



(12) 发明专利

(10) 授权公告号 CN 107409234 B

(45) 授权公告日 2020.12.25

(21) 申请号 201680013303.4

(22) 申请日 2016.03.03

(65) 同一申请的已公布的文献号
申请公布号 CN 107409234 A

(43) 申请公布日 2017.11.28

(30) 优先权数据

62/128,380 2015.03.04 US

62/128,943 2015.03.05 US

15/058,963 2016.03.02 US

(85) PCT国际申请进入国家阶段日
2017.09.01

(86) PCT国际申请的申请数据
PCT/US2016/020652 2016.03.03

(87) PCT国际申请的公布数据
WO2016/141165 EN 2016.09.09

(73) 专利权人 高通股份有限公司
地址 美国加利福尼亚

(72) 发明人 T·施托克哈默 G·K·沃克
Y-K·王

(74) 专利代理机构 永新专利商标代理有限公司
72002
代理人 陈松涛 韩宏

(51) Int.Cl.
H04N 21/2343 (2011.01)
H04N 21/262 (2011.01)
H04N 21/61 (2011.01)
H04N 21/845 (2011.01)

(56) 对比文件
CN 105594219 A, 2016.05.18
WO 2014/063730 A1, 2014.05.01
CN 104040993 A, 2014.09.10

审查员 龙玄耀

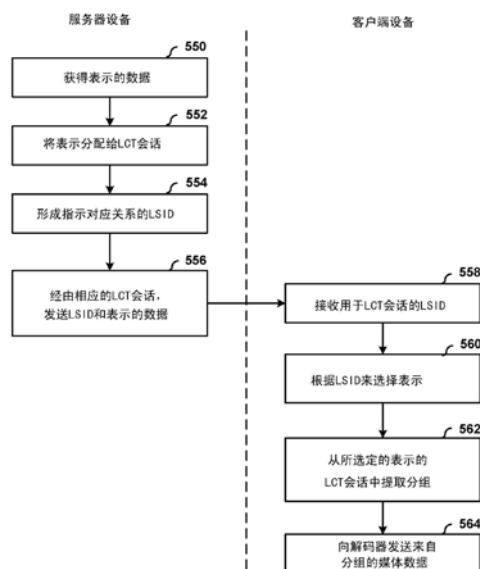
权利要求书5页 说明书36页 附图17页

(54) 发明名称

基于LCT利用DASH格式的基于文件格式的流式传输

(57) 摘要

在一个例子中,一种设备,包括:一个或多个媒体解码器,其被配置为对媒体数据进行解码;网络接口,其被配置为接收分层编码传输(LCT)会话实例描述(LSID),其中,LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括DASH媒体呈现的多个表示中的相应表示的数据以及LCT会话中的一个或多个LCT会话的数据;以及处理器,其被配置为使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费DASH媒体呈现的表示中的一个或多个表示,其中,为了开始消费,处理器被配置为经由网络接口来接收LCT会话的分组,其中,LCT会话的分组包括所述表示中的一个或多个表示的数据中的部分;并且将分组的数据提供给一个或多个媒体解码器。



1. 一种接收媒体数据的方法,所述方法包括:

根据分层编码传输LCT会话实例描述LSID来确定基于HTTP的动态自适应流DASH媒体呈现的多个表示,其中,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括所述表示中的相应表示的数据;以及

基于1)使用所述LSID并且基于2)不使用清单文件生成针对所述DASH媒体呈现的一个或多个表示的一部分数据的请求,来开始消费所述DASH媒体呈现的所述一个或多个表示,其中,所述开始消费包括:

接收包括所述LCT会话的分组的第一数据集直到第一播放时间为止,其中,所述LCT会话的所述分组包括所述表示中的所述一个或多个表示的所述部分的数据;以及

将所述分组的数据提供给媒体解码器;

接收所述清单文件;以及

使用所述清单文件来接收所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

2. 根据权利要求1所述的方法,还包括:

根据所述LSID的一个或多个内容描述符来确定所述DASH媒体呈现的所述表示的编码特性或渲染特性中的至少一个;以及

基于所确定的编码特性或者渲染特性来选择所述表示中的所述一个或多个表示。

3. 根据权利要求2所述的方法,其中,所述一个或多个编码特性或者渲染特性包括以下中的一个或多个:编解码器、可访问性信息、质量、空间分辨率、视点、评级、适配集的简档属性、样本宽高比、帧速率、音频采样速率、MIME类型、扫描类型、帧封装信息、音频信道配置、内容准备、基本属性、补充属性、或者带内事件流。

4. 根据权利要求1所述的方法,还包括:

使用所述清单文件对所述DASH媒体呈现的数据的广播和单播传送进行组合。

5. 根据权利要求1所述的方法,其中,所述DASH媒体呈现提供:位于所述DASH媒体呈现的第一组随机访问点RAP处的第一多个轻量级清单文件、以及位于第二组RAP处的第二多个完整清单文件,其中,与所述第一组RAP相比,所述第二组RAP较小。

6. 根据权利要求1所述的方法,还包括:

接收所述LCT会话的所述分组中的用于指示目标传输时间的数据。

7. 根据权利要求6所述的方法,其中,接收用于指示所述目标传输时间的所述数据包括:接收所述分组的LCT报头的拥塞控制信息字段或者所述LCT报头的报头扩展字段中的用于指示所述目标传输时间的所述数据。

8. 根据权利要求6所述的方法,其中,所述目标传输时间被表达成相对于所述LCT会话的其它分组的相对时间或者绝对挂钟时间中的一种。

9. 根据权利要求6所述的方法,其中,所述目标传输时间是相对于所述LCT会话的其它分组的目標传输时间来表达的,所述方法还包括:接收用于指示在所述分组中的至少一些分组中示意的发布时间的数据。

10. 根据权利要求1所述的方法,其中,所述一个或多个表示中的至少一个表示包括:根据DASH分段格式来格式化的初始化分段和一个或多个媒体分段,并且其中,包括用于所述初始化分段或者所述媒体分段的数据的分组还包括LCT报头。

