

(19)



(11)

**EP 3 591 908 B9**

(12)

**CORRECTED EUROPEAN PATENT SPECIFICATION**

(15) Correction information:

**Corrected version no 1 (W1 B1)**  
**Corrections, see**  
**Claims EN 1, 4**

(51) International Patent Classification (IPC):

**H04L 12/70** <sup>(2013.01)</sup> **H04L 29/06** <sup>(2006.01)</sup>  
**H04N 1/36** <sup>(2006.01)</sup> **H04N 21/242** <sup>(2011.01)</sup>  
**H04W 56/00** <sup>(2009.01)</sup> **H04L 7/00** <sup>(2006.01)</sup>  
**H04N 21/43** <sup>(2011.01)</sup> **H04N 21/6437** <sup>(2011.01)</sup>

(48) Corrigendum issued on:

**06.04.2022 Bulletin 2022/14**

(52) Cooperative Patent Classification (CPC):

**H04L 65/608; H04N 21/43072; H04N 21/6437**

(45) Date of publication and mention  
of the grant of the patent:

**09.02.2022 Bulletin 2022/06**

(86) International application number:

**PCT/CN2017/077925**

(21) Application number: **17902312.2**

(87) International publication number:

**WO 2018/170852 (27.09.2018 Gazette 2018/39)**

(22) Date of filing: **23.03.2017**

(54) **METHOD AND DEVICE FOR LIP-SPEECH SYNCHRONIZATION AMONG MULTIPLE DEVICES**

VERFAHREN UND VORRICHTUNG ZUR LIPPENSYNCHRONISATION ZWISCHEN MEHREREN  
VORRICHTUNGEN

PROCÉDÉ ET DISPOSITIF DE SYNCHRONISATION LABIALE-PAROLE ENTRE PLUSIEURS  
DISPOSITIFS

(84) Designated Contracting States:

**AL AT BE BG CH CY CZ DE DK EE ES FI FR GB  
GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO  
PL PT RO RS SE SI SK SM TR**

(74) Representative: **Epping - Hermann - Fischer**

**Patentanwaltsgesellschaft mbH  
Schloßschmidstraße 5  
80639 München (DE)**

(43) Date of publication of application:

**08.01.2020 Bulletin 2020/02**

(56) References cited:

**CN-A- 101 123 611 CN-A- 101 179 484  
CN-A- 101 237 586 CN-A- 101 288 257  
US-A1- 2003 066 094 US-A1- 2005 177 643  
US-A1- 2008 304 571**

(73) Proprietor: **Huawei Technologies Co., Ltd.**

**Longgang District  
Shenzhen, Guangdong 518129 (CN)**

(72) Inventor: **SUN, Tao**

**Shenzhen  
Guangdong 518129 (CN)**

Note: Within nine months of the publication of the mention of the grant of the European patent in the European Patent Bulletin, any person may give notice to the European Patent Office of opposition to that patent, in accordance with the Implementing Regulations. Notice of opposition shall not be deemed to have been filed until the opposition fee has been paid. (Art. 99(1) European Patent Convention).

**EP 3 591 908 B9**

**Description****TECHNICAL FIELD**

5     **[0001]** Embodiments of the present invention relate to the field of communications technologies, and in particular, to a multi-device lip synchronization method and a device.

**BACKGROUND**

10    **[0002]** As Wi-Fi speakers and Bluetooth speakers are widely applied, an increasing quantity of users favor an application in which an audio and video playing device sends audio data to a Wi-Fi speaker or a Bluetooth speaker in a wireless connection manner, and then the audio and video playing device plays a video and the Wi-Fi speaker or the Bluetooth speaker plays audio. In this application, a most important issue is to ensure lip synchronization, so that a user can obtain good use experience.

15    **[0003]** In an existing process of using a Bluetooth speaker, a Bluetooth driver module in the audio and video playing device reads audio pulse code modulation (Pulse Code Modulation, PCM) data from an audio PCM buffer of the audio and video playing device, and sends the audio PCM data to the Bluetooth speaker in real time by using a Bluetooth protocol. A Bluetooth audio driver module in the Bluetooth speaker receives the PCM data according to the Bluetooth protocol, and sends the PCM data to an audio PCM buffer of the Bluetooth speaker in real time. A PCM play driver module in the Bluetooth speaker directly reads the audio PCM data from the audio PCM buffer of the Bluetooth speaker and plays and outputs the audio PCM data.

20    **[0004]** It may be learned that the Bluetooth speaker resolves only a problem of cross-device audio output, but cannot implement lip synchronization between devices. The same is true when the foregoing method is applied to a Wi-Fi speaker. In addition, for the Wi-Fi speaker, a forwarding delay of a Wi-Fi router further exacerbates asynchronization between audio output and video output. Therefore, how to implement multi-device lip synchronization is a problem that needs to be resolved.

25    **[0005]** US 2008/304571 A1 relates to a content receiving apparatus comprising: decoding means for receiving, from a content providing apparatus provided at an encoder side, a plurality of encoded video frames to which video time-stamps based on reference clock for the encoder side are attached and a plurality of encoded audio frames to which audio time-stamps based on a reference clock for the encoder side are attached, and for decoding the encoded video frames and the encoded audio frames; storing means for storing a plurality of video frames that the decoding means has obtained by decoding the encoded video frames and a plurality of audio frames that the decoding means has obtained by decoding the encoded audio frames; calculating means for calculating a time difference resulting from a gap between a clock frequency of a reference clock for the encoder side and a clock frequency of a system time clock for the decoder side; and timing-adjusting means for adjusting a timing of outputting the plurality of video frames, one by one, in accordance with the time difference and on the basis of a timing of outputting the plurality of audio frames, one by one.

**SUMMARY**

40    **[0006]** The invention is defined by the appended independent claims. Further detailed embodiments are defined in the dependent claims. The invention resolves the problem of how to implement lip synchronization when a plurality of devices synchronously play a video and audio.

45    **[0007]** According to a first aspect, an embodiment of the present invention provides a multi-device lip synchronization method, where the method is used by a primary device to synchronously output audio and a video to at least one secondary device, and includes: receiving, by the secondary device, a Real-Time Control Protocol RTCP packet sent by the primary device, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries a program clock reference PCR cyclically collected by the primary device, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet; correcting, by the secondary device, a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay; receiving, by the secondary device based on a multicast service address and port sent by the primary device, RTPs published by the primary device, splicing the RTPs into a complete audio data frame, obtaining, from a time stamp header field of the RTP packet, a presentation time stamp PTS corresponding to the audio data frame, and putting the audio data frame into an audio pulse code modulation PCM buffer of the secondary device; and outputting, by the secondary device, the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user.

55    **[0008]** Before the receiving, by the secondary device, a Real-Time Control Protocol RTCP packet sent by the primary

device, the method further includes: receiving, by the secondary device, media description information sent by the primary device, where the media description information includes a Simple Network Time Protocol SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency; and performing, by the secondary device, system time synchronization based on the SNTP service address and port; and

the receiving, by the secondary device, a Real-Time Control Protocol RTCP packet sent by the primary device includes: receiving, by the secondary device based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0009]** In a possible design, the RTCP packet carries an RTCP session identifier, and before the correcting, by the secondary device, a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay, the method further includes: sending, by the secondary device, an RTCP session join request to the primary device, so that the primary device sends the RTCP session identifier to the secondary device.

**[0010]** In a possible design, the correcting, by the secondary device, a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay includes: calculating, by the secondary device, an STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

where

$scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and

correcting, by the secondary device, the STC of the secondary device based on the calculated STC correction value.

**[0011]** According to a second aspect, an embodiment of the present invention provides a multi-device lip synchronization method, where the method is used by a primary device to synchronously output audio and a video to at least one secondary device, and includes: collecting, by the primary device, a program clock reference PCR based on a preset collection cycle; sending, by the primary device, a Real-Time Control Protocol RTCP packet to the secondary device when determining that a preset condition is met, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device is able to correct a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay; and copying, by the primary device, an audio data frame from an audio pulse code modulation PCM buffer of the primary device, packing the audio data frame into real-time streaming packets RTPs, and publishing the RTPs to a multicast service address and port, where a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user. Before the collecting, by the primary device, a program clock reference PCR based on a preset collection cycle, the method further includes: sending, by the primary device, media description information to the secondary device, where the media description information includes a Simple Network Time Protocol SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0012]** In a possible design, the RTCP packet carries an RTCP session identifier, and before the sending, by the primary device, a Real-Time Control Protocol RTCP packet to the secondary device when determining that a preset condition is met, the method further includes: receiving, by the primary device, an RTCP session join request sent by the secondary device, and sending the RTCP session identifier to the secondary device.

**[0013]** In a possible design, the preset condition is: a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

**[0014]** In a possible design, the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr-scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

or

$$(scr\_curr-scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

where  $scr\_curr$  is a currently collected clock value,  $scr\_last$  is a previously collected clock value,  $scf\_srv$  is the program clock frequency of the primary device, and  $cycle\_read\_x$  is the preset collection cycle.

**[0015]** According to a third aspect, an embodiment of the present invention provides a secondary device, including: a first receiving module, configured to receive a Real-Time Control Protocol RTCP packet sent by a primary device, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries a program clock reference PCR cyclically collected by the primary device, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet; a correction module, configured to correct a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay; a second receiving module, configured to receive, based on a multicast service address and port sent by the primary device, RTPs published by the primary device; a processing module, configured to: splice the RTPs into a complete audio data frame, obtain, from a time stamp header field of the RTP packet, a presentation time stamp PTS corresponding to the audio data frame, and put the audio data frame into an audio pulse code modulation PCM buffer of the secondary device; and an output module, configured to output the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output.

**[0016]** This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user. The first receiving module is further configured to: before receiving the Real-Time Control Protocol RTCP packet sent by the primary device, receive media description information sent by the primary device, where the media description information includes a Simple Network Time Protocol SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency; and the processing module is further configured to perform system time synchronization based on the SNTP service address and port; and the first receiving module is specifically configured to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0017]** In a possible design, the RTCP packet carries an RTCP session identifier, and the secondary device further includes: a sending module, configured to send an RTCP session join request to the primary device before the correction module corrects the system clock STC of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay, so that the primary device sends the RTCP session identifier to the secondary device.

**[0018]** In a possible design, the correction module is specifically configured to: calculate an STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

where

$scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and correct the STC of the secondary device based on the calculated STC correction value.

**[0019]** According to a fourth aspect, an embodiment of the present invention provides a system composed by a primary device and at least one secondary device, wherein the primary device includes: a collection module, configured to collect a program clock reference PCR based on a preset collection cycle; a sending module, configured to send a Real-Time Control Protocol RTCP packet to a secondary device when it is determined that a preset condition is met, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device corrects a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay; and a processing module, configured to: copy

an audio data frame from an audio pulse code modulation PCM buffer of the primary device, pack the audio data frame into real-time streaming packets RTPs, and publish the RTPs to a multicast service address and port, where a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user. The sending module is further configured to: send media description information to the secondary device before the collection module collects the program clock reference PCR based on the preset collection cycle, where the media description information includes a Simple Network Time Protocol SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0020]** In a possible design, the RTCP packet carries an RTCP session identifier, and the primary device further includes: a receiving module, configured to: before the sending module sends the Real-Time Control Protocol RTCP packet to the secondary device when determining that the preset condition is met, receive an RTCP session join request sent by the secondary device, and send the RTCP session identifier to the secondary device.

**[0021]** In a possible design, the preset condition is: a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

**[0022]** In a possible design, the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

or

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

where  $scr\_curr$  is a currently collected clock value,  $scr\_last$  is a previously collected clock value,  $scf\_srv$  is the program clock frequency of the primary device, and  $cycle\_read\_x$  is the preset collection cycle.

**[0023]** An embodiment of the present invention provides a terminal, including a receiver, a processor, and a transmitter. The receiver is configured to receive data from outside of the terminal. The transmitter is configured to send data to an external device. The processor is configured to perform the method according to the first aspect or the second aspect.

An embodiment of the present invention further provides a computer storage medium, configured to store a computer software instruction used by the secondary device or the primary device in any of the foregoing aspects, and the computer software instruction includes a method or a program designed for performing the foregoing aspect.

**[0024]** An embodiment of the present invention further provides a data processing system, including a module configured to perform the method according to the first aspect or the second aspect.

**[0025]** An embodiment of the present invention further provides a computer program, configured to perform the method according to the first aspect or the second aspect.

## BRIEF DESCRIPTION OF DRAWINGS

**[0026]**

FIG. 1 is a schematic flowchart of Embodiment 1 of a multi-device lip synchronization method according to an embodiment of the present invention;

FIG. 2 is a schematic flowchart of Embodiment 2 of a multi-device lip synchronization method according to an embodiment of the present invention;

FIG. 3 is a schematic diagram of a principle of a multi-device lip synchronization method according to an embodiment of the present invention;

FIG. 4 is a flowchart of interaction between modules when a primary device enables a media sharing service;

FIG. 5 is a flowchart of interaction between modules when a secondary device queries a shared media resource;

FIG. 6 is a flowchart of interaction between modules when a secondary device synchronously plays a shared media resource;

FIG. 7 is a schematic structural diagram of Embodiment 1 of a secondary device according to an embodiment of the present invention;

FIG. 8 is a schematic structural diagram of Embodiment 2 of a secondary device according to an embodiment of the present invention;

FIG. 9 is a schematic structural diagram of Embodiment 1 of a primary device according to an embodiment of the present invention;

FIG. 10 is a schematic structural diagram of Embodiment 2 of a primary device according to an embodiment of the present invention;

FIG. 11 is a schematic structural diagram of Embodiment 3 of a secondary device according to an embodiment of the present invention;

FIG. 12 is a schematic structural diagram of Embodiment 3 of a primary device according to an embodiment of the present invention; and

FIG. 13 is a schematic structural diagram of Embodiment 4 of a primary device according to an embodiment of the present invention.

