A Distributed Stream Multiplexing Architecture for Multi-Chip Configuration beyond HDTV

Takayuki ONISHI†, Ken NAKAMURA†, Members, Takeshi YOSHITOME†, Nonmember, and Jiro NAGANUMA†, Member

SUMMARY This paper proposes a distributed stream multiplexing architecture for video codec LSIs with multi-chip configuration. This distributed architecture utilizes a built-in media multiplexing unit with an external stream input and inter-chip communication interfaces. Parallel protocol processing, with an autonomous inter-chip control mechanism to mix and concatenate packets through daisy-chained transfer paths, provides a complete multi-chip stream output at the end of the chain. Dispensing with external post-processing devices contributes to both high throughput and downsizing of high-end video codec systems. It is configurable for parallel encoding of super high-resolution video, multi-view/-angled HDTV vision and multiple HDTV programs. The architecture was successfully implemented in a fabricated single-chip MPEG-2 422P@HL codec LSI and utilized for the development of a super high-resolution video codec system.

Key words: super high-resolution video, multiplexing, codec LSI

1. Introduction

Recent progress in video and audio compression technology has made it possible to effectively compress video and audio with lesser drop in quality. The MPEG-2 standard [1] has been a driving force behind the development of digital high-definition television (HDTV), and the emerging H.264/AVC standard [2] has recently become popular for its better coding efficiency. In order to provide further advanced ultra-quality visual experiences toward next-generation services, some sensational applications are promising, such as super high-resolution videos beyond the HDTV level and multi-view/-angled HDTV systems providing clear 3-D vision.

For encoding and decoding of large scale data, parallel processing is considered to be a major solution. Parallel encoding methods for MPEG-2 video have been intensively studied [3], [4], where fragments of video images are transferred to multiple encoders and inter-chip information exchange is performed. Output concatenation and multiplexing, however, continue to be performed by external devices as illustrated in Fig. 1 (a). This is also the case for multi-view/-angled or multi-program systems as shown in Fig. 1 (b), an additional multiplexer (MUX) and de-multiplexer (DEMUX) is required. In actual video transmission beyond the HDTV level, final output stream reaches several hundred megabits per second and multiplexing devices take up as much space as an encoder itself [5], so that achieving higher throughputs and further downsizing are critical demands. Stream formats and levels for these extended applications are not yet specifically defined in video encoding standards, so that scalable and flexible configuration to various encoding modes and levels is also essential.

To achieve these requirements, we propose a new distributed transport stream multiplexing (TS-MUX) architecture [6] with daisy-chained inter-chip connection and master/slave task sharing. This architecture consists of each MPEG-2 MUX unit with an external TS input and inter-chip communication interfaces. TS packets are generated in parallel at each chip, then mixed or concatenated according to operation modes and states, through daisy-chained transfer paths. A master chip re-forms numbering and timestamping, providing a complete multi-chip output at the end of the chain. This makes the existing external post-processing devices dispensable, which contributes to not only downsizing but also high throughput. The operation is configurable for super high-resolution video, multi-view/-angled HDTV vision and multiple HDTV program encoding systems. The architecture can be applied to other codec sys-
tems using MPEG-2 TS regardless of video coding standards, such as the latest H.264/AVC encoders with higher compression performance. It was implemented in a fabricated single-chip MPEG-2 422P@HL codec LSI and a super high-resolution codec system using this chip was successfully developed.

2. Architecture

2.1 MUX Block Diagram

A block diagram of individual MUX is illustrated in Fig. 2. It is derived from our previous memory-based architecture for MPEG-2 system protocol LSIs [7], with some hardware modification to enhance MPEG-2 TS processing. Dedicated hardware units assist high throughput processing of standard MPEG-2 TS generation, while a RISC CPU dedicates itself to handling protocol extensions and additional requirements. When a codec’s encoding and decoding functions operate exclusively, these resources and interfaces of MUX/DEMUX are reconfigured and shared.

To realize the proposed architecture, an external TS input and inter-chip control interfaces are added for multi-chip extension, which are shown in Fig. 2 with dashed squares. They provide daisy-chained parallel operation of multiple MUXs, which behavior is described in the following sections.

2.2 Multi-Chip Approach

In prior parallel video encoding techniques, such as [4] for MPEG-2 HDTV encoding, video streams are concatenated per picture at the packetized elementary stream (PES) phase, as illustrated in Fig. 1 (a). With the proposed approach, in contrast, inter-chip concatenation and mixture is performed at the TS phase after the media multiplexing, so that it provides unified architecture together with the multi-program and multi-view/angled encoding, as shown in Fig. 1 (b).