11. 根据权利要求10所述的方法,还包括:

针对所述分组中的每一个分组,根据所述分组的所述LCT报头的代码点字段来确定:所述分组对应的分段的类型、所述分组是否包括ROUTE报头、是否能够针对所述分组来示意时间轴不连续性、所述分组是否对应于冗余初始化段、以及所述分组是否对应于辅助初始化段。

12. 根据权利要求1所述的方法,还包括:

使用所述LCT会话描述的所述分组的LCT报头的传输会话标识符(TSI)字段来确定所述LCT会话与所述表示之间的对应关系。

13. 根据权利要求1所述的方法,还包括:

根据以下中的至少一个来确定所述LCT会话的分组的数据的发布时间:所述分组的LCT报头的协议特定指示(PSI)比特、或者所述分组的所述LCT报头的扩展报头。

14. 根据权利要求1所述的方法,还包括:

根据所述分组的LCT报头的拥塞控制信息来确定所述LCT会话的分组的目标传输时间。

15. 根据权利要求1所述的方法,还包括:

根据在所述LCT会话的分组的LCT报头中示意的传输对象标识符(TOI)来确定所述DASH媒体呈现的媒体分段的序列号。

16. 一种用于接收媒体数据的设备,所述设备包括:

一个或多个媒体解码器,所述一个或多个媒体解码器被配置为对媒体数据进行解码;

网络接口,所述网络接口被配置为接收分层编码传输LCT会话实例描述LSID,其中,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括:基于HTTP的动态自适应流DASH媒体呈现的多个表示中的相应表示的数据、以及所述LCT会话中的一个或多个LCT会话的数据;以及

处理器,所述处理器被配置为基于1)使用所述LSID并且基于2)不使用清单文件生成针对所述DASH媒体呈现的一个或多个表示的一部分数据的请求,来开始消费所述DASH媒体呈现的所述一个或多个表示,其中,为了开始消费,所述处理器被配置为:

经由所述网络接口来接收包括所述LCT会话的分组的第一数据集直到第一播放时间为止,其中,所述LCT会话的所述分组包括所述表示中的所述一个或多个表示的所述部分的数据;以及

将所述分组的数据提供给所述一个或多个媒体解码器;

接收所述清单文件;以及

使用所述清单文件来接收所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

17. 根据权利要求16所述的设备,其中,所述处理器还被配置为:

根据所述LSID的一个或多个内容描述符来确定所述DASH媒体呈现的所述表示的编码特性或渲染特性中的至少一个;以及

基于所确定的编码特性或者渲染特性来选择所述表示中的所述一个或多个表示。

18. 根据权利要求16所述的设备,其中,所述处理器还被配置为:

根据所述LCT会话的所述分组来确定目标传输时间,并且使用所述目标传输时间向所述一个或多个媒体解码器提供所述分组的数据。

19. 根据权利要求16所述的设备, 其中, 所述一个或多个表示中的至少一个表示包括: 根据DASH分段格式来格式化的初始化分段和一个或多个媒体分段, 其中, 包括用于所述初始化分段或者所述媒体分段的数据的所述分组还包括LCT报头。

20. 根据权利要求19所述的设备, 其中, 所述处理器还被配置为:

针对所述分组中的每一个分组, 根据所述分组的所述LCT报头的代码点字段来确定: 所述分组对应的分段的类型、所述分组是否包括ROUTE报头、是否可以针对所述分组示意时间轴不连续性、所述分组是否对应于冗余初始化分段、以及所述分组是否对应于辅助初始化分段。

21. 根据权利要求16所述的设备, 其中, 所述处理器还被配置为:

根据所述LCT会话的所述分组的LCT报头的传输会话标识符(TSI)字段来确定所述LCT会话与所述表示之间的对应关系。

22. 根据权利要求16所述的设备, 其中, 所述处理器还被配置为:

根据以下中的至少一个来确定所述LCT会话的分组的数据的发布时间: 所述分组的LCT报头的协议特定指示(PSI)比特、或者所述分组的所述LCT报头的扩展报头。

23. 根据权利要求16所述的设备, 其中, 所述处理器还被配置为:

根据在所述LCT会话的分组的LCT报头中示意的传输对象标识符(TOI)来确定所述DASH媒体呈现的媒体分段的序列号。

24. 根据权利要求16所述的设备, 其中, 所述设备包括以下中的至少一个:

集成电路;

微处理器; 或者

无线通信设备。

25. 一种用于接收媒体数据的设备, 所述设备包括:

用于根据分层编码传输LCT会话实例描述LSID来确定基于HTTP的动态自适应流DASH媒体呈现的多个表示的单元, 其中, 所述LSID包括表示多个LCT会话的信息, 所述LCT会话中的每个LCT会话包括所述表示中的相应表示的数据; 以及

用于基于1) 使用所述LSID并且基于2) 不使用清单文件生成针对所述DASH媒体呈现的一个或多个表示的一部分数据的请求, 来开始消费所述DASH媒体呈现的所述一个或多个表示的单元, 其中, 用于开始消费的所述单元包括:

用于接收包括所述LCT会话的分组的第一数据集直到第一播放时间为止的单元, 其中, 所述LCT会话的所述分组包括所述表示中的所述一个或多个表示的所述部分的数据; 以及

用于将所述分组的数据提供给媒体解码器的单元;

用于接收所述清单文件的单元; 以及

用于使用所述清单文件来接收所述DASH媒体呈现的与所述第一数据集不同的第二数据集的单元, 其中, 所述第二数据集具有在所述第一播放时间之后的播放时间。

26. 一种其上存储有指令的计算机可读存储介质, 所述指令当被执行时, 使得用于接收媒体数据的设备的处理器执行以下操作:

根据分层编码传输LCT会话实例描述LSID来确定基于HTTP的动态自适应流DASH媒体呈现的多个表示, 其中, 所述LSID包括表示多个LCT会话的信息, 所述LCT会话中的每个LCT会话包括所述表示中的相应表示的数据;