## DESCRIPTION OF EMBODIMENTS

**[0027]** The technical solutions in the embodiments of the present invention may be applied to various communications systems in a wireless cellular network, for example, a Global System for Mobile Communications (Global System of Mobile communication, GSM), a Code Division Multiple Access (Code Division Multiple Access, CDMA) system, a Wideband Code Division Multiple Access (Wideband Code Division Multiple Access Wireless, WCDMA) system, a general packet radio service (General Packet Radio Service, GPRS) system, an LTE system, and a Universal Mobile Telecommunications System (Universal Mobile Telecommunications System, UMTS). This is not limited in the embodiments of the present invention.

**[0028]** The technical solutions in the embodiments of the present invention are mainly applied to a scenario of how to ensure lip synchronization when a primary device and one or more secondary devices synchronously play video data and audio data. Lip synchronization means that video data played by the primary device synchronizes with audio data played by the secondary device, so that a playing effect that a lip movement synchronizes with voice can be achieved when a user is watching a video. The primary device is connected to the secondary device through Wi-Fi or in another manner. The primary device plays video data and the secondary device plays audio data. The primary device supports audio and video output synchronization, such as a mobile phone, a set-top box (Set-Top Box, STB), or a television box (Over The Top, OTT). The secondary device may be a device such as a Wi-Fi speaker, and the secondary device can play audio data.

**[0029]** According to a multi-device lip synchronization method and a device that are provided in the embodiments of the present invention, one primary device can drive a plurality of secondary devices, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary devices, and provides good use experience for a user. The technical solutions provided in the embodiments of the present invention are described in detail below with reference to the accompanying drawings.

**[0030]** FIG. 1 is a schematic flowchart of Embodiment 1 of a multi-device lip synchronization method according to an embodiment of the present invention. The method is used by a primary device to synchronously output audio and a video to at least one secondary device, and as shown in FIG. 1, includes the following steps.

**[0031]** S101. The primary device collects a program clock reference (Program Clock Reference, PCR) based on a preset collection cycle.

**[0032]** S102. The primary device determines, based on the collected PCR, whether a preset condition is met, and sends a Real-Time Control Protocol (Real-Time Control Protocol, RTCP) packet to the secondary device when the preset condition is met, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol (Network Time protocol, NTP) field in the RTCP packet carries a sending time point of the RTCP packet. The preset condition is: a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold. The preset threshold is, for example, a value ranging from -60 ms to 20 ms. 20 ms is used as an example, and the preset condition is:

$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} < (\text{cycle\_read\_x} - 20);$$

; or

$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} > (\text{cycle\_read\_x} + 20),$$

where scr\_curr is a currently collected clock value, scr\_last is a previously collected clock value, scf\_srv is a program

clock frequency of the primary device, and cycle read x is the preset collection cycle.

**[0033]** S103. The secondary device receives the RTCP packet sent by the primary device, and the secondary device corrects a system clock (System Time Clock, STC) of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay.

**[0034]** Because the real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and the NTP field in the RTCP packet carries the sending time point of the RTCP packet, after the secondary device receives the RTCP packet, the secondary device may obtain the PCR, the sending time point of the RTCP packet, and a moment at which the RTCP packet is received. The RTCP delay is an RTCP transmission delay, and is a difference between the sending time point of the RTCP packet and the moment at which the RTCP packet is received. The RTCP packet carries an RTCP session identifier, and the RTCP session identifier is an identifier allocated by the primary device to distinguish between different secondary devices. Specifically, the secondary device sends an RTCP session join request to the primary device, and the primary device sends the RTCP session identifier to the secondary device after receiving the RTCP session join request. Then the RTCP session identifier is added to the RTCP packet sent by the primary device to the secondary device.

**[0035]** Specifically, the secondary device may calculate an STC correction value scr\_correct of the secondary device according to the following formula:

$$\text{scr\_correct} = \text{scr\_srv} \times \text{scf\_clt} / \text{scf\_srv} + (\text{ntp\_rcv} - \text{ntp\_snd}) \times \text{scf\_clt} / 1000,$$

where scf\_clt is the program clock frequency of the secondary device, scf\_srv is the program clock frequency of the primary device, ntp\_rcv is the moment at which the RTCP packet is received, and scr\_srv and ntp\_snd are sending time points of the PCR and the RTCP packet, respectively.

**[0036]** Then the secondary device corrects the STC of the secondary device based on the calculated STC correction value.

**[0037]** S104. The primary device copies an audio data frame from an audio pulse code modulation (Pulse Code Modulation, PCM) buffer of the primary device, packs the audio data frame into real-time streaming packets (Real Time Packet, RTP), and publishes the RTPs to a multicast service address and port, where a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame.

**[0038]** S105. The secondary device receives, based on the multicast service address and port sent by the primary device, the RTPs published by the primary device, splices the RTPs into the complete audio data frame, obtains, from the time stamp header field of the RTP packet, the presentation time stamp corresponding to the audio data frame, and puts the audio data frame into a PCM buffer of the secondary device. The audio data frame is, for example, an audio frame.

**[0039]** S106. The secondary device outputs the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame. Specifically, the secondary device synchronously plays the audio data frame in the audio PCM buffer based on the presentation time stamp of the audio data frame with reference to the STC of the secondary device.

**[0040]** According to the multi-device lip synchronization method provided in this embodiment, the primary device collects the PCR based on the preset cycle, and sends the RTCP packet to the secondary device when determining that the preset condition is met (in other words, time correction needs to be performed), where the RTCP packet carries the collected PCR and the sending time point of the RTCP packet. The secondary device corrects the STC of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay. Then the secondary device receives, based on the multicast service address and port sent by the primary device, the RTPs published by the primary device, splices the RTPs into the complete audio data frame, obtains, from the time stamp header field of the RTP packet, the presentation time stamp corresponding to the audio data frame, and puts the audio data frame into the PCM buffer of the secondary device. Finally, the secondary device outputs the audio data frame in the audio PCM buffer with reference to the STC of the secondary device and the presentation time stamp of the audio data frame. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user.

**[0041]** FIG. 2 is a schematic flowchart of Embodiment 2 of a multi-device lip synchronization method according to an embodiment of the present invention. As shown in FIG. 2, the method in this embodiment is based on the embodiment shown in FIG. 1. If it is not a first time for the secondary device and the primary device to synchronously output audio data and video data, after audio data and video data are synchronously output for the first time, the program clock frequency of the primary device may be stored in the secondary device, and a corresponding identifier of the primary device may also be stored, so that the program clock frequency of the primary device and the identifier of the primary device may be directly used for synchronous output performed for a second time. If it is the first time for the secondary device and the primary device to synchronously output audio data and video data, the primary device needs to send the

program clock frequency of the primary device to the secondary device before S103. Likewise, if it is not the first time for the secondary device and the primary device to synchronously output audio data and video data, after audio data and video data are synchronously output for the first time, a Simple Network Time Protocol (Simple Network Time protocol, SNTP) service address and port used by the secondary device to perform system time synchronization, an RTCP service address and port used by the secondary device to receive the RTCP packet, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency may be stored in the secondary device, and a corresponding identifier of the primary device may also be stored, so that the SNTP service address and port, the RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, the media format, the time stamp frequency, and the identifier of the primary device may be directly used for synchronous output performed for a second time. If it is the first time for the secondary device and the primary device to synchronously output audio data and video data, before S101, the method may further include the following steps.

**[0042]** S107. The primary device sends media description information to the secondary device, where the media description information includes an SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency. Specifically, the primary device may publish the media description information by using the Hypertext Transfer Protocol.

**[0043]** S108. The secondary device receives the media description information sent by the primary device. The secondary device obtains the SNTP service address and port, the RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, the media format, and the time stamp frequency by parsing the media description information, performs system time synchronization based on the SNTP service address and port, and then receives, based on the RTCP service address and port, the RTCP packet sent by the primary device. The multicast service address and port are used by the secondary device to receive the RTPs published by the primary device. The media format is a format used by the primary device to perform compression coding on the audio data frame, and is used to by the secondary device to determine a corresponding decoding format. The time stamp frequency is used by the secondary device to adjust a time stamp between two frames of the secondary device based on a time stamp between two frames of the primary device.

**[0044]** A subsequent procedure is the same as that shown in FIG. 1. Details are not described herein again.

**[0045]** According to the multi-device lip synchronization method provided in this embodiment, the primary device sends the media description information to the secondary device, and the secondary device performs system time synchronization based on the SNTP service address and port in the media description information. Then the primary device collects the PCR according to the preset cycle, and sends the RTCP packet to the secondary device when determining that the preset condition is met (in other words, time correction needs to be performed), where the RTCP packet carries the collected PCR and the sending time point of the RTCP packet. The secondary device corrects the STC of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay. Then the secondary device receives, based on the multicast service address and port sent by the primary device, the RTPs published by the primary device, splices the RTPs into the complete audio data frame, obtains, from the time stamp header field of the RTP packet, the presentation time stamp corresponding to the audio data frame, and puts the audio data frame into the PCM buffer of the secondary device. Finally, the secondary device outputs the audio data frame in the audio PCM buffer with reference to the STC of the secondary device and the presentation time stamp of the audio data frame. Therefore, the primary device drives the secondary device, so that audio data and video data are synchronously output. This implements lip synchronization between the primary device and the secondary device, and provides good use experience for a user.

**[0046]** A schematic diagram of a principle of audio output from the primary device to the secondary device is described in detail below with reference to the accompanying drawings, module composition inside the primary device and the secondary device, and an interaction process between modules.

**[0047]** FIG. 3 is a schematic diagram of a principle of a multi-device lip synchronization method according to an embodiment of the present invention. As shown in FIG. 3, in this embodiment of the present invention, for a primary device, a media information publishing module, a system time server, a program clock server, and an RTP packet server are newly added to an existing primary device (a media player), and for a secondary device, a media information downloading module, a system time client, a program clock client, an RTP packet client, and a program clock driver module are newly added to an existing secondary device. The media information publishing module, the system time server, the program clock server, the RTP packet server, the media information downloading module, the system time client, the program clock client, the RTP packet client, and the program clock driver module each may be a software module. The media information publishing module is configured to publish media description information, where the media description information includes an SNTP service address and port, an RTCP service address and port, a multicast service address and port, a program clock frequency of the primary device, a media format, and a time stamp frequency. The media information downloading module is configured to receive the media description information. The system time server and the system time client implement time synchronization between the secondary device and the primary device.



The program clock server and the program clock client implement high-precision synchronization between program clocks of the secondary device and the primary device, for example, ensure that a program clock deviation is less than 20 ms (about a 1-frame PCM output delay). When the program clock client cannot modulate a clock frequency to match the program clock server, corresponding clock frequency conversion needs to be performed on an SCR and a PTS. The RTP packet server and the RTP packet client establish an RTP channel for transmitting an audio data frame from the primary device to the secondary device. The program clock driver module of the secondary device is configured to perform output with reference to a clock of the program clock client when a PCM play driver module plays audio data.

**[0048]** It should be noted that as shown in FIG. 3, after the secondary device is connected to the primary device, an original PCM play driver module of the primary device may be disconnected and not used, and the secondary device plays synchronously output audio. After the secondary device is connected to the primary device, an original audio target decoder and an original audio PCM buffer (FIFO) of the secondary device are also disconnected, and the audio output by the primary device is played.

**[0049]** Specifically, an audio and video synchronization process of the primary device is as follows: An audio and video synchronization module of the primary device calculates a program clock reference value according to a synchronization policy (a PCR reference, a video stream reference, an audio stream reference, or an audio and video reference) of a target program and a lip synchronization algorithm and based on a PCR, an audio presentation time stamp (Presentation Time Stamp, PTS), and a video PTS that are input. Then the audio and video synchronization module compares the program clock reference value with a current system clock sample (SCR). When an offset exceeds a threshold (which is usually 100 ms to 200 ms), the program clock reference value is used to correct a system clock to obtain a corrected PCR. During output of a media stream such as audio or a video, a display output time sequence of the audio or the video is arranged based on a comparison between a PTS and the SCR. For example, when the PTS is less than or equal to the current sampling SCR, a current audio frame is output; or if the PTS is greater than the current sampling SCR, no audio frame is output. After the primary device is connected to the secondary device, a video is normally output, and audio is not played by using the PCM play driver module, but is output to the secondary device by using the RTP packet server. Then the secondary device synchronously outputs the audio.

**[0050]** FIG. 4 is a flowchart of interaction between modules when a primary device enables a media sharing service. As shown in FIG. 4, the following steps are included.

**[0051]** S201. After a monitoring module of the primary device is connected to a secondary device or finds a secondary device, the monitoring module enables a media sharing service procedure, including: enabling system time synchronization (by a system time server), enabling program clock synchronization (by a program clock server), enabling multicast (by an RTP packet server), and enabling media information publishing (by a media information publishing module).

**[0052]** S202. The media information publishing module publishes media description information, where the media description information includes an SNTP service address and port, an RTCP service address and port, a multicast service address and port, a program clock frequency of the primary device, a media format, and a time stamp frequency.

**[0053]** Then the system time server performs system time synchronization based on the SNTP service address and port. The program clock server receives, based on the RTCP service address and port, an RTCP packet sent by the primary device, and performs program clock synchronization based on the RTCP packet. The RTP packet server publishes RTPs based on the multicast service address and port.

**[0054]** FIG. 5 is a flowchart of interaction between modules when a secondary device queries a shared media resource. As shown in FIG. 5, the following steps are included.

**[0055]** S301. A monitoring module of the secondary device sends a request for querying server media information to a media information downloading module.

**[0056]** S302. The media information downloading module sends a media description information request to a media information publishing module, and the media information publishing module returns a media description information request response that carries media description information, where the media description information includes an SNTP service address and port, an RTCP service address and port, a multicast service address and port, a program clock frequency of a primary device, a media format, and a time stamp frequency.

