ST2110 NETWORK-ATTACHED DISPLAY FOR VIRTUALIZING PRODUCTION

Thomas True

NVIDIA Corporation, United States

ABSTRACT

GPU installations in post-production and broadcast facilities require video to be sent remotely conforming to the SMPTE ST2110 standards for plug-and-play device connectivity. Furthermore, in such facilities, workstations are now being virtualized and hosted in the data centre with what would have previously been a physically attached Serial Digital Video (SDI) reference display replaced by an IP network connection.

We present a SMPTE ST2110 network-attached reference display using commercial off-the-shelf (COTS) hardware that presents the GPU and NIC as a display on the computer workstation desktop. Colour-correct rendered video frames are transmitted directly from GPU device memory. Alignment of the GPU scan out as well as the synchronization of the audio, video and ancillary data streams is maintained by PTP.

NMOS integration permits the discovery and registration of the network-attached reference display as a source node on the media network such that it can be remotely managed and controlled by software greatly simplifying the administration compared with a physically connected display and associated cables. Content creation applications send ST2110 compliant video, audio and ancillary data flows on the media-network from a virtualized computer workstation desktop without the cost of directly integrating ST2110 transmission capabilities.

INTRODUCTION

In this paper we present an IP network-attached display based upon established industry standards that utilizes the combination of the Graphics Processing Unit (GPU) and Network Interface Card (NIC) to create a virtual display on the workstation's desktop. Application video frames rendered on the GPU are combined with associated audio data and transmitted on the IP network to a capable receiving device. This network-attached reference display provides a solution for virtualizing production since both the GPU and NIC can be virtualized.

MOTIVATION

GPU-based post-production applications require a standards conformant video reference display in order to provide a high-quality preview of the created material. The need for this reference display to be network attached is driven by technological trends impacting the content creation pipeline in both post-production and broadcast facilities.

Networked Media Systems

Limitations of the traditional SDI infrastructure along with advances in common of-the-shelf (COTS) hardware have moved broadcast and post-production facilities to adopt an all-IP infrastructure with media transport between devices based upon the SMPTE ST2110 suite of standards for the transport of digital media over IP networks 'SMPTE (1)' via Real-time Transport Protocol (RTP) based streams 'IETF (2)'. With this move to network-based media production systems, media content creation workflows are now software defined and resources can be quickly reconfigured to make maximum use of both hardware and software. This has led to the need to easily orchestrate data flows. The Networked Media Open Specifications (NMOS) is a collection of standards that provide an industry standard solution for interoperability and management of IP connected media devices 'AMWA (3)'.

GPU Virtualization

Furthermore, in post-production and broadcast facilities GPU-based workstations as networked media systems are being moved from the desktops of content creators into the data centre in order to provide a more secure, higher utilized and easier to maintain infrastructure. In this environment, assets never leave the confines of the data centre while a virtual user application experience is delivered via the IP network to one or more end users on a thin client running on a laptop or small form factor computer system. Traditional reference displays physically connected to the GPU with a SDI, HDMI or DP to the GPU do not typically work well in this environment due to lack of virtualization support for the video I/O interface or physical cabling constraints. As a result, no longer is there an SDI, HDMI, or DP connection, only a TCP/IP network connection to the desktop.

SOLUTION

Our solution pictured in Figure 1 creates a SMPTE ST2110 compatible network attached reference display. High-quality colour-accurate frames for display by the Windows Operating System display manager are rendered on the GPU prior to transmission by the NIC as a SMPTE ST2110-20 'SMPTE (4)' uncompressed video stream. Accompanying audio data is transferred to the NIC for transmission as a SMPTE ST2110-30 'SMPTE (5)' stream. IEEE 1588-2008 Precision Time Protocol (PTP) 'IEEE (6)' is used for synchronization while industry-standard Session Description Protocol (SDP) 'IETF (7)' manifests are used for configuration along with NMOS for device registration and control.