基于1)使用所述LSID并且基于2)不使用清单文件生成针对所述DASH媒体呈现的一个或多个表示的一部分数据的请求,来开始消费所述DASH媒体呈现的所述一个或多个表示,其中,使所述处理器开始消费的所述指令包括使所述处理器执行以下操作的指令:

接收包括所述LCT会话的分组的第一数据集直到第一播放时间为止,其中,所述LCT会话的所述分组包括所述表示中的所述一个或多个表示的所述部分的数据;以及

将所述分组的数据提供给媒体解码器;

接收所述清单文件;以及

使用所述清单文件来接收所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

27.一种发送媒体数据的方法,所述方法包括:

构造分层编码传输LCT会话实例描述LSID,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括基于HTTP的动态自适应流DASH媒体呈现的多个表示中的相应表示的数据,其中,所述LSID指示所述LCT会话与所述表示之间的对应关系;

输出所述LSID;

在不接收基于清单文件的针对所述表示的数据的请求的情况下,在相对应的LCT会话中输出包括所述表示的所述数据的第一数据集直到第一播放时间为止;

输出所述清单文件;以及

响应于基于所述清单文件的请求,输出所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

28.一种用于发送媒体数据的设备,所述设备包括:

网络接口,所述网络接口用于输出多个分层编码传输LCT会话的数据;以及

处理器,所述处理器被配置为:

构造LCT会话实例描述LSID,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括基于HTTP的动态自适应流DASH媒体呈现的多个表示中的相应表示的数据,其中,所述LSID指示所述LCT会话与所述表示之间的对应关系;

经由所述网络接口来输出所述LSID;

在不接收基于清单文件的针对所述表示的数据的请求的情况下,经由所述网络接口在相对应的LCT会话中输出包括所述表示的所述数据的第一数据集直到第一播放时间为止;

输出所述清单文件;以及

响应于基于所述清单文件的请求,输出所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

29.一种用于发送媒体数据的设备,所述设备包括:

用于构造分层编码传输LCT会话实例描述LSID的单元,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括基于HTTP的动态自适应流DASH媒体呈现的多个表示中的相应表示的数据,其中,所述LSID指示所述LCT会话与所述表示之间的对应关系;

用于输出所述LSID的单元;

用于在不接收基于清单文件的针对所述表示的数据的请求的情况下,在相对应的LCT会话中输出包括所述表示的所述数据的第一数据集直到第一播放时间为止的单元;

用于输出所述清单文件的单元;以及

用于响应于基于所述清单文件的请求,输出所述DASH媒体呈现的与所述第一数据集不同的第二数据集的单元,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

30.一种其上存储有指令的计算机可读存储介质,所述指令当被执行时,使得用于发送媒体数据的设备的处理器执行以下操作:

构造分层编码传输LCT会话实例描述LSID,所述LSID包括表示多个LCT会话的信息,所述LCT会话中的每个LCT会话包括基于HTTP的动态自适应流DASH媒体呈现的多个表示中的相应表示的数据,其中,所述LSID指示所述LCT会话与所述表示之间的对应关系;

输出所述LSID;

在不接收基于清单文件的针对所述表示的数据的请求的情况下,在相对应的LCT会话中输出包括所述表示的所述数据的第一数据集直到第一播放时间为止;

输出所述清单文件;以及

响应于基于所述清单文件的请求,输出所述DASH媒体呈现的与所述第一数据集不同的第二数据集,其中,所述第二数据集具有在所述第一播放时间之后的播放时间。

基于LCT利用DASH格式的基于文件格式的流式传输

[0001] 本申请要求享有于2015年3月4日提交的美国临时申请No.62/128,380和2015年3月5日提交的美国临时申请No.62/128,943的优先权,这两个申请中的每一个的全部内容以引用方式并入本文。

技术领域

[0002] 本公开内容涉及编码的视频数据的存储和传输。

背景技术

[0003] 数字视频能力可以并入到各种各样的设备中,其包括数字电视、数字直接广播系统、无线广播系统、个人数字助理(PDA)、膝上型计算机或台式计算机、数码相机、数字记录设备、数字媒体播放器、视频游戏设备、视频游戏控制台、蜂窝或卫星无线电话、视频会议设备等等。数字视频设备实施视频压缩技术,例如,在由MPEG-2、MPEG-4、ITU-T H.263或ITU-T H.264/MPEG-4、第10部分、高级视频编码(AVC)、ITU-T H.265/MPEG-H第2部分(其还被称为高效率视频编码(HEVC))所规定的标准、以及这些标准的扩展里所描述的那些技术,以更高效地发送和接收数字视频信息。

[0004] 视频压缩技术执行空间预测和/或时间预测以减少或者消除视频序列中固有的冗余。对于基于块的视频编码而言,可以将视频帧或片划分成宏块。每个宏块还可以进一步划分。使用相对于相邻宏块的空间预测,对帧内编码(I)帧或者片中的宏块进行编码。帧间编码(P或B)帧或者片中的宏块可以使用相对于相同帧或者片中的相邻宏块的空间预测或者相对于其它参考帧的时间预测。

[0005] 在对视频数据进行编码之后,可以对视频数据进行分组化以进行传输或者存储。可以将视频数据组合成符合各种标准中的任何一种的视频文件,例如,基于国际标准化组织(ISO)的媒体文件格式以及其扩展(例如,AVC)。

发明内容

[0006] 概括而言,本公开内容描述了用于在不使用清单文件(例如,用于表示的媒体呈现描述(MPD))的情况下访问分层编码传输(LCT)所携带的一个或多个表示的媒体数据的各种技术。例如,LCT会话实例描述(LSID)可以包括用于启动访问媒体数据的服务和/或该服务的连续操作的清单文件数据中的至少一些。例如,LSID可以包括用于指示表示的属性的信息。另外地或替代地,LCT会话的分组可以包括具有分配的数据的LCT报头,其中该分配的数据指示分组如何与表示的段相关,例如,哪些分组对应于每个分段(segment)。用此方式,客户端设备可以在未接收到清单文件的情况下(或者在此之前),开始消费(consumption)表示中的一个或多个表示。