**[0057]** S303. The media information downloading module parses the media description information to obtain all service addresses and ports, the program clock frequency of the primary device, the media format, and the time stamp frequency.

**[0058]** S304. The monitoring module of the secondary device configures the SNTP service address and port, and starts a time client to perform SNTP time synchronization.

**[0059]** Finally, the time client and a system time server perform time synchronization.

**[0060]** FIG. 6 is a flowchart of interaction between modules when a secondary device synchronously plays a shared media resource. As shown in FIG. 6, the following steps are included.

**[0061]** S401. A monitoring module of the secondary device sends a program clock synchronization request to a program clock client.

**[0062]** S402. The program clock client sends an RTCP session join request to a program clock server, and the program clock server returns an RTCP session identifier.

**[0063]** S403. The program clock server collects a PCR based on a preset collection cycle, and sends a PCR sampling value to the program clock client when a preset condition is met.

**[0064]** S404. The program clock client corrects an STC of the secondary device based on the PCR, a program clock frequency of a primary device, a program clock frequency of the secondary device, and an RTCP delay.

**[0065]** S405. The monitoring module of the secondary device sends a request for starting to receive an audio data frame to an RTP packet client.

**[0066]** S406. The RTP packet client sends a multicast group join request to an RTP packet server.

**[0067]** S407. The RTP packet server sends RTPs to the RTP packet client.

**[0068]** S408. The RTP packet client receives the RTPs, splices the RTPs into a complete audio data frame, obtains a presentation time stamp corresponding to the audio data frame from a time stamp header field of the RTP packet, and puts the audio data frame into a PCM buffer of the secondary device.

**[0069]** S409. A PCM play driver module outputs the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame.

**[0070]** In the embodiments of the present invention, function modules of the primary device and the secondary device may be obtained through division based on the foregoing method examples. For example, each function module may be obtained through division based on each function, or two or more functions may be integrated into one processing module. The integrated module may be implemented in a form of hardware, or may be implemented in a form of a software function module. It should be noted that the module division in the embodiments of the present invention is an example, and is merely logical function division. There may be another division manner in actual implementation.

**[0071]** FIG. 7 is a schematic structural diagram of Embodiment 1 of a secondary device according to an embodiment of the present invention. As shown in FIG. 7, the secondary device in this embodiment may include a first receiving module 11, a correction module 12, a second receiving module 13, a processing module 14, and an output module 15. The first receiving module 11 is configured to receive an RTCP packet sent by a primary device, where an RTP time stamp header field in the RTCP packet carries a program clock reference PCR cyclically collected by the primary device, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet. The correction module 12 is configured to correct a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay. The second receiving module 13 is configured to receive, based on a multicast service address and port sent by the primary device, RTPs published by the primary device. The processing module 14 is configured to: splice the RTPs into a complete audio data frame, obtain, from a time stamp header field of the RTP packet, a presentation time stamp PTS corresponding to the audio data frame, and put the audio data frame into an audio pulse code modulation PCM buffer of the secondary device. The output module 15 is configured to output the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame.

**[0072]** FIG. 8 is a schematic structural diagram of Embodiment 2 of a secondary device according to an embodiment of the present invention. As shown in FIG. 8, based on the secondary device shown in FIG. 7, further, the secondary device in this embodiment may further include a sending module 16. The sending module 16 is configured to send an RTCP session join request to the primary device before the correction module 12 corrects the system clock STC of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay, so that the primary device sends an RTCP session identifier to the secondary device.

**[0073]** Further, the correction module 12 is specifically configured to: calculate an STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

where  $scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and correct the STC of the secondary device based on the calculated STC correction value.

**[0074]** Further, the first receiving module 11 is further configured to: before receiving the Real-Time Control Protocol RTCP packet sent by the primary device, receive media description information sent by the primary device, where the media description information includes an SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency. The processing module 14 is further configured to perform system time synchronization based on the SNTP service address and port. The first receiving module 11 is specifically configured to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0075]** The secondary device in the embodiment shown in FIG. 7 or FIG. 8 may be configured to execute the technical

solution of the method embodiment shown in FIG. 1 or FIG. 2, and an implementation principle and a technical effect thereof are similar and are not described herein again.

**[0076]** FIG. 9 is a schematic structural diagram of Embodiment 1 of a primary device according to an embodiment of the present invention. As shown in FIG. 9, the primary device in this embodiment may include a collection module 21, a sending module 22, and a processing module 23. The collection module 21 is configured to collect a program clock reference PCR based on a preset collection cycle. The sending module 22 is configured to send a Real-Time Control Protocol RTCP packet to a secondary device when it is determined that a preset condition is met, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device corrects a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay. The processing module 23 is configured to: copy an audio data frame from an audio pulse code modulation PCM buffer of the primary device, pack the audio data frame into real-time streaming packets RTPs, and publish the RTPs to a multicast service address and port, where a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port.

**[0077]** FIG. 10 is a schematic structural diagram of Embodiment 2 of a primary device according to an embodiment of the present invention. As shown in FIG. 10, based on the primary device shown in FIG. 9, further, the primary device in this embodiment may further include a receiving module 24. The receiving module 24 is configured to: before the sending module 22 sends the RTCP packet to the secondary device when it is determined that the preset condition is met, receive an RTCP session join request sent by the secondary device, and send an RTCP session identifier to the secondary device.

**[0078]** Further, the sending module 22 is further configured to: send media description information to the secondary device before the collection module 21 collects the program clock reference PCR based on the preset collection cycle, where the media description information includes an SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0079]** The preset condition is: a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

**[0080]** Optionally, the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

or

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

where  $scr\_curr$  is a currently collected clock value,  $scr\_last$  is a previously collected clock value,  $scf\_srv$  is the program clock frequency of the primary device, and  $cycle\_read\_x$  is the preset collection cycle.

**[0081]** The primary device in the embodiments shown in FIG. 9 and FIG. 10 may be configured to execute the technical solution of the method embodiment shown in FIG. 1 or FIG. 2, and an implementation principle and a technical effect thereof are similar and are not described herein again.

**[0082]** FIG. 11 is a schematic structural diagram of Embodiment 3 of a secondary device according to an embodiment of the present invention. As shown in FIG. 11, the secondary device in this embodiment may include a receiver 31, a processor 32, and a transmitter 33. The receiver 31 is configured to receive an RTCP packet sent by a primary device, where an RTP time stamp header field in the RTCP packet carries a program clock reference PCR cyclically collected by the primary device, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet. The processor 32 is configured to correct a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay. The receiver 31 is further configured to receive, based on a multicast service address and port sent by the primary device, RTPs published by the primary device. The processor 32 is further configured to: splice the RTPs into a complete audio data frame, obtain, from a time stamp header field of the RTP packet, a presentation time stamp PTS corresponding to the audio data frame, and put the audio data frame into an audio pulse code modulation PCM buffer of the secondary device. The transmitter 33 is configured to output the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame.

**[0083]** Further, the transmitter 33 is further configured to send an RTCP session join request to the primary device before the correction module corrects the system clock STC of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay, so that the primary device sends an RTCP session identifier to the secondary device.

**[0084]** Further, the processor 32 is specifically configured to: calculate an STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

where  $scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and correct the STC of the secondary device based on the calculated STC correction value.

**[0085]** Further, the receiver 31 is further configured to: before receiving the Real-Time Control Protocol RTCP packet sent by the primary device, receive media description information sent by the primary device, where the media description information includes an SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency. The processor 32 is further configured to perform system time synchronization based on the SNTP service address and port. The receiver 31 is specifically configured to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0086]** The secondary device in this embodiment may be configured to execute the technical solution of the method embodiment shown in FIG. 1 or FIG. 2, and an implementation principle and a technical effect thereof are similar and are not described herein again.

**[0087]** FIG. 12 is a schematic structural diagram of Embodiment 3 of a primary device according to an embodiment of the present invention. As shown in FIG. 12, the primary device in this embodiment may include a processor 41 and a transmitter 42. The processor 41 is configured to collect a program clock reference PCR based on a preset collection cycle. The transmitter 42 is configured to send a Real-Time Control Protocol RTCP packet to a secondary device when it is determined that a preset condition is met, where a real-time streaming packet RTP time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol NTP field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device corrects a system clock STC of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and an RTCP delay. The processor 41 is further configured to: copy an audio data frame from an audio pulse code modulation PCM buffer of the primary device, pack the audio data frame into real-time streaming packets RTPs, and publish the RTPs to a multicast service address and port, where a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port.

**[0088]** FIG. 13 is a schematic structural diagram of Embodiment 4 of a primary device according to an embodiment of the present invention. As shown in FIG. 13, based on the primary device shown in FIG. 12, further, the primary device in this embodiment may further include a receiver 43. The receiver 43 is configured to: before the transmitter 42 sends the RTCP packet to the secondary device when it is determined that the preset condition is met, receive an RTCP session join request sent by the secondary device, and send an RTCP session identifier to the secondary device.

**[0089]** Further, the transmitter 42 is further configured to: send media description information to the secondary device before the processor 41 collects the program clock reference PCR based on the preset collection cycle, where the media description information includes an SNTP service address and port, an RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

**[0090]** The preset condition is: a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

**[0091]** Optionally, the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

where scr\_curr is a currently collected clock value, scr\_last is a previously collected clock value, scf\_srv is the program clock frequency of the primary device, and cycle\_read\_x is the preset collection cycle.

[0092] The primary device in the embodiments shown in FIG. 12 and FIG. 13 may be configured to execute the technical solution of the method embodiment shown in FIG. 1 or FIG. 2, and an implementation principle and a technical effect thereof are similar and are not described herein again.

[0093] The embodiments in this specification are all described in a progressive manner, for same or similar parts in the embodiments, refer to these embodiments. Each embodiment focuses on a difference from other embodiments. Especially, a system or apparatus embodiment is basically similar to a method embodiment, and therefore is described briefly. For related parts, refer to partial descriptions in the method embodiment.

[0094] A person of ordinary skill in the art may understand that each aspect of the embodiments of the present invention or a possible implementation of each aspect may be specifically implemented as a system, a method, or a computer program product. Therefore, each aspect of the embodiments of the present invention or a possible implementation of each aspect may use forms of hardware only embodiments, software only embodiments (including firmware, resident software, and the like), or embodiments with a combination of software and hardware, which are uniformly referred to as "circuit", "module", or "system" herein. In addition, each aspect of the embodiments of the present invention or the possible implementation of each aspect may take a form of a computer program product, where the computer program product is computer-readable program code stored in a computer-readable medium.

[0095] The computer-readable medium may be a computer-readable signal medium or a computer-readable storage medium. The computer-readable storage medium includes but is not limited to an electronic, magnetic, optical, electromagnetic, infrared, or semi-conductive system, device, or apparatus, or any appropriate combination thereof, such as a random access memory (RAM), a read-only memory (ROM), an erasable programmable read only memory (EPROM or flash memory), an optical fiber, and a compact disc read-only memory (CD-ROM).

[0096] A processor in a computer reads the computer-readable program code stored in the computer-readable medium, so that the processor can perform a function and an action specified in each step or a combination of steps in a flowchart; an apparatus is generated to implement a function and an action specified in each block or a combination of blocks in a block diagram. All computer-readable program code may be executed on a local computer of a user, or some may be executed on a local computer of a user as a standalone software package, or some may be executed on a local computer of a user while some is executed on a remote computer, or all the code may be executed on a remote computer or a server. It should also be noted that in some alternative implementation solutions, each step in the flowcharts or functions specified in each block in the block diagrams may not occur in the illustrated order. For example, two steps or blocks that depend on a function and are shown in sequence may be actually executed concurrently, or sometimes these blocks may be executed in reverse order.

## Claims

1. A multi-device lip synchronization method, wherein the method is applied to a primary device and at least one secondary device synchronously output video and audio, wherein the primary device outputs video and the at least one secondary device outputs audio, wherein the primary device is connected to the at least one secondary device through Wi-Fi, and the method comprises:

receiving (S102), by the secondary device, a Real-Time Transport Control Protocol, RTCP, packet sent by the primary device, wherein a Real-Time Transport Protocol, RTP, time stamp header field in the RTCP packet carries a program clock reference, PCR, cyclically collected by the primary device, and a Network Time Protocol, NTP, field in the RTCP packet carries a sending time point of the RTCP packet;

correcting (S103), by the secondary device, a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay;

receiving (S105), by the secondary device based on a multicast service address and port sent by the primary device, RTPs published by the primary device, splicing the RTPs into a complete audio data frame, obtaining, from a time stamp header field of the RTP packet, a presentation time stamp, PTS, corresponding to the audio data frame, and storing the audio data frame into an audio pulse code modulation, PCM, buffer of the secondary device; and

outputting (S106), by the secondary device, the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame, wherein before the receiving (S102), by the secondary device, a Real-Time Transport Control Protocol, RTCP, packet sent by the primary device, the method further comprises:

receiving (S107), by the secondary device, media description information sent by the primary device, wherein the media description information comprises a Simple Network Time Protocol, SNTP, service address and port, a RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency;

performing (S108), by the secondary device, system time synchronization based on the SNTP service address and port; and

the receiving (S102), by the secondary device, a Real-Time Transport Control Protocol, RTCP, packet sent by the primary device comprises:

receiving, by the secondary device based on the RTCP service address and port, the RTCP packet sent by the primary device.

2. The method according to claim 1, wherein the RTCP packet carries a RTCP session identifier, and before the correcting (S103), by the secondary device, a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay, the method further comprises: sending, by the secondary device, a RTCP session join request to the primary device, so that the primary device sends the RTCP session identifier to the secondary device.

3. The method according to claim 1 or 2, wherein the correcting (S103), by the secondary device, a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay comprises:

calculating, by the secondary device, a STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

wherein

$scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and

correcting, by the secondary device, the STC of the secondary device based on the calculated STC correction value.