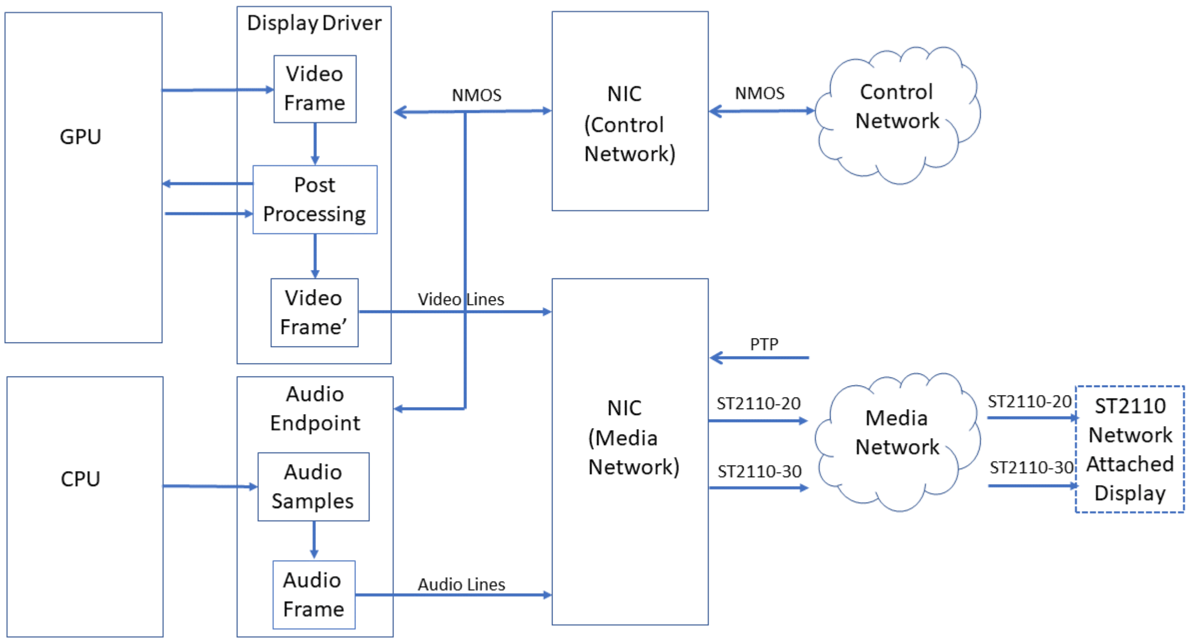

Figure 1 – Solution Architecture

Display Driver

The Display Driver receives the RGB pixel data of the rendered frame for the display in GPU device memory. Prior to packetization and ultimately network transmission, the RGB pixels are converted by the GPU from the native Windows operating system format to the sampling, depth and colorspace specified for the outgoing video stream in the SDP manifest. As a highly parallel floating-point processor with thousands of compute engines, the GPU can perform this conversion instantly on all the image pixels in parallel. The final video essence pixels are then transferred directly from GPU device memory to the NIC via Direct Memory Access (DMA) for transmission as RTP packet payload on the created SMPTE ST2110-20 stream. Meanwhile, the NIC DMAs the RTP packet headers from system memory. The granularity of each DMA transaction is a group of lines. For HD, the size of each DMA is 8 lines while for UHD it is 4 lines in order to stay within the standard UDP size limit of 1460 octets or bytes.

Audio Endpoint

Audio sample data is processed by a Windows operating system audio sink endpoint device. When this audio device is selected in the operating system control panel to be the active audio device, the audio endpoint receives the uncompressed 16- or 24-bit PCM audio samples in a host memory buffer. The NIC then DMA's this audio essence data as RTP packet payload along with the RTP headers for network transmission in the SMPTE ST2110-30 audio stream as specified in the SDP manifest.