[0007] 在一个例子中,一种接收媒体数据的方法,包括:根据分层编码传输(LCT)会话实例描述(LSID)来确定基于HTTP的动态自适应流(DASH)媒体呈现的多个表示,其中,该LSID包括表示多个LCT会话的信息,该LCT会话中的每个LCT会话包括所述表示中的相应表示的

数据;以及使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的所述表示中的一个或多个表示,其中,开始消费包括:接收LCT会话的分组,其中,LCT会话的分组包括所述表示中的所述一个或多个表示的数据中的部分;以及将分组的数据提供给媒体解码器。

[0008] 在另一个例子中,一种用于接收媒体数据的设备,包括:一个或多个媒体解码器,其被配置为对媒体数据进行解码;网络接口,其被配置为接收分层编码传输(LCT)会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据、以及LCT会话中的一个或多个LCT会话的数据;以及处理器,其被配置为使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的所述表示中的一个或多个表示,其中,为了开始消费,所述处理器被配置为:经由网络接口来接收LCT会话的分组,其中,LCT会话的分组包括所述表示中的所述一个或多个表示的数据中的部分;以及将分组的数据提供给一个或多个媒体解码器。

[0009] 在另一个例子中,一种用于接收媒体数据的设备,包括:用于根据分层编码传输(LCT)会话实例描述(LSID)来确定基于HTTP的动态自适应流(DASH)媒体呈现的多个表示的单元,其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括所述表示中的相应表示的数据;以及用于使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的所述表示中的一个或多个表示的单元,其中,用于开始消费的单元包括:用于接收LCT会话的分组的单元,其中,LCT会话的分组包括所述表示中的所述一个或多个表示的数据中的部分;以及用于将分组的数据提供给媒体解码器的单元。

[0010] 在另一个例子中,一种其上存储有指令的计算机可读存储介质,所述指令当被执行时,使得用于接收媒体数据的设备的处理器执行以下操作:根据分层编码传输(LCT)会话实例描述(LSID)来确定基于HTTP的动态自适应流(DASH)媒体呈现的多个表示,其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括所述表示中的相应表示的数据;以及使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的所述表示中的一个或多个表示,其中,用于使所述处理器开始消费的指令包括用于使所述处理器执行以下操作的指令:接收LCT会话的分组,其中,LCT会话的分组包括所述表示中的所述一个或多个表示的数据中的部分;以及将分组的数据提供给媒体解码器。

[0011] 在另一个例子中,一种用于发送媒体数据的方法,包括:构造分层编码传输(LCT)会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示LCT会话与所述表示之间的对应关系;输出该LSID;以及在相对应的LCT会话中输出所述表示的数据。

[0012] 在另一个例子中,一种用于发送媒体数据的设备,包括:网络接口。其用于输出多个分层编码传输(LCT)会话的数据;以及处理器,其被配置为:构造LCT会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示LCT会话与所述表示之间的对应关系;经由网络接口来输出该LSID;经由网络接口在相对应的LCT会话中输出所述表示的数据。

[0013] 在另一个例子中,一种用于发送媒体数据的设备,包括:用于构造分层编码传输(LCT)会话实例描述(LSID)的单元,其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示LCT会话与所述表示之间的对应关系;用于输出该LSID的单元;以及用于在相对应的LCT会话中输出所述表示的数据的单元。

[0014] 在另一个例子中,一种其上存储有指令的计算机可读存储介质,所述指令当被执行时,使得用于发送媒体数据的设备的处理器执行以下操作:构造分层编码传输(LCT)会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示LCT会话与所述表示之间的对应关系;输出该LSID;以及在相对应的LCT会话中输出所述表示的数据。

附图说明

[0015] 图1是示出实施用于通过网络来流式传输媒体数据的技术的示例性系统的框图。

[0016] 图2是示出示例性多媒体内容的元素的概念图。

[0017] 图3是示出一种示例性视频文件的元素的概念图,其可以对应于表示的一个分段。

[0018] 图4是示出在广播传送中经常出现的示例性场景的概念图。

[0019] 图5是示出可以执行本公开内容的技术的示例性系统的概念图。

[0020] 图6是示出可以执行本公开内容的技术的示例性系统的概念图。

[0021] 图7是示出基于FLUTE的DASH的例子中,服务条目的不同方面的概念图。

[0022] 图8是示出基于ROUTE的DASH的例子中,服务条目的不同方面的概念图。

[0023] 图9是根据本公开内容的技术,示出可以用于携带数据的根据RFC5651的一组示例性报头字段的概念图。

[0024] 图10是示出当对象的前缀可以发布到下一层进行解码时,用于信令的各种选项的概念图。

[0025] 图11是示出用于数据接收的示例性模型的概念图。

[0026] 图12是示出用于接收媒体数据的示例性系统的概念图。

[0027] 图13是示出一种示例性发送过程的概念图。

[0028] 图14是示出一种示例性混合DASH客户端模型的概念图。

[0029] 图15是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的示例性方法的流程图。

[0030] 图16是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的另一种示例性方法的流程图。

[0031] 图17是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的另一种示例性方法的流程图。

具体实施方式

[0032] 概括地说,本公开内容描述了一种技术,其实现了使用单向协议的能力(例如,单向传输文件传送(FLUTE)、单向传输实时对象传送(ROUTE)或者用于异步分层编码的对象传

送和面向NACK的可靠多播协议 (FCAST)、以及基于ISO基本媒体文件格式 (ISO BMFF) 和可能的其它格式 (例如, MPEG-2传输流 (TS)) 的基于HTTP的动态自适应流 (DASH) 分段), 以便创建支持媒体流的传送和播放的基于IP的广播媒体传送系统, 其在于:

[0033] ●彼此之间同步 (由ISO BMFF处理)。

[0034] ●随机访问 (由传送协议中的特定信令进行处理)。

[0035] ●将被播放以使得不发生重新缓冲和停滞。

[0036] ●与通过宽带和单播提供和供应的媒体流组合。

[0037] ●实现多程序传送和复用。

[0038] ●实现低启动延迟和快速信道更改。

[0039] ●在发送器和接收器处实现拼接内容。

[0040] ●在广播分发系统中提供DASH的所有特征。

[0041] 具体而言, 本公开内容的技术允许在不接收用于表示的清单文件 (例如, 媒体呈现描述 (MPD)) 的情况下 (或在其之前), 经由分层编码传输 (LCT) 会话来接收一个或多个表示的媒体数据。用此方式, 本公开内容的技术可以减少与开始媒体数据的消费 (例如, 执行服务启动) 相关联的延迟。也就是说, 如果没有这些技术, 则客户端设备将需要等待清单文件的接收, 这可能造成较差的用户体验 (例如, 看到黑屏和/或听不到音频播放)。使用这些技术可以减少延迟, 以使得改善用户体验 (例如, 允许音频和/或视频数据的更快速播放)。

[0042] 本公开内容的技术可以应用于遵循根据以下文件格式中的任何一种封装的视频数据的视频文件: ISO基本媒体文件格式、可伸缩视频编码 (SVC) 文件格式、高级视频编码 (AVC) 文件格式、第三代合作伙伴计划 (3GPP) 文件格式、和/或多视角视频编码 (MVC) 文件格式或者其它类似的视频文件格式。也就是说, 可以根据这些各种格式中的任何一种来形成表示的分段。通常, 分段代表相应表示的独立可接收的媒体文件。

[0043] 在HTTP流式传输中, 经常使用的操作包括HEAD、GET和部分GET。HEAD操作获取与给定的统一资源定位符 (URL) 或者统一资源名称 (URN) 相关联的文件的报头, 而不获取与该URL或URN相关联的有效载荷。GET操作获取与给定的URL或URN相关联的整个文件。部分GET操作接收字节范围作为输入参数, 并获取文件的连续数量的字节, 其中该字节的数量与所接收的字节范围相对应。因此, 可以提供电影片段以进行HTTP流式传输, 这是由于部分GET操作可以获得一个或多个单个的电影片段。在电影片段中, 可以存在不同轨道的若干轨道片段。在HTTP流式传输中, 媒体呈现可以是客户端可访问的结构化的数据集合。客户端可以请求和下载媒体数据信息, 以向用户呈现流服务。

[0044] 在使用HTTP流式传输对3GPP数据进行流式传输的例子中, 可能存在针对多媒体内容的视频和/或音频数据的多个表示。如下面所解释的, 不同的表示可以对应于不同的编码特性 (例如, 视频编码标准的不同简档或者水平)、不同的编码标准或者编码标准的扩展 (例如, 多视角和/或可伸缩扩展) 或者不同的比特率。可以用媒体呈现描述 (MPD) 数据结构定义这些表示的清单。媒体呈现可以对应于HTTP流式传输客户端设备可访问的结构化的数据集合。HTTP流式传输客户端设备可以请求和下载媒体数据信息, 以向客户端设备的用户呈现流服务。可以用MPD数据结构来描述媒体呈现, 其中MPD数据结构可以包括MPD的更新。

[0045] 媒体呈现可以包含一个或多个时段 (period) 的序列。可以通过MPD中的Period (时段) 元素来规定时段。每个时段可以具有MPD中的属性start (开始)。MPD可以包括针对每个

时段的start属性和availableStartTime(可用开始时间)属性。对于实时服务而言,该时段的start属性和MPD属性availableStartTime的总和,可以以UTC格式来指示该时段的可用时间(具体而言,相应的时段中的每个表示的第一媒体分段)。对于按需服务而言,第一时段的start属性可以是0。对于任何其它时段而言,start属性可以指示相应的时段的开始时间相对于第一时段的开始时间的偏移。每个时段可以扩展到下个时段的开始,或者扩展到媒体呈现的结束为止(在最后的时段的情况下)。时段开始时间可以是精确的。它们可以反映源自于播放所有在先时段的实际时间。

[0046] 每个时段可以包含相同媒体内容的一个或多个表示。表示可以是音频或视频数据的多个替代编码版本中的一个。这些表示可以依据编码类型(例如,用于视频数据的比特率、分辨率、和/或编解码器、以及用于音频数据的比特率、语言和/或编解码器)而不同。可以使用术语表示来指代:与多媒体内容的特定时段相对应、并以特定的方式进行编码的经编码的音频或视频数据的一部分。

[0047] 可以将特定时段的表示分配给通过MPD中的属性所指示的组,其中该属性指示这些表示属于的适配集。相同适配集中的表示通常认为是彼此替代的,其在于:客户端设备可以动态地和无缝地在这些表示之间切换,例如,以执行带宽适配。例如,可以将特定时段的视频数据的每个表示分配给相同的适配集,以使得可以选择这些表示中的任何一个来解码,以呈现针对相应的时段的多媒体内容的媒体数据(例如,视频数据或音频数据)。在一些例子中,一个时段中的媒体内容可以通过来自组0的一个表示(如果存在的话),或者来自每个非零组的至多一个表示的组合来表示。一个时段的每个表示的定时数据(timing data)可以相对于该时段的开始时间来表达。

[0048] 一个表示可以包括一个或多个分段。每个表示可以包括初始化分段,或者表示的每个分段可以是自初始化的。当存在时,初始化分段可以包含用于访问该表示的初始化信息。通常,初始化分段不包含媒体数据。分段可以通过诸如统一资源定位符(URL)、统一资源名称(URN)或者统一资源标识符(URI)之类的标识符来唯一地引用。MPD可以为每个分段提供标识符。在一些例子中,MPD还可以用range(范围)属性的形式来提供字节范围,其中该range属性可以对应于:用于可通过URL、URN或URI访问的文件内的分段的数据。