4. A multi-device lip synchronization method, wherein the method is applied to a primary device and at least one secondary device synchronously output video and audio, wherein the primary device outputs video and the at least one secondary device outputs audio, wherein the primary device is connected to the at least one secondary device through Wi-Fi, and the method comprises:

collecting (S101), by the primary device, a program clock reference, PCR, based on a preset collection cycle; sending (S102), by the primary device, a Real-Time Transport Control Protocol, RTCP, packet to the secondary device when determining that a preset condition is met, wherein a Real-Time Transport Protocol, RTP, time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol, NTP, field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device is able to correct a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay; and

copying (S104), by the primary device, an audio data frame from an audio pulse code modulation, PCM, buffer of the primary device, packing the audio data frame into Real-Time Transport Protocols, RTPs, and publishing the RTPs to a multicast service address and port, wherein a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port, wherein before the collecting (S101), by the primary device, a program clock reference, PCR, based on a preset collection cycle, the method further comprises:

sending (S107), by the primary device, media description information to the secondary device, wherein the media description information comprises a Simple Network Time Protocol, SNTP, service address and port, a RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by

the primary device.

5. The method according to claim 4, wherein the RTCP packet carries a RTCP session identifier, and before the sending (S102), by the primary device, a Real-Time Transport Control Protocol, RTCP, packet to the secondary device when determining that a preset condition is met, the method further comprises:  
 5 receiving, by the primary device, a RTCP session join request sent by the secondary device, and sending the RTCP session identifier to the secondary device.

6. The method according to any one of claims 4 or 5, wherein the preset condition is:  
 10 a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

7. The method according to claim 6, wherein the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

or

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

wherein

scr\_curr is a currently collected clock value, scr\_last is a previously collected clock value, scf\_srv is the program clock frequency of the primary device, and cycle\_read\_x is the preset collection cycle.

8. A secondary device, comprising:

a first receiving module (11), configured to receive a Real-Time Transport Control Protocol, RTCP, packet sent by primary device, wherein a Real-Time Transport Protocol, RTP, time stamp header field in the RTCP packet carries a program clock reference, PCR, cyclically collected by the primary device, and a Network Time Protocol, NTP, field in the RTCP packet carries a sending time point of the RTCP packet;

a correction module (12), configured to correct a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay;

a second receiving module (13), configured to receive, based on a multicast service address and port sent by the primary device, RTPs published by the primary device;

a processing module (14), configured to: splice the RTPs into a complete audio data frame, obtain, from a time stamp header field of the RTP packet, a presentation time stamp, PTS, corresponding to the audio data frame, and put the audio data frame into an audio pulse code modulation, PCM, buffer of the secondary device; and  
 40 an output module (15), configured to output the audio data frame in the audio PCM buffer based on the STC of the secondary device and the presentation time stamp of the audio data frame, wherein the first receiving module (11) is further configured to:

before receiving the Real-Time Transport Control Protocol, RTCP, packet sent by the primary device, receive media description information sent by the primary device, wherein the media description information comprises a Simple Network Time Protocol, SNTP, service address and port, a RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency;

the processing module (14) is further configured to perform system time synchronization based on the SNTP service address and port; and

the first receiving module (11) is specifically configured to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

9. The secondary device according to claim 8, wherein the RTCP packet carries a RTCP session identifier, and the secondary device further comprises:

a sending module (16), configured to send a RTCP session join request to the primary

device before the correction module (12) corrects the system clock, STC, of the secondary device based on the PCR, the program clock frequency of the primary device, the program clock frequency of the secondary device, and the RTCP delay, so that the primary device sends the RTCP session identifier to the secondary device.

10. The secondary device according to claim 8 or 9, wherein the correction module (12) is specifically configured to:

calculate a STC correction value  $scr\_correct$  of the secondary device according to the following formula:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

wherein  $scf\_clt$  is the program clock frequency of the secondary device,  $scf\_srv$  is the program clock frequency of the primary device,  $ntp\_rcv$  is a moment at which the RTCP packet is received, and  $scr\_srv$  and  $ntp\_snd$  are sending time points of the PCR and the RTCP packet, respectively; and correct the STC of the secondary device based on the calculated STC correction value.

11. A system for multi-device lip synchronization, composed by a primary device and at least one secondary device, the primary device comprising:

a collection module (21), configured to collect a program clock reference, PCR, based on a preset collection cycle; a sending module (22), configured to send a Real-Time Transport Control Protocol, RTCP, packet to the secondary device when a preset condition is met, wherein a Real-Time Transport Protocol, RTP, time stamp header field in the RTCP packet carries the PCR, and a Network Time Protocol, NTP, field in the RTCP packet carries a sending time point of the RTCP packet, so that the secondary device corrects a system clock, STC, of the secondary device based on the PCR, a program clock frequency of the primary device, a program clock frequency of the secondary device, and a RTCP delay; and

a processing module (23), configured to: copy an audio data frame from an audio pulse code modulation, PCM, buffer of the primary device, pack the audio data frame into Real-Time Transport Protocols, RTPs, and publish the RTPs to a multicast service address and port, wherein a time stamp header field of the RTP carries a presentation time stamp corresponding to the audio data frame, so that the secondary device receives the RTPs based on the multicast service address and port, wherein the sending module (22) is further configured to: send media description information to the secondary device before the collection module (21) collects the program clock reference PCR based on the preset collection cycle, wherein the media description information comprises a Simple Network Time Protocol SNTP service address and port, a RTCP service address and port, the multicast service address and port, the program clock frequency of the primary device, a media format, and a time stamp frequency, the SNTP service address and port is used by the secondary device to perform system time synchronization, and the RTCP service address and port is used by the secondary device to receive, based on the RTCP service address and port, the RTCP packet sent by the primary device.

12. The system according to claim 11, wherein the RTCP packet carries a RTCP session identifier, and the primary device further comprises:

a receiving module (24), configured to: before the sending module (22) sends the Real-Time Transport Control Protocol, RTCP, packet to the secondary device when the preset condition is met, receive a RTCP session join request sent by the secondary device, and send the RTCP session identifier to the secondary device.

13. The system according to any one of claims 11 or 12, wherein the preset condition is:

a deviation between an actual PCR collection time interval of the primary device and the preset collection cycle is greater than a preset threshold.

14. The system according to claim 13, wherein the preset threshold is 20 ms, and the preset condition is:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

or



$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} > (\text{cycle\_read\_x} + 20),$$

wherein scr\_curr is a currently collected clock value, scr\_last is a previously collected clock value, scf\_srv is the program clock frequency of the primary device, and cycle\_read\_x is the preset collection cycle.

## Patentansprüche

1. Lippensynchronisationsverfahren mit mehreren Vorrichtungen, wobei das Verfahren auf eine primäre Vorrichtung angewendet wird und wenigstens eine sekundäre Vorrichtung Video und Audio synchron ausgibt, wobei die primäre Vorrichtung Video ausgibt und die wenigstens eine sekundäre Vorrichtung Audio ausgibt, wobei die primäre Vorrichtung über WLAN mit der wenigstens einen sekundären Vorrichtung verbunden ist und das Verfahren Folgendes umfasst:

Empfangen (S102), durch die sekundäre Vorrichtung, eines Echtzeit-Transportsteuerprotokoll(Real-Time Transport Control Protocol - RTCP)-Pakets, das durch die primäre Vorrichtung gesendet wird, wobei ein Echtzeit-Transportprotokoll(Real-Time Transport Protocol - RTP)-Zeitstempel-Header-Feld in dem RTCP-Paket eine Programmtaktreferenz (program clock reference - PCR) trägt, die durch die primäre Vorrichtung zyklisch gesammelt wird, und ein Netzwerkzeitprotokoll(Network Time Protocol - NTP)-Feld in dem RTCP-Paket einen Sendezeitpunkt des RTCP-Pakets trägt;

Korrigieren (S103), durch die sekundäre Vorrichtung, eines Systemtakts (system clock - STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung;

Empfangen (S105), durch die sekundäre Vorrichtung basierend auf einer Multicast-Dienstadresse und einem durch die primäre Vorrichtung gesendeten Port, von durch die primäre Vorrichtung veröffentlichten RTPs, Spleißen der RTPs in einen vollständigen Audiodatenrahmen, Erhalten, von einem Zeitstempel-Header-Feld des RTP-Pakets, eines Präsentationszeitstempels (presentation time stamp - PTS), der dem Audiodatenrahmen entspricht, und Speichern des Audiodatenrahmens in einem Audio-Pulscodemodulations(PCM)-Puffer der sekundären Vorrichtung; und

Ausgeben (S106), durch die sekundäre Vorrichtung, des Audiodatenrahmens in dem Audio-PCM-Puffer basierend auf dem STC der sekundären Vorrichtung und dem Präsentationszeitstempel des Audiodatenrahmens, wobei vor dem Empfangen (S102), durch die sekundäre Vorrichtung, eines durch die primäre Vorrichtung gesendeten Echtzeit-Transportsteuerprotokoll(RTCP)-Pakets, das Verfahren ferner Folgendes umfasst:

Empfangen (S107), durch die sekundäre Vorrichtung, von durch die primäre Vorrichtung gesendeten Medienbeschreibungsinformationen, wobei die Medienbeschreibungsinformationen eine Simple-Network-Zeitprotokoll(Simple Network Time Protocol - SNTP)-Dienstadresse und einen Port, eine RTCP-Dienstadresse und einen Port, die Multicast-Dienstadresse und den Port, die Programmtaktfrequenz der primären Vorrichtung, ein Medienformat und eine Zeitstempelfrequenz umfassen;

Durchführen (S108), durch die sekundäre Vorrichtung, einer Systemzeitsynchronisation basierend auf der SNTP-Dienstadresse und dem Port; und

wobei das Empfangen (S102), durch die sekundäre Vorrichtung, eines durch die primäre Vorrichtung gesendeten Echtzeit-Transportsteuerprotokoll(RTCP)-Pakets, Folgendes umfasst:

Empfangen, durch die sekundäre Vorrichtung basierend auf der RTCP-Dienstadresse und dem Port, des durch die primäre Vorrichtung gesendeten RTCP-Pakets.

2. Verfahren nach Anspruch 1, wobei das RTCP-Paket einen RTCP-Sitzungsbezeichner trägt und vor dem Korrigieren (S103), durch die sekundäre Vorrichtung, eines Systemtakts (STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung das Verfahren ferner Folgendes umfasst:  
Senden, durch die sekundäre Vorrichtung, einer RTCP-Sitzungsbeitrittsanforderung an die primäre Vorrichtung, so dass die primäre Vorrichtung den RTCP-Sitzungsbezeichner an die sekundäre Vorrichtung sendet.
3. Verfahren nach Anspruch 1 oder 2, wobei das Korrigieren (S103), durch die sekundäre Vorrichtung, eines Systemtakts (STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung Folgendes umfasst:

fasst:

Berechnen, durch die sekundäre Vorrichtung, eines STC-Korrekturwerts  $scr\_correct$  der sekundären Vorrichtung gemäß der folgenden Formel:

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

wobei

$scf\_clt$  die Programmtaktfrequenz der sekundären Vorrichtung ist,  $scf\_srv$  die Programmtaktfrequenz der primären Vorrichtung ist,  $ntp\_rcv$  ein Moment ist, an dem das RTCP-Paket empfangen wird, und  $scr\_srv$  und  $ntp\_snd$  Sendezeitpunkte der PCR beziehungsweise des RTCP-Pakets sind; und

Korrigieren, durch die sekundäre Vorrichtung, des STC der sekundären Vorrichtung basierend auf dem berechneten STC-Korrekturwert.

4. Lippensynchronisationsverfahren mit mehreren Vorrichtungen, wobei das Verfahren auf eine primäre Vorrichtung angewendet wird und wenigstens eine sekundäre Vorrichtung Video und Audio synchron ausgibt, wobei die primäre Vorrichtung Video ausgibt und die wenigstens eine sekundäre Vorrichtung Audio ausgibt, wobei die primäre Vorrichtung über WLAN mit der wenigstens einen sekundären Vorrichtung verbunden ist und das Verfahren Folgendes umfasst:

Sammeln (S101), durch die primäre Vorrichtung, einer Programmtaktreferenz (PCR) basierend auf einem voreingestellten Sammelzyklus;

Senden (S102), durch die primäre Vorrichtung, eines Echtzeit-Transportsteuerprotokoll(RTCP)-Pakets an die sekundäre Vorrichtung, wenn bestimmt wird, dass eine voreingestellte Bedingung erfüllt ist, wobei ein Echtzeit-Transportprotokoll(RTP)-Zeitstempel-Header-Feld in dem RTCP-Paket die PCR trägt und ein Netzwerkzeitprotokoll(NTP)-Feld in dem RTCP-Paket einen Sendezeitpunkt des RTCP-Pakets trägt, so dass die sekundäre Vorrichtung einen Systemtakt (STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung korrigieren kann; und

Kopieren (S104), durch die primäre Vorrichtung, eines Audiodatenrahmens aus einem Audio-Pulsmodulations(PCM)-Puffer der primären Vorrichtung, Packen des Audiodatenrahmens in Echtzeit-Transportprotokolle (Real-Time Transport Protocols - RTPs) und Veröffentlichen der RTPs an eine Multicast-Dienstadresse und den Port, wobei ein Zeitstempel-Header-Feld des RTP einen Präsentationszeitstempel trägt, der dem Audiodatenrahmen entspricht, so dass die sekundäre Vorrichtung die RTPs basierend auf der Multicast-Dienstadresse und dem Port empfängt, wobei vor dem Sammeln (S101), durch die primäre Vorrichtung, einer Programmtaktreferenz (PCR) basierend auf einem voreingestellten Sammelzyklus das Verfahren ferner Folgendes umfasst: Senden (S107), durch die primäre Vorrichtung, von Medienbeschreibungsinformationen an die sekundäre Vorrichtung, wobei die Medienbeschreibungsinformationen eine Simple-Network-Zeitprotokoll(SNTP)-Dienstadresse und den Port, eine RTCP-Dienstadresse und den Port, die Multicast-Dienstadresse und den Port, die Programmtaktfrequenz der primären Vorrichtung, ein Medienformat und eine Zeitstempelfrequenz umfassen, die SNTP-Dienstadresse und der Port durch die sekundäre Vorrichtung verwendet werden, um eine Systemzeitsynchronisation durchzuführen, und die RTCP-Dienstadresse und der Port durch die sekundäre Vorrichtung verwendet werden, um basierend auf der RTCP-Dienstadresse und dem Port das durch die primäre Vorrichtung gesendete RTCP-Paket zu empfangen.