Two operation modes are arranged to represent each of the configurations in Fig. 1 for various multi-chip applications, which are listed in Table 1. For encoding of super high-resolution video beyond HDTV level, when images are horizontally split but a complete sequential video stream is required, each chip’s output needs to be concatenated per picture, as illustrated in Fig. 3 (a). In this case, “concatenation mode” is appropriate. In contrast, as depicted in Fig. 3 (b), when images are split into multiple HDTV videos and transmitted in parallel to multiple HDTV decoders, then “mixture mode” can be used to transmit with multiple TS channels. For multi-program or multi-view/angled applications, streams are transmitted with multiple TS channels and “mixture mode” operates as a substitute for external TS-MUX.

2.3 Concatenation Mode

The multi-chip configuration for concatenation mode is illustrated in Fig. 4. TS I/O interfaces are connected between each neighboring two chips, and inter-chip controls are connected ring-wise to transfer a “token”. Each slave chip encodes a split picture in its charge and its MUX transforms the video PES into TS packets. The MUX in a master chip additionally handles audio/user data PES inputs, and is also responsible for complete multi-chip stream output.

State transition diagram of each slave MUX is shown in Fig. 5 (a). Each time, only one MUX which has a token is allowed to output self-generated video TS packets. The

<table>
<thead>
<tr>
<th>Table 1: Operation modes for multi-chip configuration.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application</td>
</tr>
<tr>
<td>Super high resolution video</td>
</tr>
<tr>
<td>Horizontally split, sequential output</td>
</tr>
<tr>
<td>Multiple HDTV split, parallel output</td>
</tr>
<tr>
<td>Multi-program</td>
</tr>
<tr>
<td>Multi-view/angled</td>
</tr>
<tr>
<td>Configuration</td>
</tr>
<tr>
<td>Concatenation mode</td>
</tr>
<tr>
<td>Mixture mode</td>
</tr>
</tbody>
</table>

![Fig. 2](image url) A MUX block diagram.

![Fig. 3](image url) Two split methods for super-high resolution video.
other slaves operate in "TS through state", simply relaying TS packets from the previous MUX to the next. When a certain slave gets a token from the previous chip, its MUX transits its state into "local output state". Here the encoded video PES is transferred to the MUX in constant bit rate from a shaper, until the end of one split picture. The MUX generates and sends TS packets, and when an "end of picture" signal comes from the internal encoder, the MUX sends the remaining video TS to flush the buffer, puts the token to the next chip, and returns to the TS through state. Padding the last video TS to flush leads to bit loss, the average of which can be calculated as:

$$\frac{\text{payload size}}{2} \times N \times \frac{\text{pictures per second}}{\text{pictures per second}}$$ (1)

where N is the number of chips. For 4-chip 30-fps systems, the average loss will be 8.8-kbps, which is almost negligible compared with the high video bit rate.

State transition diagram of the master MUX is illustrated in Fig.5 (b). When in TS through state, the scheduler interprets external TS as video TS and does the output scheduling. When the token comes, the scheduler switches the source and chooses the self-generated video TS, until the end of picture. Other TS types such as audio, user data, program specific information (PSI)/program clock reference (PCR) and null packets, are scheduled as normal. Since video TS packets generated in parallel have inconsistent serial numbers, a continuity counter field in each video TS is renumbered at the final stage.

As mentioned above, split video streams are concatenated at the TS phase, which achieves super high-resolution stream output without external devices. This method can also apply to other video coding standards, such as H.264/AVC, as long as each split image is encoded as a "slice" and concatenated afterwards. When decoding is also performed in parallel with multiple video decoders, each decoder simply extracts a video elementary stream (ES) of the picture region in its charge, and no extra mechanism is needed for the DEMUX.

2.4 Mixture Mode

For mixture mode, the multi-chip configuration and corresponding MUX operation are illustrated in Fig.6 and 7 respectively. Here a multi-program application is shown for instance, with each chip handling video, audio and user data of each program. Each slave produces video/audio/user data/PSI/PCR TS packets with individual program IDs (PIDs), but no null packets. The scheduler transmits these packets to the next chip.
TS packets, and also external TS packets from the previous chip whenever they arrive, to the next chip. The master is responsible for adding null packets. To avoid timing jitter that occurred during the transfer, shared PCR values are stamped at the final stage.

When encoders operate with super high-resolution video or multi-view/-angled vision such as MPEG-2 multi-view profile, each slave handles only video and external TS. Shapers which control the video PES output rate also need to be controlled according to the multi-chip rate control; however, other operations remain to be the same as mentioned above. At the decoding side, DEMUX of each decoder simply extracts the video TS of its charge and transfers them to internal video decoder, thus no extra function is required.