[0049] 针对不同类型的媒体数据,可以选择不同的表示以基本同时获取。例如,客户端设备可以选择从其获取分段的音频表示、视频表示和定时文本表示。在一些例子中,客户端设备可以选择特定的适配集来执行带宽适配。也就是说,客户端设备可以选择包括视频表示的适配集、包括音频表示的适配集、和/或包括定时文本的适配集。替代地,客户端设备可以选择用于某些类型的媒体(例如,视频)的适配集,并且直接选择用于其它类型的媒体(例如,音频和/或定时文本)的表示。

[0050] 图1是示出实施本公开内容的技术以通过网络来流式传输媒体数据的示例性系统10的框图。系统10的各个元素通常可以对应于图5和图6中所示出的例子的类似元素,如下面所进一步详细讨论的。同样,客户端设备40的组件(component)通常可以对应于图11、12和图14的组件,也如下面所进一步详细讨论的。

[0051] 在图1的例子中,系统10包括内容准备设备20、服务器设备60和客户端设备40。客户端设备40和服务器设备60通过网络74来通信地耦合,其中网络74可以包括互联网。在一些例子中,内容准备设备20和服务器设备60还可以通过网络74或者另一个网络来耦合,或

者可以直接地通信耦合。在一些例子中,内容准备设备20和服务器设备60可以包括相同的设备。

[0052] 在图1的例子中,内容准备设备20包括音频源22和视频源24。例如,音频源22可以包括麦克风,该麦克风产生表示将由音频编码器26进行编码的经捕获的音频数据的电信号。替代地,音频源22可以包括:用于存储先前记录的音频数据的存储介质、诸如计算机化合成器之类的音频数据发生器、或者音频数据的任何其它源。视频源24可以包括:视频照相机,其产生由视频编码器28进行编码的视频数据、利用先前记录的视频数据进行编码的存储介质、诸如计算机图形源之类的视频数据生成单元、或者视频数据的任何其它源。内容准备设备20并不需要在所有例子中都通信地耦合到服务器设备60,而可以将多媒体内容存储在由服务器设备60进行读取的单独介质中。

[0053] 原始音频和视频数据可以包括模拟或数字数据。在音频编码器26和/或视频编码器28对模拟数据进行编码之前,可以对其进行数字化。音频源22可以在说话参与者正在说话时从该说话参与者获得音频数据,并且视频源24可以同时地获得说话参与者的视频数据。在其它例子中,音频源22可以包括具有存储的音频数据的计算机可读存储介质,视频源24可以包括具有存储的视频数据的计算机可读存储介质。用此方式,本公开内容中所描述的技术可以应用于实况、流式传输、实时音频和视频数据,或者也可以应用于归档的、预先录制的音频和视频数据。

[0054] 与视频帧相对应的音频帧通常是包含音频数据的音频帧,该音频数据由音频源22与包含在视频帧内的由视频源24所捕获(或者生成)的视频数据同时地捕获(或生成)。例如,虽然说话参与者通常通过说话产生音频数据,音频源22捕获音频数据,同时视频源24捕获说话参与者的视频数据(也就是说,在音频源22捕获音频数据时)。因此,音频帧可以在时间上与一个或多个特定的视频帧相对应。因此,与视频帧相对应的音频帧通常对应于以下情形:在相同时间捕获音频数据和视频数据,并且音频帧和视频帧分别包括在相同时间捕获的音频数据和视频数据。

[0055] 在一些例子中,音频编码器26可以对每个经编码的音频帧中的时间戳进行编码,其中该时间戳表示对经编码的音频帧的音频数据进行记录的时间,类似地,视频编码器28可以对每个经编码的视频帧中的时间戳进行编码,其中该时间戳表示对经编码的视频帧的视频数据进行记录的时间。在这些例子中,与视频帧相对应的音频帧可以包括:包括时间戳的音频帧和包括相同的时间戳的视频帧。内容准备设备20可以包括内部时钟,其中音频编码器26和/或视频编码器28可以根据该内部时钟来生成时间戳,或者音频源22和视频源24可以使用该内部时钟将音频数据和视频数据分别与时间戳进行关联。

[0056] 在一些例子中,音频源22可以向音频编码器26发送与记录音频数据的时间相对应的数据,视频源24可以向视频编码器28发送与记录视频数据的时间相对应的数据。在一些例子中,音频编码器26可以对经编码的音频数据中的序列标识符进行编码,以指示经编码的音频数据的相对时间顺序,而无需指示记录该音频数据的绝对时间,类似地,视频编码器28也可以使用序列标识符来指示经编码的视频数据的相对时间顺序。类似地,在一些例子中,可以使序列标识符与时间戳相映射或者相关。

[0057] 音频编码器26通常产生经编码的音频数据的流,而视频编码器28产生经编码的视频数据的流。每个单独的数据流(无论是音频还是视频)可以称为基本流。基本流是表示的

单个数字编码的(可能被压缩的)组分(component)。例如,表示的经编码的视频或者音频部分可以是基本流。在将基本流封装在视频文件内之前,可以将其转换成分组化基本流(PES)。在相同的表示之内,可以使用流ID来区分属于一个基本流的PES分组与其它PES分组。基本流的数据的基本单元是分组化基本流(PES)分组。因此,经编码的视频数据通常与基本视频流相对应。类似地,音频数据与一个或多个相应的基本流相对应。

[0058] 诸如ITU-T H.264/AVC和即将推出的高效视频编码(HEVC)标准之类的很多视频编码标准,定义了针对无错比特流的语法、语义和解码过程,其中任何一个都符合某种简档或级别。视频编码标准通常并不指定编码器,但编码器的任务是保证生成的比特流对于解码器是符合标准的。在视频编码标准的背景下,“简档”与应用于它们的算法、特征或工具和约束的一个子集相对应。如H.264标准所规定的,例如,“简档”是H.264标准所指定的整个比特流语法的子集。“级别”与解码器资源消耗的限制(例如,解码器存储器和计算)相对应,其与图片的分辨率、比特速率和块处理速率相关。可以使用profile_idc(简档指示符)值来示意(signal)简档,而使用level_idc(级别指示符)值来示意级别。