5. Verfahren nach Anspruch 4, wobei das RTCP-Paket einen RTCP-Sitzungsbezeichner trägt und vor dem Senden (S102), durch die primäre Vorrichtung, eines Echtzeit-Transportsteuerprotokoll(RTCP)-Pakets an die sekundäre Vorrichtung, wenn bestimmt wird, dass eine voreingestellte Bedingung erfüllt ist, das Verfahren ferner Folgendes umfasst:

Empfangen, durch die primäre Vorrichtung, einer durch die sekundäre Vorrichtung gesendeten RTCP-Sitzungsbeitrittsanforderung und Senden des RTCP-Sitzungsbezeichners an die sekundäre Vorrichtung.

6. Verfahren nach einem der Ansprüche 4 oder 5, wobei die voreingestellte Bedingung Folgendes ist: eine Abweichung zwischen einem tatsächlichen PCR-Sammelzeitintervall der primären Vorrichtung und dem voreingestellten Sammelzyklus ist größer als eine voreingestellte Schwelle.

7. Verfahren nach Anspruch 6, wobei die voreingestellte Schwelle 20 ms beträgt und die voreingestellte Bedingung

Folgendes ist:

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20);$$

oder

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

wobei  $scr\_curr$  ein aktuell gesamelter Taktwert ist,  $scr\_last$  ein zuvor gesamelter Taktwert ist,  $scf\_srv$  die Programmtaktfrequenz der primären Vorrichtung ist und  $cycle\_read\_x$  der voreingestellte Sammelzyklus ist.

**8. Zweite Vorrichtung, die Folgendes umfasst:**

ein erstes Empfangsmodul (11), das konfiguriert ist, um ein durch die primäre Vorrichtung gesendetes Echtzeit-Transportsteuerprotokoll(RTCP)-Paket zu empfangen, wobei ein Echtzeit-Transportprotokoll(RTP)-Zeitstempel-Header-Feld in dem RTCP-Paket eine Programmtaktreferenz (PCR) trägt, die durch die primäre Vorrichtung zyklisch gesammelt wird, und ein Netzwerkzeitprotokoll(NTP)-Feld in dem RTCP-Paket einen Sendezeitpunkt des RTCP-Pakets trägt;

ein Korrekturmodul (12), das konfiguriert ist, um einen Systemtakt (STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung zu korrigieren;

ein zweites Empfangsmodul (13), das konfiguriert ist, um basierend auf einer Multicast-Dienstadresse und dem durch die primäre Vorrichtung gesendeten Port durch die primäre Vorrichtung veröffentlichte RTPs zu empfangen;

ein Verarbeitungsmodul (14), das für Folgendes konfiguriert ist: Spleißen der RTPs in einen vollständigen Audiodatenrahmen, Erhalten, von einem Zeitstempel-Header-Feld des RTP-Pakets, eines Präsentationszeitstempels (PTS), der dem Audiodatenrahmen entspricht, und Setzen des Audiodatenrahmens in einen Audio-Pulsmodulations(PCM)-Puffer der sekundären Vorrichtung; und

ein Ausgabemodul (15), das konfiguriert ist, um den Audiodatenrahmen in dem Audio-PCM-Puffer basierend auf dem STC der sekundären Vorrichtung und dem Präsentationszeitstempel des Audiodatenrahmens auszugeben, wobei das erste Empfangsmodul (11) ferner für Folgendes konfiguriert ist:

vor dem Empfangen des durch die primäre Vorrichtung gesendeten Echtzeit-Transportsteuerprotokoll(RTCP)-Pakets, Empfangen von Medienbeschreibungsinformationen, die durch die primäre Vorrichtung gesendet werden, wobei die Medienbeschreibungsinformationen eine Simple-Network-Zeitprotokoll(SNTP)-Dienstadresse und den Port, eine RTCP-Dienstadresse und den Port, die Multicast-Dienstadresse und den Port, die Programmtaktfrequenz der primären Vorrichtung, ein Medienformat und eine Zeitstempelfrequenz umfassen;

das Verarbeitungsmodul (14) ferner konfiguriert ist, um eine Systemzeitsynchronisation basierend auf der SNTP-Dienstadresse und dem Port durchzuführen; und

das erste Empfangsmodul (11) speziell konfiguriert ist, um basierend auf der RTCP-Dienstadresse und dem Port das durch die primäre Vorrichtung gesendete RTCP-Paket zu empfangen.

**9. Sekundäre Vorrichtung nach Anspruch 8, wobei das RTCP-Paket einen RTCP-Sitzungsbezeichner trägt und die sekundäre Vorrichtung ferner Folgendes umfasst:**

ein Sendemodul (16), das konfiguriert ist, um eine RTCP-Sitzungsbeitrittsanforderung an die primäre Vorrichtung zu senden, bevor das Korrekturmodul (12) den Systemtakt (STC) der sekundären Vorrichtung basierend auf der PCR, der Programmtaktfrequenz der primären Vorrichtung, der Programmtaktfrequenz der sekundären Vorrichtung und der RTCP-Verzögerung korrigiert, so dass die primäre Vorrichtung den RTCP-Sitzungsbezeichner an die sekundäre Vorrichtung sendet.

**10. Die sekundäre Vorrichtung nach Anspruch 8 oder 9, wobei das Korrekturmodul (12) speziell für Folgendes konfiguriert ist:**

Berechnen eines STC-Korrekturwerts  $scr\_correct$  der sekundären Vorrichtung gemäß der folgenden Formel:

$$\text{scr\_correct} = \text{scr\_srv} \times \text{scf\_clt} / \text{scf\_srv} + (\text{ntp\_rcv} - \text{ntp\_snd}) \times \text{scf\_clt} / 1000,$$

wobei

scf\_clt die Programmtaktfrequenz der sekundären Vorrichtung ist, scf\_srv die Programmtaktfrequenz der primären Vorrichtung ist, ntp\_rcv ein Moment ist, an dem das RTCP-Paket empfangen wird, und scr\_srv und ntp\_snd Sendezeitpunkte der PCR beziehungsweise des RTCP-Pakets sind; und

Korrigieren des STC der sekundären Vorrichtung basierend auf dem berechneten STC-Korrekturwert.

11. System für die Lippensynchronisation mit mehreren Vorrichtungen, das aus einer primären Vorrichtung und wenigstens einer sekundären Vorrichtung besteht, wobei die primäre Vorrichtung Folgendes umfasst:

ein Sammelmodul (21), das konfiguriert ist, um eine Programmtaktreferenz (PCR) basierend auf einem voreingestellten Sammelzyklus zu sammeln;

ein Sendemodul (22), das konfiguriert ist, um ein Echtzeit-Transportsteuerprotokoll(RTCP)-Paket an die sekundäre Vorrichtung zu senden, wenn eine voreingestellte Bedingung erfüllt ist, wobei ein Echtzeit-Transportprotokoll(RTP)-Zeitstempel-Header-Feld in dem RTCP-Paket die PCR trägt und ein Netzwerkzeitprotokoll(NTP)-Feld in dem RTCP-Paket einen Sendezeitpunkt des RTCP-Pakets trägt, so dass die sekundäre Vorrichtung einen Systemtakt (STC) der sekundären Vorrichtung basierend auf der PCR, einer Programmtaktfrequenz der primären Vorrichtung, einer Programmtaktfrequenz der sekundären Vorrichtung und einer RTCP-Verzögerung korrigiert; und

ein Verarbeitungsmodul (23), das für Folgendes konfiguriert ist: Kopieren eines Audiodatenrahmens aus einem Audio-Pulscodemodulation(PCM)-Puffer der primären Vorrichtung, Packen des Audiodatenrahmens in Echtzeit-Transportprotokolle (RTPs) und Veröffentlichen der RTPs an eine Multicast-Dienstadresse und den Port, wobei ein Zeitstempel-Header-Feld des RTP einen Präsentationszeitstempel trägt, der dem Audiodatenrahmen entspricht, so dass die sekundäre Vorrichtung die RTPs basierend auf der Multicast-Dienstadresse und dem Port empfängt, wobei das Sendemodul (22) ferner für Folgendes konfiguriert ist:

Senden von Medienbeschreibungsinformationen an die sekundäre Vorrichtung, bevor das Sammelmodul (21) die Programmtaktreferenz (PCR) basierend auf dem voreingestellten Sammelzyklus sammelt, wobei die Medienbeschreibungsinformationen eine Simple-Network-Zeitprotokoll(SNTP)-Dienstadresse und den Port, eine RTCP-Dienstadresse und den Port, die Multicast-Dienstadresse und den Port, die Programmtaktfrequenz der primären Vorrichtung, ein Medienformat und eine Zeitstempelfrequenz umfassen, die SNTP-Dienstadresse und der Port durch die sekundäre Vorrichtung verwendet werden, um die Systemzeitsynchronisation durchzuführen, und die RTCP-Dienstadresse und der Port durch die sekundäre Vorrichtung verwendet werden, um basierend auf der RTCP-Dienstadresse und dem Port das durch die primäre Vorrichtung gesendete RTCP-Paket zu empfangen.

12. System nach Anspruch 11, wobei das RTCP-Paket einen RTCP-Sitzungsbezeichner trägt und die primäre Vorrichtung ferner Folgendes umfasst:

ein Empfangsmodul (24), das für Folgendes konfiguriert ist: bevor das Sendemodul (22) das Echtzeit-Transportsteuerprotokoll(RTCP)-Paket an die sekundäre Vorrichtung sendet, wenn die voreingestellte Bedingung erfüllt ist, Empfangen einer durch die sekundäre Vorrichtung gesendeten RTCP-Sitzungsbeitrittsanforderung und Senden des die RTCP-Sitzungsbezeichners an die sekundäre Vorrichtung.

13. System, nach einem der Ansprüche 11 oder 12, wobei die voreingestellte Bedingung Folgendes ist: eine Abweichung zwischen einem tatsächlichen PCR-Sammelzeitintervall der primären Vorrichtung und dem voreingestellten Sammelzyklus ist größer als eine voreingestellte Schwelle.

14. System nach Anspruch 13, wobei die voreingestellte Schwelle 20 ms beträgt und die voreingestellte Bedingung Folgendes ist:

$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} < (\text{cycle\_read\_x} - 20);$$

oder

$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} > (\text{cycle\_read\_x} + 20),$$

wobei

scr\_curr ein aktuell gesammelter Taktwert ist, scr\_last ein zuvor gesammelter Taktwert ist, scf\_srv die Programmtaktfrequenz der primären Vorrichtung ist und cycle\_read\_x der voreingestellte Sammelzyklus ist.

5

## Revendications

1. Procédé de synchronisation labiale multi-dispositifs, le procédé étant appliqué à un dispositif primaire et au moins un dispositif secondaire émettant une vidéo et un audio de manière synchrone, le dispositif principal émettant une vidéo et l'au moins un dispositif secondaire émettant un audio, le dispositif primaire étant connecté à l'au moins un dispositif secondaire par Wi-Fi, et le procédé comprenant :

la réception (S102), par le dispositif secondaire, d'un paquet de protocoles de commande de transport en temps réel, RTCP, envoyé par le dispositif primaire, un champ d'en-tête d'estampille temporelle de protocole de transport en temps réel, RTP, dans le paquet RTCP transportant une référence d'horloge de programme, PCR, collectée cycliquement par le dispositif primaire, et un champ de protocole de synchronisation de réseau, NTP, dans le paquet RTCP transportant un point de synchronisation d'envoi du paquet RTCP ;

la correction (S103), par le dispositif secondaire, d'une horloge système, STC, du dispositif secondaire sur la base de la PCR, d'une fréquence d'horloge de programme du dispositif primaire, d'une fréquence d'horloge de programme du dispositif secondaire et d'un retard RTCP ;

la réception (S105), par le dispositif secondaire sur la base d'une adresse et d'un port de service de multidiffusion envoyés par le dispositif primaire, de RTP publiés par le dispositif primaire, le collage des RTP en une trame de données audio complète, l'obtention, à partir d'un champ d'en-tête d'estampille temporelle du paquet RTP, d'une estampille temporelle de présentation, PTS, correspondant à la trame de données audio, et le stockage de la trame de données audio dans un tampon de modulation par code d'impulsion, PCM, audio du dispositif secondaire ; et

l'émission (S106), par le dispositif secondaire, de la trame de données audio dans le tampon PCM audio sur la base de la STC du dispositif secondaire et de l'estampille temporelle de présentation de la trame de données audio, avant la réception (S102), par le dispositif secondaire, d'un paquet de protocoles de commande de transport en temps réel, RTCP, envoyé par le dispositif primaire, le procédé comprenant en outre :

la réception (S107), par le dispositif secondaire, d'informations de description multimédia envoyées par le dispositif primaire, les informations de description multimédia comprenant une adresse et un port de service de protocole de synchronisation de réseau simple, SNTP, une adresse et un port de service RTCP, l'adresse et le port de service de multidiffusion, la fréquence d'horloge de programme du dispositif principal, un format multimédia et une fréquence d'estampille temporelle ;

l'exécution (S108), par le dispositif secondaire, d'une synchronisation temporelle du système sur la base de l'adresse et du port de service SNTP ; et

la réception (S102), par le dispositif secondaire, d'un paquet de protocoles de commande de transport en temps réel, RTCP, envoyé par le dispositif primaire comprenant :

la réception, par le dispositif secondaire sur la base de l'adresse et du port de service RTCP, du paquet RTCP envoyé par le dispositif primaire.

