discuss and perform the processing of video content. Two of the main technical issues are high-data volume and time synchronization when transferring over an asynchronous network such as the current Internet.

For 4K resolution, the data volume ranges from 4.2 Gb/s for 4:2:2 subsampling [5], 10-bit color depth and 24 frames per second to over 9.6 Gb/s for RGB (no subsampling), 12-bit color depth and 30 frames per second. Overhead in packet headers needs to be added.

Real-time video streaming requires that the speed of rendering on the receiver side matches the rate of video source on the sender side. When the sender and receiver are connected over an asynchronous network, such as Ethernet, the receiver cannot directly synchronize its clock with the sender.

Therefore, the receiver needs to implement a technique that adjusts the speed of data arrival and maintains a continuous stream of video and audio data to the rendering device, preferably with minimal added latency for remote interactive applications.

We implemented the proposed architecture in a device called MVTP-4K (Modular Video Transfer Platform).

The structure of this paper is as follows. In Section 2, we summarize the main system requirements for real-time high-definition video transfers. The proposed system architecture is described in Section 3. Our practical experience is described in Section 4. Related work is referred to in Section 5 and our conclusions and thought about future directions are provided in Section 6.

2. Requirements and design constraints

In order to satisfy the needs of the targeted applications, we set the following set of requirements to be fulfilled by the proposed solution:

* Corresponding author.
E-mail addresses: halak@cesnet.cz (J. Halák), michalk@cesnet.cz (M. Krsek), ubik@cesnet.cz (S. Ubik), zejdlp@cesnet.cz (P. Žejdl), felix.nevrela@cinepost.cz (F. Nevřela).

0167-739X/$ – see front matter
doi:10.1016/j.future.2010.11.014

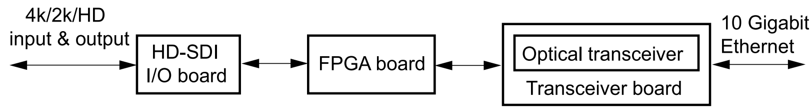

Fig. 1. Hardware architecture.

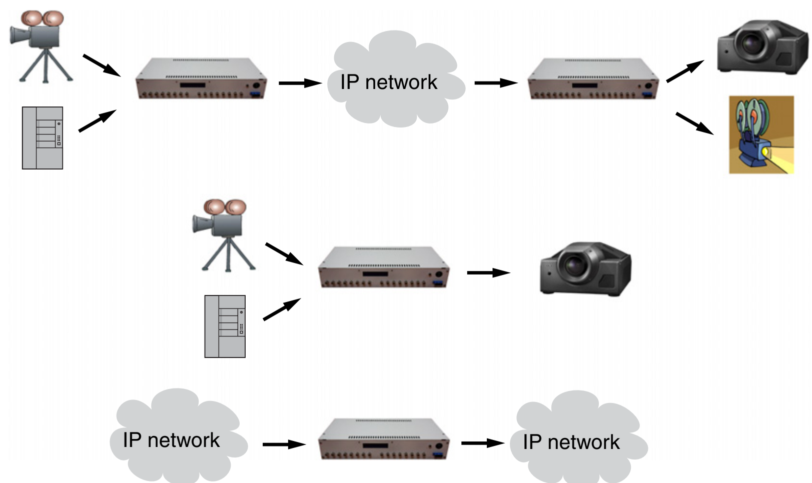

Fig. 2. Operation modes—network transfer (above), video processor (center), network processor (bottom).

- Real-time transfer of uncompressed 4K video content over a long-distance asynchronous network with no observable visual impairments.
- Support of at least 24 frames per second, 12-bit color depth and no subsampling (RGB).
- Video input and output in multiple HD-SDI channels.
- 10 Gb Ethernet network interface.
- Pixel-correct synchronization between picture quadrants.
- Audio synchronized with video.
- Small added latency (see below).

The rationale behind these requirements, in addition to those already mentioned in Section 1, is as follows.

The use of HD-SDI channels for transmission of high definition video is now a common industry practice. Three variants are currently in use — HD-SDI [7], dual-link HD-SDI [8] and 3G-SDI [9]. Mapping of digital video signals into these channels is specified for HD (1920 × 1080) [10,8], 2K (2048 × 1080, D-Cinema operational level 2 and 3) [8] and for other lower resolution formats. The 4K format (4096 × 2160, D-Cinema operational level 1 or 3840 × 2160, UHDTV1 [11]) is typically transferred in four quadrants, each in 2K or HD format carried over a separate HD-SDI channel.

Asynchronous Ethernet technology is currently more frequently deployed in 10 Gb/s networks than synchronous SONET/SDH, due to its simplicity and therefore lower cost. Ethernet will likely play even more important role in future 40 Gb/s and 100 Gb/s networks, although often enveloped in a synchronous Optical Transport Network (OTN) [12].

Empirical experience has shown that the maximum acceptable one-way latency for remote interactive work not perceived by users as a limiting factor is around 150 ms [6]. This latency can easily be caused just by network propagation delay. Therefore, the video transfer system should add minimal further latency. Buffering of one frame at 24 frames per second adds 41 ms.

3. Architecture

3.1. Hardware

Real-time processing of multi-gigabit data rates is difficult on PC-based platforms with standard operating systems not designed for real-time operation. We were looking for a real-time design that is scalable to higher data rates (such as for 8K or UHDTV2 format), higher network speeds (such as 40 and 100 Gb/s) and that can be integrated with commonly requested video processing functions, such as encryption, transcoding or compression. This implies highly parallel and truly real-time data paths. DSP (digital signal processor) and FPGA (field programmable gate arrays) are the standard technologies in this area. We selected FPGA, due to its high data bandwidth and our design having no requirements with regard to floating-point operations.

The selected FPGA circuit needs to have a sufficient number of fast channels for input and output of the HD-SDI data. The sustained speed needs to be 1.485 Gb/s for HD-SDI and 2.97 Gb/s for 3G-SDI. For the 4K format, we need four or eight HD-SDI interfaces depending on the exact format and interface speed. Xilinx, Altera and Lattice all have FPGA circuits that satisfy these requirements. We selected Xilinx Virtex 5, which allowed us to reuse some design blocks that we have developed for network monitoring devices.

The hardware architecture is shown in Fig. 1. The HD-SDI board converts electrical levels and timing between input and output HD-SDI channels on the one side and Virtex RocketIO channels on the other side. The FPGA board processes the video signal and is connected to an optical transceiver, which converts electrical and optical signals for network transmission.

This architecture allows operation in several modes illustrated in Fig. 2. Two devices can be used to transfer video content over a network or a single device can be used as a video processor or a network processor.

3.2. Packetization

There are several possible ways of mapping the HD-SDI data into network packets. Three options are described below. The resulting bit rate for 4K (all four quadrants) at the physical layer is summarized in Table 1. These rates also include embedded audio and the packet header overhead. We assume 24 frames per second as per the D-Cinema format.

Complete HD-SDI data.

One solution is to transfer all HD-SDI data. One complete line of the 2K frame (one quadrant) consists of 2750 samples. One