SMPTE ST2110

Many TCP/HTTP-based solutions (NDI, MPEG Dash, HLS) exist for application streaming of compressed video and audio. However, these solutions do not support real-time transmission, multicast, frame-accurate switching or the uncompressed video streams required by professional workflows. For these reasons, we elected to use the standard SMPTE ST2110 framework for our network attached reference display. In the case of both video and audio the NIC implements the packet pacing that complies with the strict timing and traffic flow requirements of the SMPTE ST 2110-21 specification 'SMPTE (8)'.

Synchronization

Integration of our solution as a content generation node within today's all-IP post-production and broadcast facilities requires the audio and video essence streams to be synchronized via RTP timestamps aligned to the media clock 'SMPTE (9)'. As a synthetically generated essence, the RTP timestamp of each video frame is the time point at which the first video sample of the frame arrives at the NIC for transmission while for the audio stream, the RTP timestamp is a sample of the audio RTP clock when the first audio sample arrives at the NIC for transmission. RTP timestamps are samples of the RTP clock aligned with the media clock specified in the SDP as IEEE 1588-2008 Precision Time Protocol (PTP).

In order to implement hardware PTP and meet the strict timing requirements of the SMPTE ST2110-20 specification, our solution integrates the NVIDIA BlueField-2 Data Processing Unit (DPU) 'NVIDIA (10)'. This NIC has up to 8 onboard ARM cores on which we run ptp4l in a Docker container 'Docker (11)' synchronized to the PTP Grand Master. When the display and audio drivers query the current PTP time, the request then passes through to *ptp4l* running on the NIC. This overcomes two problems with the native PTP support on Windows in both bare metal and virtualized environments.

1) Lack of hardware timestamps.
2) System noise resulting in out-of-spec timing jitter.

Linux PTP *ptp4l* running on the DPU with hardware time stamps provides for accurate packet paced timing with time-triggered scheduling by synchronizing the PTP hardware clock on the NIC to the master clock while *phc2sys* syncs the system clock to the PTP hardware clock.

Control

Initial set up and control is provided via an industry standard SDP file. However, this is insufficient for network orchestration in a typical broadcast or post-production facility. For this reason, we have integrated NMOS IS-04 'AMWA (12)' support for device discover and registration and NMOS IS-05 'AMWA (13)' for connection management using the SONY Open-Source Software (OSS) NMOS C++ implementation 'SONY (14)' as a starting point. The Display Driver at start up registers itself as a NMOS Sender Node in the registry with an accompanying manifest that describes the network display output transport parameters. From that point onward, NMOS Receiver Nodes can connect to receive the outgoing video and audio stream for display or further downstream processing in the workflow.

BENEFITS

Zero Application Integration Time

Our solution uniquely combines the core GPU and NIC technologies to enable all post-production and broadcast applications with zero integration time by the application developer to stream high-quality colour-accurate synchronized rendered frames, along with accompanying audio, in real-time to a remote network-attached display device as compliant SMPTE ST2110 streams. This removes the need and complexity for application developers to integrate support for SMPTE ST2110 directly into applications or via transmit plugins.

Ease of Use

The network attached display is essentially connected to the desktop using the traditional desktop GUI and appears to the user as a traditional display device removing the need for any specific operator training. Under the hood the streaming Display Driver sends the GPU generated pixels through the NIC and onto the network as a SMPTE ST2110 stream.

Seamless Virtualization

The fact that our solution works exactly the same and provides the same user experience in both bare-metal and virtual GPU environments make it especially useful for remote high-quality reference displays. Figure 2 illustrates one use case of our SMPTE ST2110 network-attached reference display in a virtualized environment where a video editing application is running on a virtual GPU workstation hosted in a data centre server. Displays 1 and 2 are the standard remote workstation desktop displays provided by the remoting software stack where the application GUI with content bin, timeline and editing controls are displayed. Meanwhile, Display 3 is our high-quality colour and time accurate SMPTE ST2110 network-attached virtual reference display driven by the combination of the GPU and the NIC. On this display, the user sees the rendered final frame content from the virtualized application in full resolution with ultimate quality.