2. Procédé selon la revendication 1, le paquet RTCP transportant un identifiant de session RTCP, et avant la correction (S103), par le dispositif secondaire, d'une horloge système, STC, du dispositif secondaire sur la base de la PCR, d'une fréquence d'horloge de programme du dispositif primaire, d'une fréquence d'horloge de programme du dispositif secondaire et d'un retard RTCP, le procédé comprenant en outre :

l'envoi, par le dispositif secondaire, d'une demande de participation à la session RTCP au dispositif principal, de sorte que le dispositif principal envoie l'identifiant de session RTCP au dispositif secondaire.

3. Procédé selon la revendication 1 ou 2, la correction (S103), par le dispositif secondaire, d'une horloge système, STC, du dispositif secondaire sur la base de la PCR, d'une fréquence d'horloge de programme du dispositif primaire, d'une fréquence d'horloge de programme du dispositif secondaire, et d'un retard RTCP comprenant :

le calcul, par le dispositif secondaire, d'une valeur de correction STC scr\_correct du dispositif secondaire selon la formule suivante :

$$\text{scr\_correct} = \text{scr\_srv} \times \text{scf\_clt} / \text{scf\_srv} + (\text{ntp\_rcv} - \text{ntp\_snd}) \times \text{scf\_clt} / 1000,$$

scf\_clt étant la fréquence d'horloge de programme du dispositif secondaire, scf\_srv étant la fréquence d'horloge de programme du dispositif primaire, ntp\_rcv étant un moment auquel le paquet RTCP est reçu et scr\_srv et ntp\_snd étant des points de synchronisation d'envoi de la PCR et du paquet RTCP, respectivement ; et la correction, par le dispositif secondaire, de la STC du dispositif secondaire sur la base de la valeur de correction STC calculée.

4. Procédé de synchronisation labiale multi-dispositifs, le procédé étant appliqué à un dispositif primaire et au moins un dispositif secondaire émettant une vidéo et un audio de manière synchrone, le dispositif principal émettant une vidéo et l'au moins un dispositif secondaire émettant un audio, le dispositif primaire étant connecté à l'au moins un dispositif secondaire par Wi-Fi, et le procédé comprenant :

la collecte (S101), par le dispositif primaire, d'une référence d'horloge de programme, PCR, sur la base d'un cycle de collecte prédéfini ;

l'envoi (S102), par le dispositif primaire, d'un paquet de protocoles de commande de transport en temps réel, RTCP, au dispositif secondaire lors de la détermination du fait qu'une condition prédéfinie est satisfaite, un champ d'en-tête d'estampille temporelle de protocole de transport en temps réel, RTP, dans le paquet RTCP transportant la PCR, et un champ de protocole de synchronisation de réseau, NTP, dans le paquet RTCP transportant un point de synchronisation d'envoi du paquet RTCP, de sorte que le dispositif secondaire peut corriger une horloge système, STC, du dispositif secondaire sur la base de la PCR, une fréquence d'horloge de programme du dispositif primaire, une fréquence d'horloge de programme du dispositif secondaire et un retard RTCP ; et

la copie (S104), par le dispositif primaire, d'une trame de données audio à partir d'un tampon de modulation par code d'impulsion, PCM, audio du dispositif primaire, l'empaquetage de la trame de données audio dans des protocoles de transport en temps réel, RTP, et la publication des RTP sur une adresse et un port de service de multidiffusion, un champ d'en-tête d'estampille temporelle du RTP transportant une estampille temporelle de présentation correspondant à la trame de données audio, de sorte que le dispositif secondaire reçoit les RTP sur la base de l'adresse et du port de service de multidiffusion, avant la collecte (S101), par le dispositif primaire, d'une référence d'horloge de programme, PCR, sur la base d'un cycle de collecte prédéfini, le procédé comprenant en outre :

l'envoi (S107), par le dispositif primaire, d'informations de description multimédia au dispositif secondaire, les informations de description multimédia comprenant une adresse et un port de service de protocole de synchronisation de réseau simple, SNTP, une adresse et un port de service RTCP, l'adresse et le port de service de multidiffusion, la fréquence d'horloge de programme du dispositif principal, un format multimédia et une fréquence d'estampille temporelle, l'adresse et le port de service SNTP étant utilisés par le dispositif secondaire pour exécuter la synchronisation temporelle du système, et l'adresse et le port de service RTCP étant utilisés par le dispositif secondaire pour recevoir, sur la base de l'adresse et du port de service RTCP, le paquet RTCP envoyé par le dispositif primaire.

5. Procédé selon la revendication 4, le paquet RTCP transportant un identifiant de session RTCP, et avant l'envoi (S102), par le dispositif primaire, d'un paquet de protocoles de commande de transport en temps réel, RTCP, au dispositif secondaire lors de la détermination du fait qu'une condition prédéfinie est satisfaite, le procédé comprenant en outre :

la réception, par le dispositif primaire, d'une demande de participation à la session RTCP envoyée par le dispositif secondaire et l'envoi de l'identifiant de session RTCP au dispositif secondaire.

6. Procédé selon l'une quelconque des revendications 4 ou 5, la condition prédéfinie étant :

un écart entre un intervalle de temps de collecte PCR réel du dispositif primaire et le cycle de collecte prédéfini supérieur à un seuil prédéfini.

7. Procédé selon la revendication 6, le seuil prédéfini étant de 20 ms et la condition prédéfinie étant :

$$(\text{scr\_curr} - \text{scr\_last}) \times 1000 / \text{scf\_srv} < (\text{cycle\_read\_x} - 20) ;$$

ou

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

scr\_curr étant une valeur d'horloge actuellement collectée, scr\_last étant une valeur d'horloge précédemment collectée, scf\_srv étant la fréquence d'horloge de programme du dispositif primaire et cycle\_read\_x étant le cycle de collecte prédéfini.

**8.** Dispositif secondaire, comprenant :

un premier module de réception (11), configuré pour recevoir un paquet de protocoles de commande de transport en temps réel, RTCP, envoyé par un dispositif primaire, un champ d'en-tête d'estampille temporelle de protocole de transport en temps réel, RTP, dans le paquet RTCP transportant une référence d'horloge de programme, PCR, collectée cycliquement par le dispositif primaire, et un champ de protocole de synchronisation de réseau, NTP, dans le paquet RTCP transportant un point de synchronisation d'envoi du paquet RTCP ;

un module de correction (12), configuré pour corriger une horloge système, STC, du dispositif secondaire sur la base de la PCR, une fréquence d'horloge de programme du dispositif primaire, une fréquence d'horloge de programme du dispositif secondaire et un retard RTCP ;

un second module de réception (13), configuré pour recevoir, sur la base d'une adresse et d'un port de service de multidiffusion envoyés par le dispositif primaire, des RTP publiés par le dispositif primaire ;

un module de traitement (14), configuré pour : coller les RTP en une trame de données audio complète, obtenir, à partir d'un champ d'en-tête d'estampille temporelle du paquet RTP, une estampille temporelle de présentation, PTS, correspondant à la trame de données audio, et mettre la trame de données audio dans un tampon de modulation par code d'impulsion, PCM, audio du dispositif secondaire ; et

un module d'émission (15), configuré pour émettre la trame de données audio dans le tampon PCM audio sur la base de la STC du dispositif secondaire et de l'estampille temporelle de présentation de la trame de données audio, le premier module de réception (11) étant en outre configuré pour :

avant la réception du paquet de protocoles de commande de transport en temps réel, RTCP, envoyé par le dispositif primaire, recevoir les informations de description multimédia envoyées par le dispositif primaire, les informations de description multimédia comprenant une adresse et un port de service de protocole de synchronisation de réseau simple, SNTP, une adresse et un port de service RTCP, l'adresse et le port de service de multidiffusion, la fréquence d'horloge de programme du dispositif primaire, un format multimédia et une fréquence d'estampille temporelle ;

le module de traitement (14) est en outre configuré pour exécuter une synchronisation temporelle du système sur la base de l'adresse et du port de service SNTP ; et

le premier module de réception (11) étant spécifiquement configuré pour recevoir, sur la base de l'adresse et du port de service RTCP, le paquet RTCP envoyé par le dispositif primaire.

**9.** Dispositif secondaire selon la revendication 8, le paquet RTCP transportant un identifiant de session RTCP, et le dispositif secondaire comprenant en outre :

un module d'envoi (16), configuré pour envoyer une demande de participation à la session RTCP au dispositif primaire avant que le module de correction (12) corrige l'horloge système, STC, du dispositif secondaire sur la base de la PCR, de la fréquence d'horloge de programme du dispositif primaire, de la fréquence d'horloge du programme du dispositif secondaire et du retard RTCP, de sorte que le dispositif primaire envoie l'identifiant de session RTCP au dispositif secondaire.

**10.** Dispositif secondaire selon la revendication 8 ou 9, le module de correction (12) étant spécifiquement configuré pour :

calculer une valeur de correction STC scr\_correct du dispositif secondaire selon la formule suivante :

$$scr\_correct = scr\_srv \times scf\_clt / scf\_srv + (ntp\_rcv - ntp\_snd) \times scf\_clt / 1000,$$

scf\_clt étant la fréquence d'horloge de programme du dispositif secondaire, scf\_srv étant la fréquence d'horloge de programme du dispositif primaire, ntp\_rcv étant un moment auquel le paquet RTCP est reçu et scr\_srv et ntp\_snd étant des points de synchronisation d'envoi de la PCR et du paquet RTCP, respectivement ; et corriger la STC du dispositif secondaire sur la base de la valeur de correction STC calculée.

11. Système de synchronisation labiale multi-dispositifs, composé d'un dispositif primaire et d'au moins un dispositif secondaire, le dispositif primaire comprenant :

un module de collecte (21), configuré pour collecter une référence d'horloge de programme, PCR, sur la base d'un cycle de collecte prédéfini ;

un module d'envoi (22), configuré pour envoyer un paquet de protocoles de commande de transport en temps réel, RTCP, au dispositif secondaire lorsqu'une condition prédéfinie est satisfaite, un champ d'en-tête d'estampille temporelle de protocole de transport en temps réel, RTP, dans le paquet RTCP transportant la PCR, et un champ de protocole de synchronisation de réseau, NTP, dans le paquet RTCP transportant un point de synchronisation d'envoi du paquet RTCP, de sorte que le dispositif secondaire corrige une horloge système, STC, du dispositif secondaire sur la base de la PCR, d'une fréquence d'horloge de programme du dispositif primaire, d'une fréquence d'horloge de programme du dispositif secondaire et d'un retard RTCP ; et

un module de traitement (23), configuré pour : copier une trame de données audio à partir d'un tampon de modulation par code d'impulsion, PCM, audio du dispositif primaire, empaqueter la trame de données audio dans des protocoles de transport en temps réel, RTP, et publier les RTP sur une adresse et un port de service de multidiffusion, un champ d'en-tête d'estampille temporelle du RTP transportant une estampille temporelle de présentation correspondant à la trame de données audio, de sorte que le dispositif secondaire reçoit les RTP sur la base de l'adresse et du port de service de multidiffusion, le module d'envoi (22) étant en outre configuré pour :

envoyer des informations de description multimédia au dispositif secondaire avant que le module de collecte (21) collecte la référence d'horloge de programme PCR sur la base du cycle de collecte prédéfini, les informations de description multimédia comprenant une adresse et un port de service de protocole de synchronisation de réseau simple, SNTP, une adresse et un port de service RTCP, l'adresse et le port de service de multidiffusion, la fréquence d'horloge de programme du dispositif principal, un format multimédia et une fréquence d'estampille temporelle, l'adresse et le port de service SNTP étant utilisés par le dispositif secondaire pour exécuter la synchronisation temporelle du système, et l'adresse et le port de service RTCP étant utilisés par le dispositif secondaire pour recevoir, sur la base de l'adresse et du port de service RTCP, le paquet RTCP envoyé par le dispositif primaire.

12. Système selon la revendication 11, le paquet RTCP transportant un identifiant de session RTCP, et le dispositif primaire comprenant en outre :

un module de réception (24), configuré pour : avant que le module d'envoi (22) envoie le paquet de protocoles de commande de transport en temps réel, RTCP, au dispositif secondaire lorsque la condition prédéfinie est satisfaite, recevoir une demande de participation à la session RTCP envoyée par le dispositif secondaire et envoyer l'identifiant de session RTCP au dispositif secondaire.

13. Système selon l'une quelconque des revendications 11 ou 12, la condition prédéfinie étant : un écart entre un intervalle de temps de collecte PCR réel du dispositif primaire et le cycle de collecte prédéfini supérieur à un seuil prédéfini.

14. Système selon la revendication 13, le seuil prédéfini étant de 20 ms et la condition prédéfinie étant :

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv < (cycle\_read\_x - 20) ;$$

ou

$$(scr\_curr - scr\_last) \times 1000 / scf\_srv > (cycle\_read\_x + 20),$$

scr\_curr étant une valeur d'horloge actuellement collectée, scr\_last étant une valeur d'horloge précédemment collectée, scf\_srv étant la fréquence d'horloge de programme du dispositif primaire et cycle\_read\_x étant le cycle de collecte prédéfini.