3. Implementation and Evaluation

The MUX architecture is included in the fabricated single-chip MPEG-2 422P@HL CODEC LSI [8] using 0.13-µm 8-level metal CMOS technology. A micrograph of the LSI is shown in Fig. 8. In the outlined area, 6% of the floor dimensions is allocated for MUX/DEMUX. The overhead implementing the new architecture is around 2% compared with the conventional MUX/DEMUX, which is almost negligible with a large downsizing advantage of overall systems. The maximum output rate is 270-Mbit/s through an 8-bit parallel TS output interface, and dispensing with low-speed external devices also contributes to adaptability to higher throughput.

We have developed an experimental super high-resolution encoder system [9] with the proposed architecture and successfully encodes 4K × 2K camera images in real time into a minimum 60-Mbps transport stream. Four encoder boards are installed inside as shown in Fig. 9, and each board processes a split image with inter-board bit rate allocation control. Encoder chip interfaces are interconnected with serial cables and MUX is performed in mixture mode as illustrated in Fig. 3 (b), since a super high resolution decoder system is feasibly built up with multiple HDTV decoder LSIs [8]. The overall super high resolution codec is installed in a 1-U (460 × 440 × 44-mm) chassis as illustrated in Fig. 10, to which downsizing the proposed architecture significantly contributes.

Figure 11 shows the bit rate of individual video and total TS, observed at the final stream output from the master. Individual video PES output rate is set to 30-Mbps which is
multiplexed into the final TS of 145-Mbps. Occasional sags in the bit rate is due to the bit rate control and occur when actual amount of video PES is less than expected, so that provision of video PES to the MUX is suspended and null TS packets are filled instead.

To check the uniformity of TS output, evaluation is performed with 6,000 TS packets described by the arrowed line in Fig. 11. Figure 12 shows the consecutiveness and interval of video TS packets that are produced at each encoder chip. Figure 12 (a) proves that no TS packets from the same encoder chip are located bumper-to-bumper in the final output. Figure 12 (b) illustrates that intervals of TS packets from the same encoder chip is uniformly distributed, proving that there are no unevenly distributed segments and all TS packets are uniformly mixed through the daisy-chained paths.

4. Conclusion

A distributed and scalable media multiplexing (MUX) architecture for multi-chip encoding systems was proposed. It consists of a MPEG-2 media multiplexing unit with an external transport stream input and inter-chip communication interfaces. Parallel protocol processing, together with an autonomous inter-chip control mechanism to mix and concatenate packets through daisy-chained transfer paths, provides a complete multi-chip output, without external stream handling devices. It realizes the downsizing and high throughput needed by the extended visual applications such as super high-resolution video beyond the HDTV level, multi-channel and multi-view HDTV systems. The evaluation with the fabricated single-chip MPEG-2 422P@HL CODEC LSI and the experimental super high-resolution encoder system showed a good distributed multiplexing performance with the multi-chip configuration. The implementation of the architecture will lead to a new dimension in high-quality, high-resolution digital image creation and circulation. In order to achieve higher throughput and more parallelism for more demanding application, inter-chip cooperative packet scheduling algorithm with faster interfaces will be required for the future.

References

Takayuki Onishi received the B.E. and M.E. degrees in information and communication engineering from the University of Tokyo, Japan, in 1997 and 1999, respectively. In 1999, he joined NTT Cyber Space Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Yokosuka, Japan, and has been engaged in research and development on high quality image coding and transmission. He is currently a research engineer at Visual Media Communications Project of NTT Cyber Space Laboratories, Kanagawa, Japan.

Ken Nakamura received the B.E. and M.E. degrees in science and technology from Keio University, Japan, in 1995 and 1997, respectively. In 1997, he joined NTT Human Interface Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Yokosuka, Japan, and has been engaged in research and development on high quality image coding and processing. He is currently a senior research engineer at Visual Media Communications Project of NTT Cyber Space Laboratories, Kanagawa, Japan.

Takeshi Yoshitome received the B.S. and M.S. degrees in computer science from Tsukuba University, Japan, in 1982 and 1984, respectively. In 1984, he joined the Electrical Communication Laboratories, Nippon Telegraph and Telephone Corporation (NTT), Kanagawa, Japan, and has been engaged in research and development of image processing systems. He is currently a senior research engineer at Visual Media Communications Project of NTT Cyber Space Laboratories, Kanagawa, Japan.

Jiro Naganuma received the B.S. degree in electrical engineering from Tokushima University, Tokushima, Japan, in 1981, and the Ph. D. degree in information science from Kyoto University, Kyoto, Japan, in 1995. In 1981, he joined Musashino Electrical Communication Laboratories, Nippon Telegraph and Telephone Public Corporation (NTT), Tokyo, Japan. He is currently a senior research engineer supervisor at Visual Media Communications Project of NTT Cyber Space Laboratories, Kanagawa, Japan. He is engaged in the research of high-level VLSI architectures and design methodologies for multimedia embedded system LSIs. He is a member of the the IPSJ, the IEEE, and the IEEE Computer Society.