[0059] 例如,H.264标准认识到,在由给定简档的语法所施加的界限之内,根据比特流中的语法元素所采用的值(例如,解码图像的指定大小),可能仍然需要编码器和解码器的性能的较大变化。此外,H.264标准还认识到,在很多应用中,实现一个能够处理特定简档内的语法的所有假设用法的解码器既不实际也不经济。因此,H.264标准将“级别”规定成被施加在比特流中的语法元素的值上的一组指定约束。这些约束可以是针对值的简单限制。替代地,这些约束可以采取对值的算术组合(例如,图像宽度乘以图像高度乘以每秒解码的图像数量)的约束的形式。此外,H.264标准还提供了:针对每个支持的简档,各个实施可以支持不同的级别。

[0060] 符合某个简档的解码器通常支持在该简档中定义的所有特征。例如,作为编码特征,B图像编码在H.264/AVC的基线简档中不受支持,但在H.264/AVC的其它简档中被支持。符合某个级别的解码器应当能够对于不需要该级别所定义的限制之外的资源的任何比特流进行解码。简档和级别的定义可以有助于实现互操作性。例如,在视频传输期间,可以针对整个传输会话,协商和协定一对简档和级别定义。具体而言,在H.264/AVC中,级别可以规定关于以下各项的限制:需要进行处理的宏块的数量、解码图片缓冲器(DPB)大小、编码图片缓冲器(CPB)大小、垂直运动矢量范围、每两个连续MB的最大运动矢量数、以及B块是否可以具有小于8x8像素的子宏块分区。用此方式,解码器可以判断该解码器是否能够适当地对比特流进行解码。

[0061] 在图1的例子中,内容准备设备20的封装单元30从视频编码器28接收包括经编码的视频数据的基本流,并且从音频编码器26接收包括经编码的音频数据的基本流。在一些例子中,视频编码器28和音频编码器26均可以包括:用于根据经编码的数据来形成PES分组的分组器。在其它例子中,视频编码器28和音频编码器26均可以与相应的分组器接口连接,以根据经编码的数据来形成PES分组。在其它例子中,封装单元30可以包括:用于根据经编码的音频和视频数据来形成PES分组的分组器。

[0062] 视频编码器28可以以各种方式,对多媒体内容的视频数据进行编码,以按照各种比特率和使用各种特性(例如,像素分辨率、帧速率、符合各种编码标准、符合用于各个编码标准的各个简档和/或简档的级别、具有一个或多个视角(例如,用于二维或三维播放)的表

示、或者其它这些特性),来产生多媒体内容的不同表示。如本公开内容中所使用的,表示可以包括音频数据、视频数据、文本数据(例如,用于闭路字幕)中的一个或者其它这种数据。该表示可以包括基本流,例如,音频基本流或者视频基本流。每个EPS分组可以包括用于标识该PES分组所属于的基本流的stream_id。封装单元30负责将基本流组合到各个表示的视频文件(例如,分段)中。

[0063] 封装单元30从音频编码器26和视频编码器28接收用于表示的基本流的PES分组,并且根据这些PES分组来形成相应的网络抽象层(NAL)单元。在H.264/AVC(高级视频编码)的例子中,将经编码的视频分段组织成NAL单元,其中这些NAL单元提供用于寻址诸如视频通话、存储、广播或者流式传输之类的应用的“网络友好的”视频表示。可以将NAL单元分类成视频编码层(VCL)NAL单元和非VCL NAL单元。VCL单元可以包含核心压缩引擎,可以包括块、宏块和/或片级数据。其它NAL单元可以是非VCL NAL单元。在一些例子中,一个时间实例的编码的图片(其通常呈现为基本编码图片)可以包含在访问单元中,其中该访问单元可以包括一个或多个NAL单元。

[0064] 非VCL NAL单元可以包括参数集NAL单元和SEI NAL单元等等。参数集可以包含序列级别报头信息(在序列参数集(SPS)中)和不频繁改变的图片级别报头信息(在图片参数集(PPS)中)。使用参数集(例如,PPS和SPS),不频繁改变的信息不需要针对每一个序列或者图片都进行重复,因此可以提高编码效率。此外,参数集的使用可以实现重要报头信息的带外传输,其避免了用于差错恢复的冗余传输的需求。在带外传输示例中,可以在与其它NAL单元(例如,SEI NAL单元)不同的信道上,发送参数集NAL单元。

[0065] 补充增强信息(SEI)可以包含对来自VCL NAL单元的经编码的图片样本进行解码不需要的信息,但其可以有助于实现与解码、显示、差错恢复和其它目的有关的过程。SEI消息可以包含在非VCL NAL单元中。SEI消息是一些标准规范的规范化部分,因此其对于标准兼容的解码器实施来说并不总是强制性的。SEI消息可以是序列级别SEI消息或者图片级别SEI消息。可以在SEI消息中包含一些序列级别消息,例如,在SVC的例子中的可伸缩性信息SEI消息和MVC中的视图可伸缩性信息SEI消息。这些示例性SEI消息可以传送例如与以下各项有关的信息:操作点的提取和操作点的特性。此外,封装单元30可以形成清单文件,例如,用于描述表示的特性的媒体呈现描述符(MPD)。封装单元30可以根据可扩展标记语言(XML)对MPD进行格式化。

[0066] 封装单元30可以向输出接口32提供用于多媒体内容的一个或多个表示的数据以及清单文件(例如,MPD)。输出接口32可以包括网络接口或者用于向存储介质写入的接口,例如,通用串行总线(USB)接口、CD或DVD写入器或者刻录器、针对磁存储介质或者闪速存储介质的接口、或者用于存储或发送媒体数据的其它接口。封装单元30可以向输出接口32提供多媒体内容的每一个表示的数据,其中输出接口32可以经由网络传输或存储介质来向服务器设备60发送数据。在图1的例子中,服务器设备60包括用于存储各种多媒体内容64的存储介质62,其中每一个多媒体内容64包括相应的清单文件66和一个或多个表示68A-68N(表示68)。在一些例子中,输出接口32还可以直接向网络74发送数据。