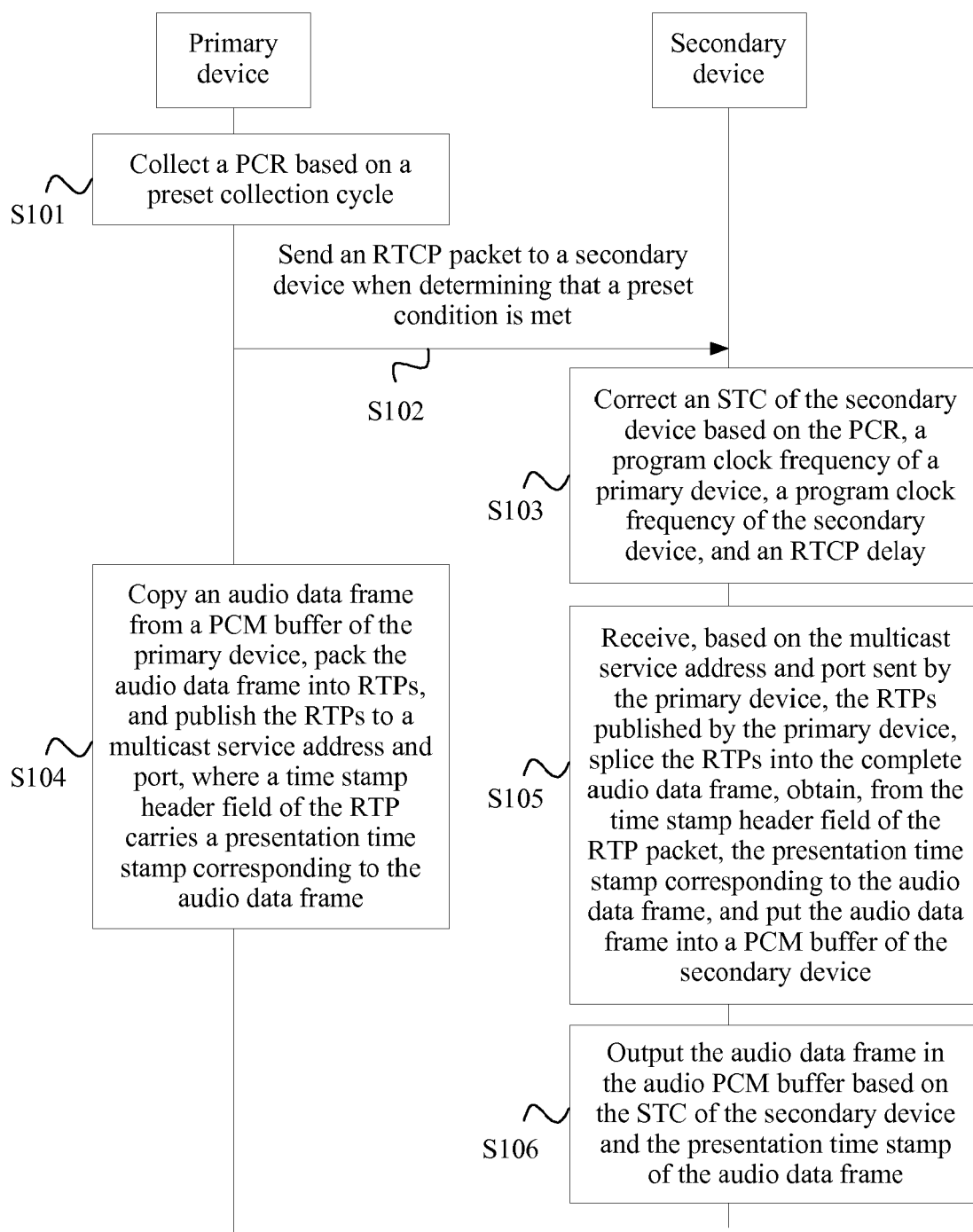


FIG. 1

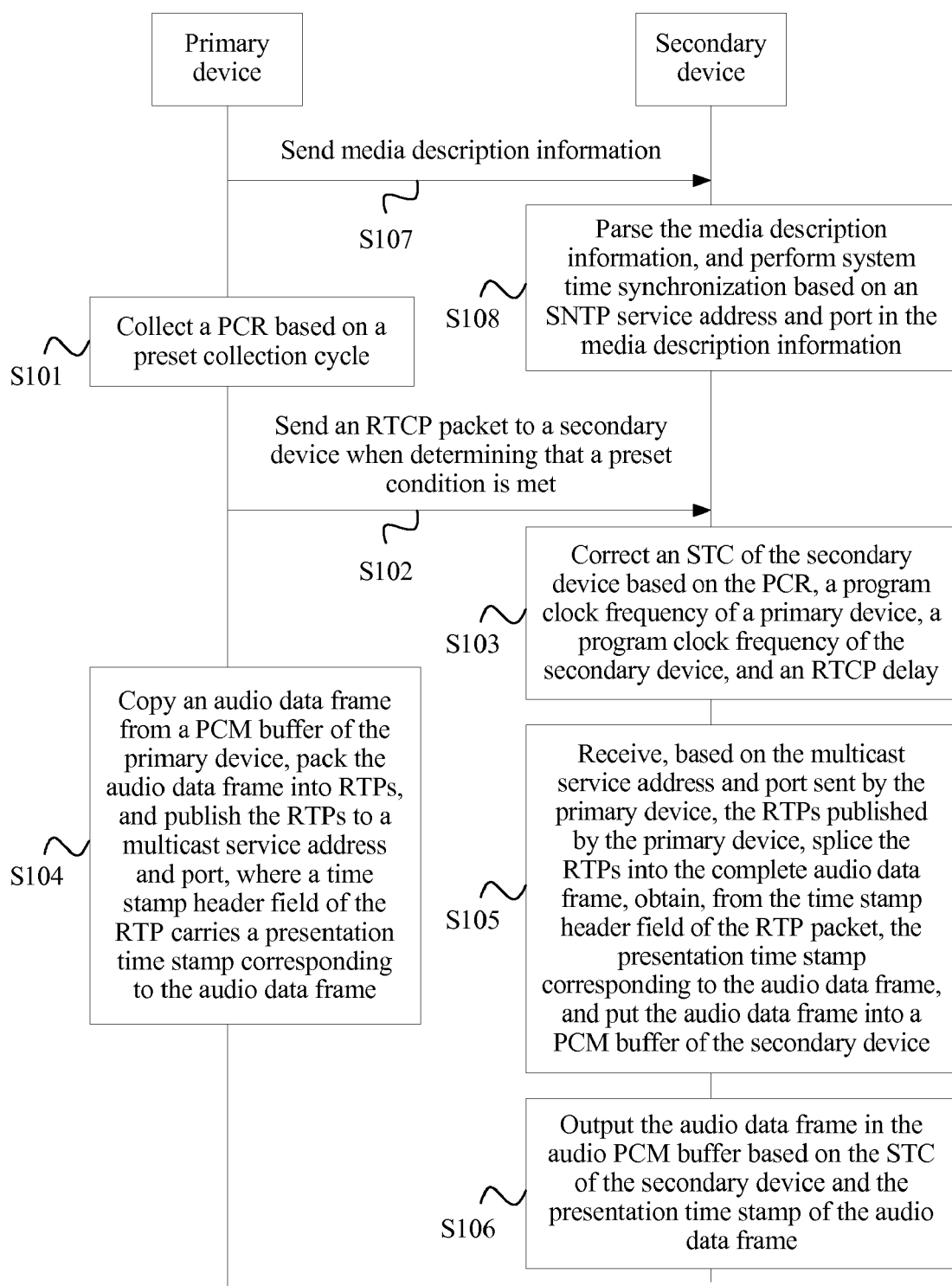


FIG. 2

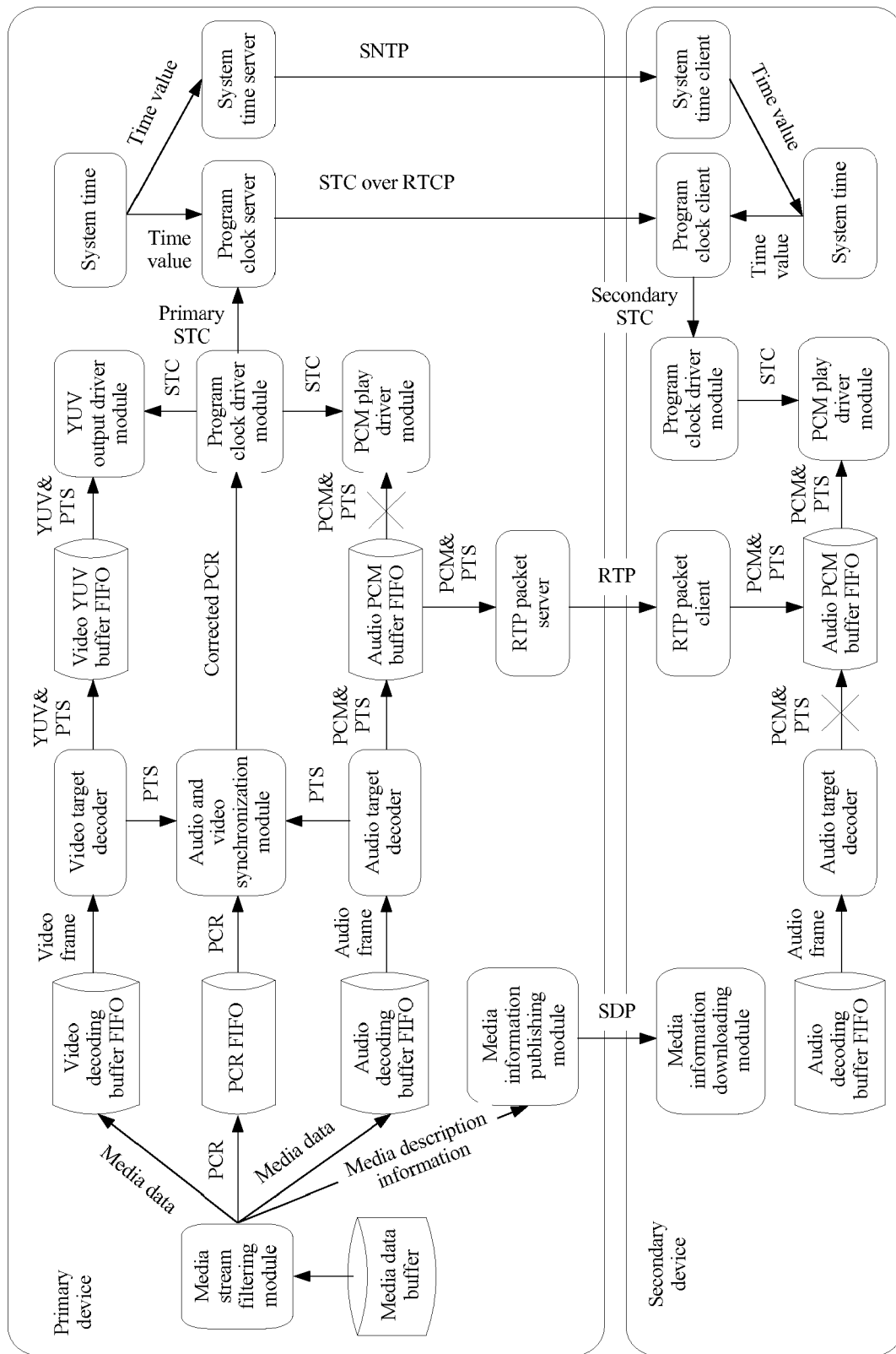


FIG. 3

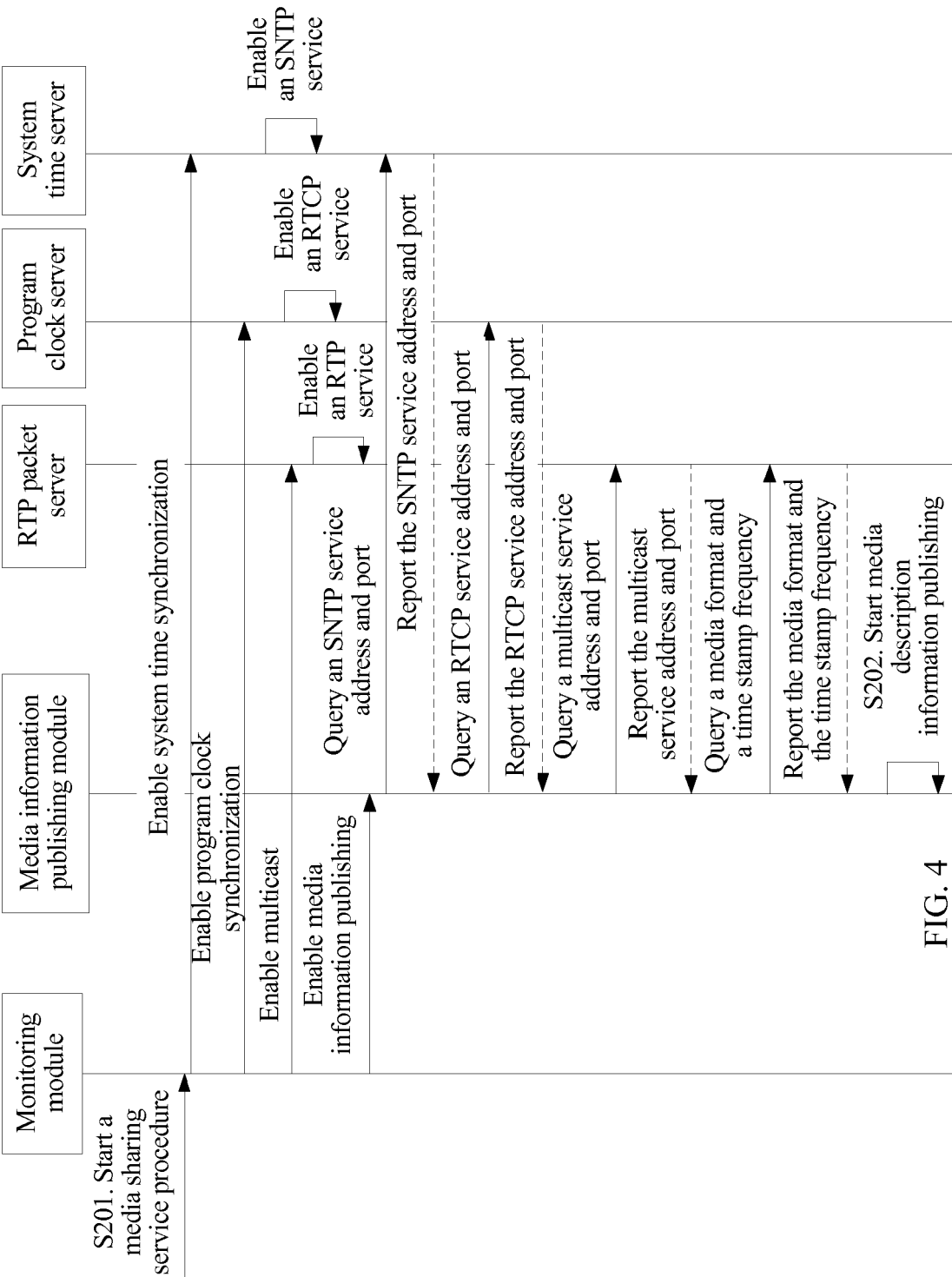


FIG. 4

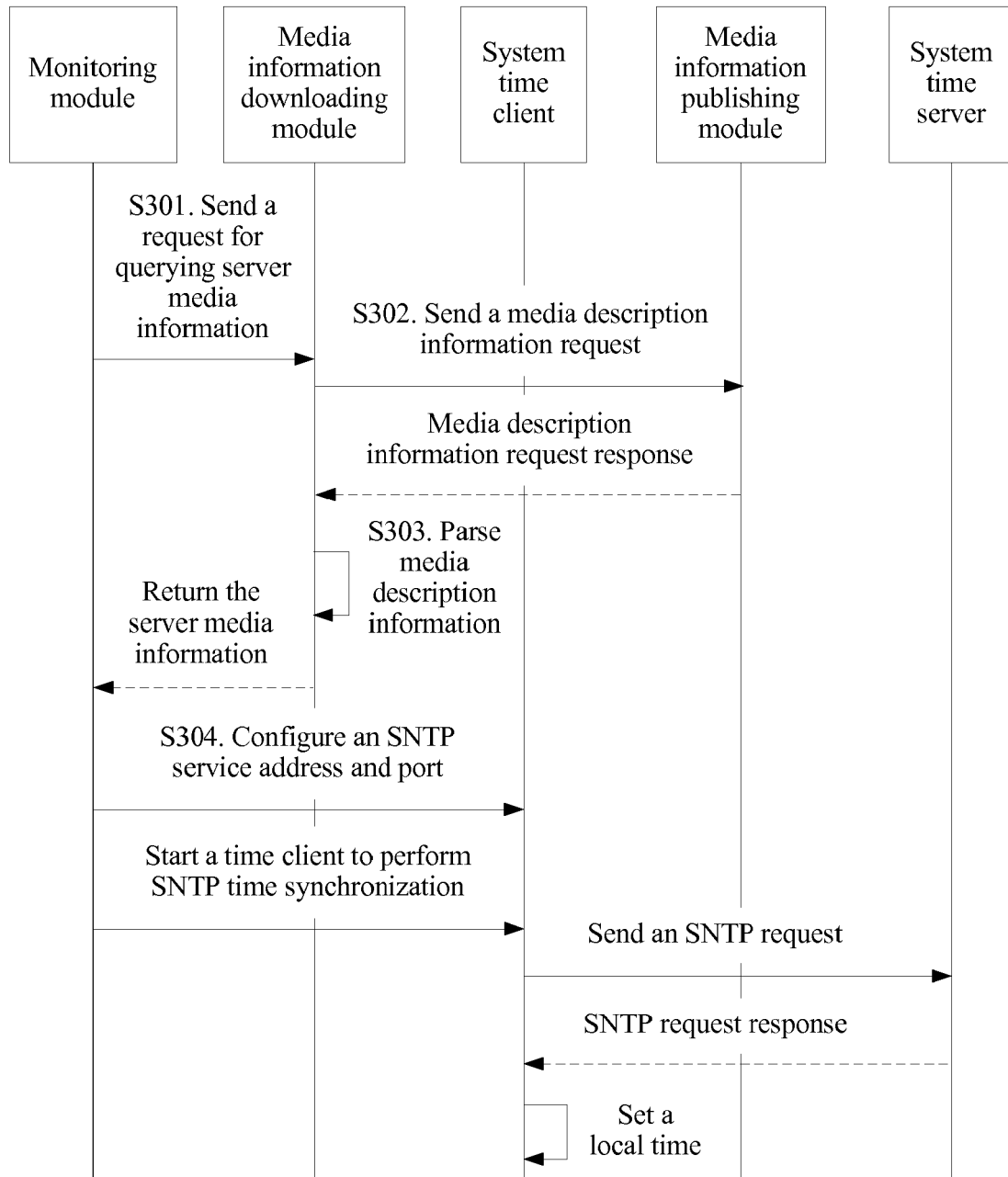


FIG. 5

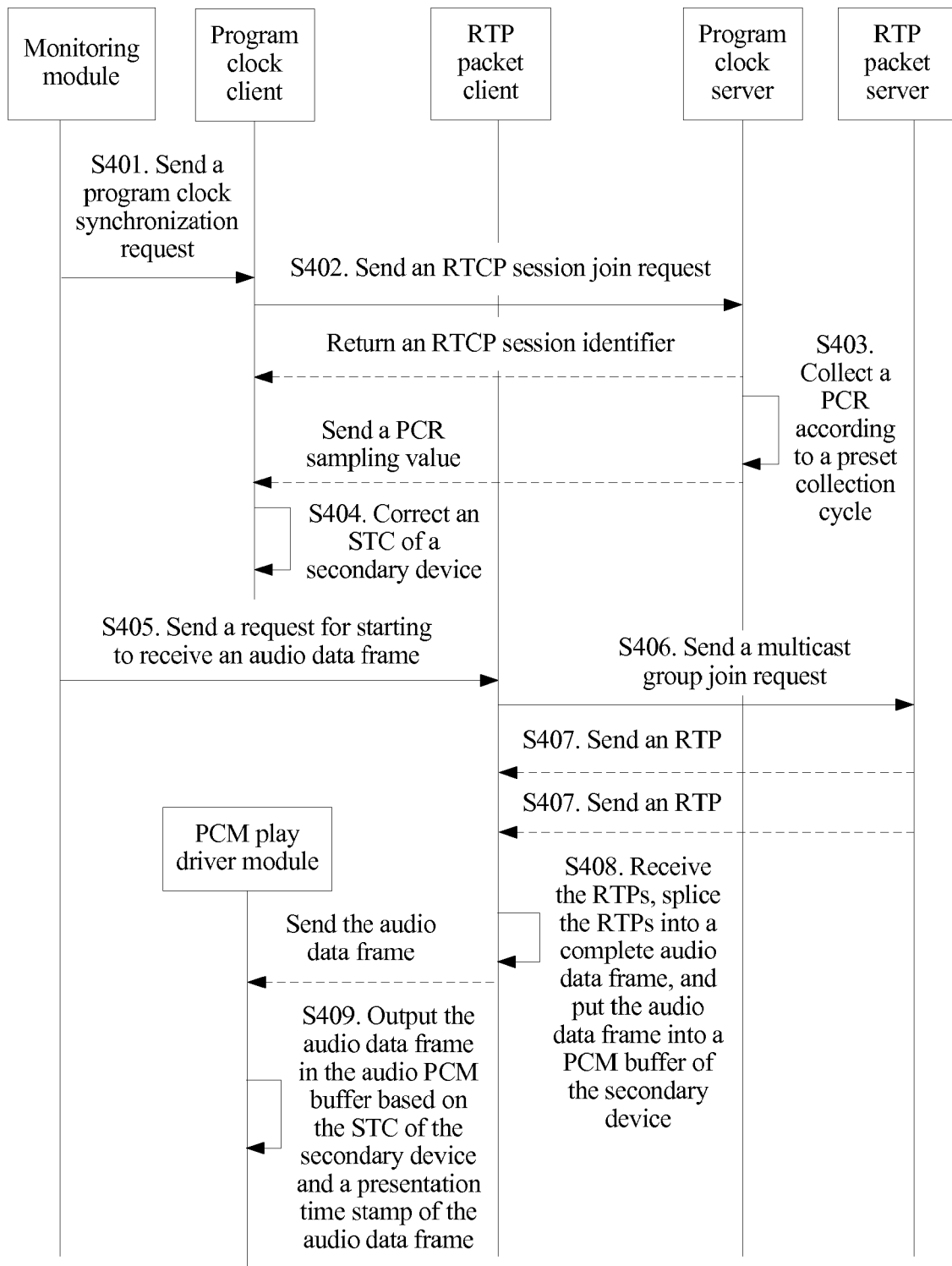


FIG. 6

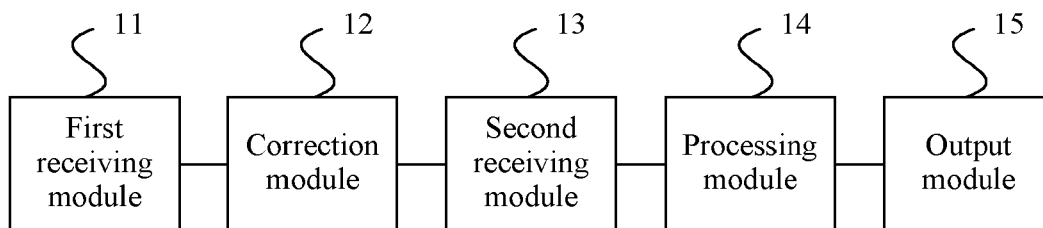


FIG. 7

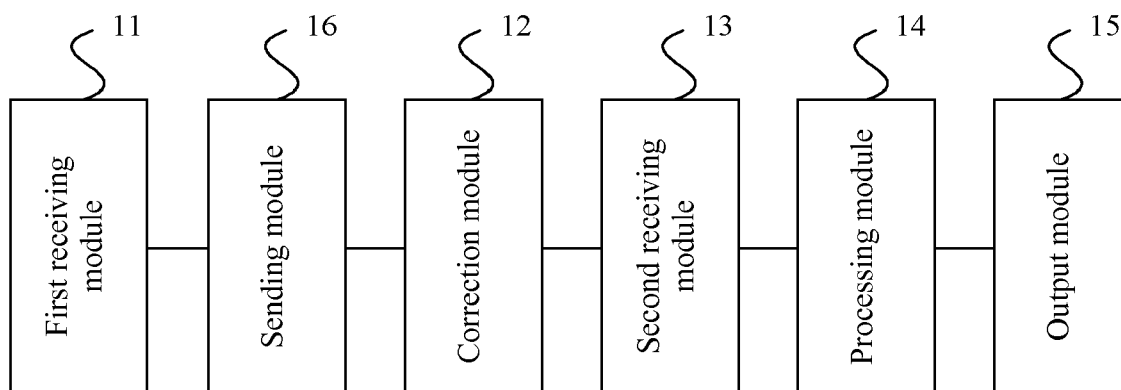


FIG. 8