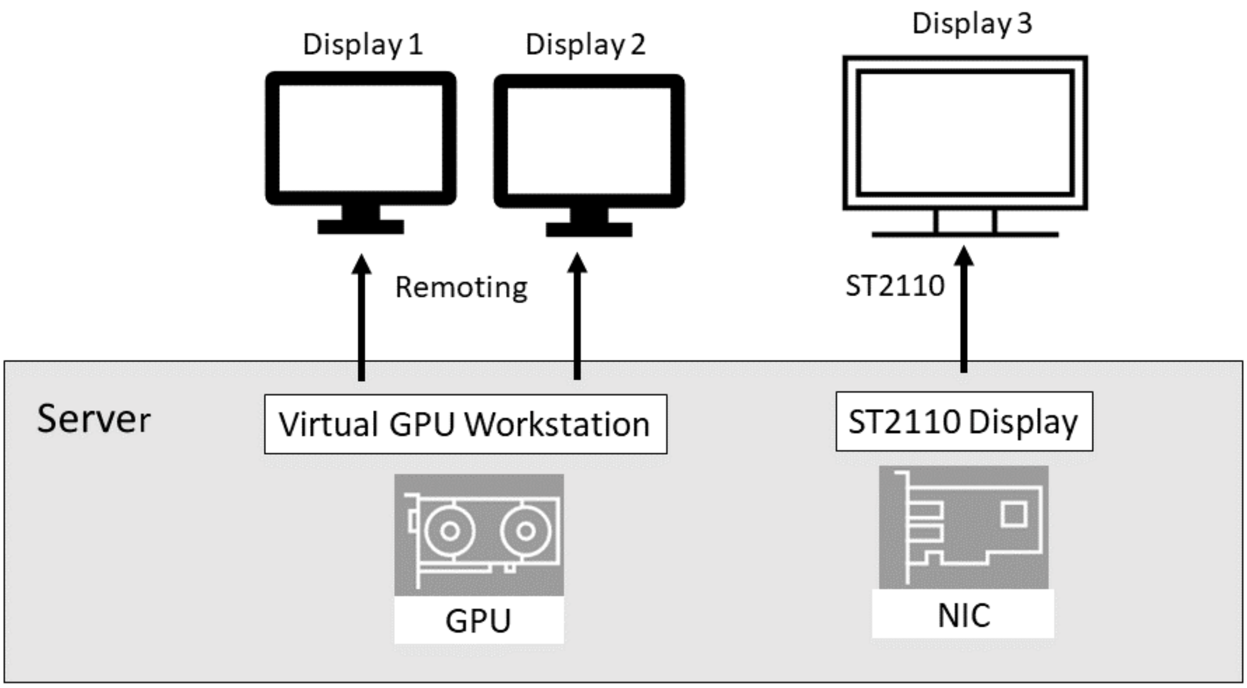

Figure 2 – ST2110 Virtualized Reference Display

FUTURE WORK

As we move forward with our implementation of a SMPTE ST2110 Network-Attached Display we plan to add the following capabilities to address additional needs of post-production and broadcast workflows.

Compressed Video Support

Our solution currently works well in a Local Area Network (LAN) environment when sufficient bandwidth is available. With the addition of SMPTE ST2110-22 'SMPTE (15)' compressed video support, our hope is that our solution will work in a Wide Area Network (WAN) environment between production facilities or to support remote workers.

Tighter GPU Integration

Tighter GPU integration will permit the support of deeper pixel formats and HDR along with more precise frame synchronization to the GPU. While currently the ST2110 video and audio output streams are tightly synchronized via PTP, the precise time at which the next frame to be displayed is rendered into the GPU frame buffer prior to network transmission is not.