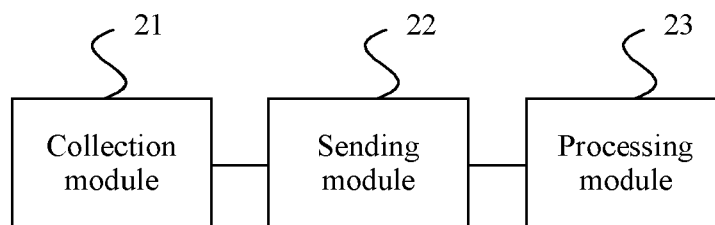


FIG. 9

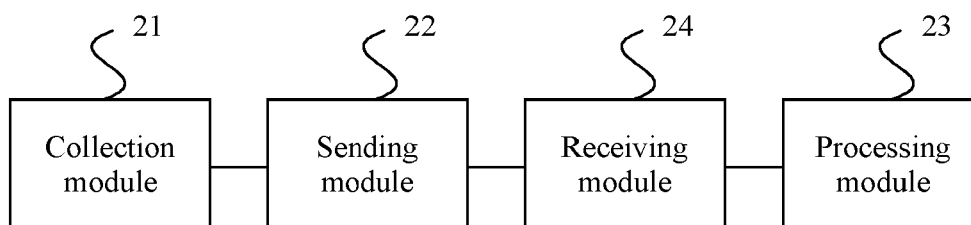


FIG. 10

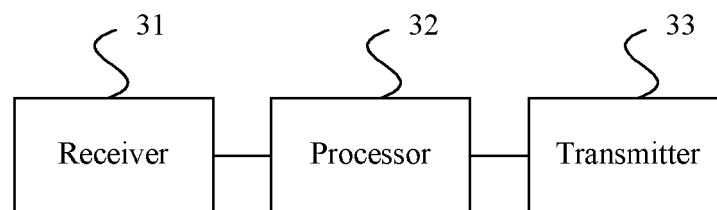


FIG. 11

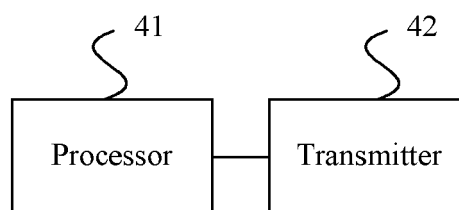


FIG. 12

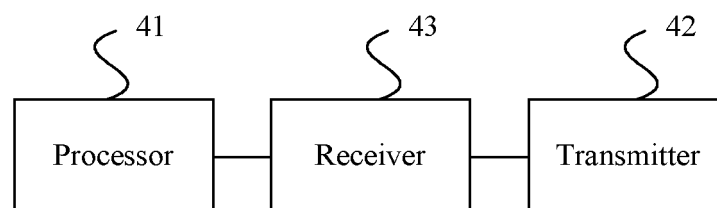


FIG. 13



**REFERENCES CITED IN THE DESCRIPTION**

*This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.*

**Patent documents cited in the description**

- US 2008304571 A1 [0005]