ST2110-40 Ancillary Data Support

We see additional use cases in virtual production and other areas if support for the transmission of ST2110-40 ancillary data streams 'SMPTE (16)' is added.

NMOS IS-06 Support

For ultimate control and orchestration of the network attached display, we must add support for NMOS IS-06 'AMWA (17)' to provide network control. This will permit receiver nodes to change the parameters of the network attached display within the Windows operating system display and audio controls and have the changes reflected in the IP network video and audio streams.

REFERENCES

1. Society of Motion Picture and Television Engineers (SMPTE) 2018. SMPTE ST 2110 FAQ, https://www.smpte.org/smpte-st-2110-faq.
2. Internet Engineering Task Force (IETF) RFE 3550 RTP: A Transport Protocol for Real-Time Applications, https://www.ietf.org/rfc/rfc3551.txt.
3. Advanced Media Workflow Association (AMWA), March 2021, Networked Media Systems – the Big Picture -, https://static.amwa.tv/networked-media-systems-big-picture-2021-03-05.pdf.
4. Society of Motion Picture and Television Engineers (SMPTE), 2017. ST 2110-20:2017 SMPTE Standard – Professional Media Over Managed IP Networks: Uncompressed Active Video, https://ieeexplore.ieee.org/document/8165971.
5. Society of Motion Picture and Television Engineers (SMPTE), 2017. ST 2110-30:2017 SMPTE Standard – Professional Media Over Managed IP Networks: PCM Digital Audio, https://ieeexplore.ieee.org/document/8167392.

6. IEEE Instrumentation and Measurement Society 2008. IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems.
7. Internet Engineering Task Force (IETF), 2006. RFC 4566 SDP: Session Description Protocol, https://www.ietf.org/rfc/rfc4566.txt
8. Society of Motion Picture and Television Engineers (SMPTE), 2017. ST 2110-21:2017 SMPTE Standard – Professional Media Over Managed IP Networks: Traffic Shaping and Delivery Timing for Video, https://ieeexplore.ieee.org/document/8165971.
9. Society of Motion Picture and Television Engineers (SMPTE), 2017. ST 2110-10:2017 SMPTE Standard – Professional Media Over Managed IP Networks: System Timing and Definitions, https://ieeexplore.ieee.org/document/8165974.
10. NVIDIA Corporation, 2020. NVIDIA BLUEFIELD-2 DPU https://www.nvidia.com/content/dam/en-zz/Solutions/Data-Center/documents/datasheet-nvidia-bluefield-2-dpu.pdf
11. Docker, 2021. Docker Container Overview, https://docs.docker.com/get-started/overview/.
12. Advanced Media Workflow Association (AMWA), IS-04 NMOS Discovery and Registration Specification, https://specs.amwa.tv/is-04/.
13. Advanced Media Workflow Association(AMWA), IS-05 NMOS Device and Connection Management Specification, https://specs.amwa.tv/is-05/.
14. SONY Corporation, An NMOS C++ Implementation, https://github.com/sony/nmos-cpp.
15. Society of Motion Picture and Television Engineers (SMPTE), 2019. ST 2110-22:2019 SMPTE Standard – Professional Media Over Managed IP Networks: Constant Bit-Rate Compressed Video, https://ieeexplore.ieee.org/document/8809389.
16. Society of Motion Picture and Television Engineers (SMPTE), 2018. ST 2110-40:2018 SMPTE Standard – Professional Media Over Managed IP Networks: SMPTE ST 291-1 Ancillary Data, https://ieeexplore.ieee.org/document/8353279.
17. Advanced Media Workflow Association (AMWA), IS-06 NMOS Network Control Specification, https://specs.amwa.tv/is-06/.

ACKNOWLEDGEMENTS

The author would like to thank his colleagues for their contributions to this work. He would also like to thank the International Broadcasting Convention for permission to publish this paper.

.