[0067] 在一些例子中,可以将表示68分成一些适配集。也就是说,表示68的各个子集可以包括各自的共同特性集,例如,编解码器、简档和级别、分辨率、视点的数量、用于分段的文件格式、可以标识要用该表示显示的文本的语言或者其它特性和/或要(例如,由扬声器)进

行解码和呈现的音频数据的文本类型信息、可以描述适配集中的表示的风景的相机角度或者真实世界的相机视角的相机角度信息、用于描述特定观众的内容适合性的评级信息等等。

[0068] 清单文件66可以包括用于指示与特定的适配集相对应的表示68的子集以及这些适配集的共同特性的数据。清单文件66还可以包括用于表示适配集的各个表示的各自特性(例如,比特率)的数据。用此方式,适配集可以提供简化的网络带宽适配。可以使用清单文件66的适配集元素的子元素,来指示适配集中的表示。

[0069] 服务器设备60包括请求处理单元70和网络接口72。在一些例子中,服务器设备60可以包括多个网络接口。此外,服务器设备60的特征中的任何一个或全部,可以在内容传送网络的其它设备(例如,路由器、网桥、代理设备、交换机或其它设备)上实施。在一些例子中,内容传送网络的中间设备可以对多媒体内容64的数据进行高速缓存,并且包括基本符合服务器设备60的这些的组件。通常,网络接口72被配置为经由网络74来发送和接收数据。

[0070] 请求处理单元70被配置为从客户端设备(例如,客户端设备40)接收针对于存储介质62的数据的网络请求。例如,请求处理单元70可以实施超文本传输协议(HTTP)版本1.1,如在RFC 2616,“Hypertext Transfer Protocol-HTTP/1.1”by R.Fielding et al, Network Working Group, IETF, June 1999中所描述的。也就是说,请求处理单元70可以被配置为接收HTTP GET或部分GET请求,并响应于这些请求来提供多媒体内容64的数据。这些请求可以例如使用分段的URL,来指示表示68中的一个的分段。在一些例子中,这些请求还可以指定该分段的一个或多个字节范围,因此包括部分GET请求。此外,请求处理单元70还可以被配置为服务HTTP HEAD请求,以提供表示68中的一个的分段的报头数据。无论如何,请求处理单元70可以被配置为对这些请求进行处理,以向请求方设备(例如,客户端设备40)提供请求的数据。

[0071] 另外地或替代地,请求处理单元70可以被配置为经由广播或多播协议(例如,eMBMS)来传送媒体数据。内容准备设备20可以以如上所述基本相同的方式,来生成DASH分段和/或子分段,但服务器设备60可以使用eMBMS或另一种广播或多播网络传输协议来传送这些分段或子分段。例如,请求处理单元70可以被配置为从客户端设备40接收多播组加入请求。也就是说,服务器设备60可以向与特定的媒体内容(例如,直播事件的广播)相关联的客户端设备(其包括客户端设备40)通告与多播组相关联的互联网协议(IP)地址。转而,客户端设备40可以提交加入该多播组的请求。可以贯穿网络74(例如,构成网络74的路由器)来传播该请求,以使得致使这些路由器将去往与该多播组相关联的IP地址的流量指引到订阅方客户端设备(例如,客户端设备40)。

[0072] 如图1的例子中所示出的,多媒体内容64包括清单文件66,后者可以对应于媒体呈现描述(MPD)。清单文件66可以包含不同的替代表示68的描述(例如,具有不同质量的视频服务),该描述可以包括例如编解码器信息、简档值、级别值、比特率和表示68的其它描述性特性。客户端设备40可以获取媒体呈现的MPD,以确定如何访问表示68的分段。

[0073] 具体而言,获取单元52可以获取客户端设备40的配置数据(未示出),以确定视频解码器48的解码能力和视频输出44的渲染能力。该配置数据还可以包括客户端设备40的用户所选择的语言偏好中的任何一个或全部、与客户端设备40的用户所设置的深度偏好相对应的一个或多个相机视角、和/或客户端设备40的用户所选择的评级偏好。例如,获取单元

52可以包括被配置为提交HTTP GET和部分GET请求的web浏览器或媒体客户端。获取单元52可以对应于由客户端设备40的一个或多个处理器或处理单元(未示出)执行的软件指令。在一些例子中,针对获取单元52所描述的功能的全部或者一部分可以用硬件、或者硬件、软件和/或固件的组合来实施,其中在该情况下,可以提供必要的硬件来执行用于软件或固件的指令。

[0074] 获取单元52可以将客户端设备40的解码和渲染能力与清单文件66的信息所指示的表示68的特性进行比较。初始时,获取单元52可以获取清单文件66的至少一部分,以确定表示68的特性。例如,获取单元52可以请求用于描述一个或多个适配集的特性的清单文件66的一部分。获取单元52可以选择表示68的一个子集(例如,适配集),其中该子集具有客户端设备40的编码和渲染能力所满足的特性。然后,获取单元52可以确定该适配集中的表示的比特率,确定网络带宽的当前可用量,并且从具有该网络带宽所能够满足的比特率的表示之一之中获取分段。

[0075] 通常,较高比特率的表示可以产生较高质量的视频播放,而当可用网络带宽减小时,较低比特率的表示可以提供足够质量的视频播放。因此,当可用的网络带宽相对较高时,获取单元52可以从相对较高的比特率表示中获取数据,而当可用的网络带宽较低时,获取单元52可以从相对较低的比特率表示中获取数据。用此方式,客户端设备40可以通过网络74,对多媒体数据进行流式传输,同时还适于改变网络74的网络带宽可用性。

[0076] 另外地或替代地,获取单元52可以被配置为根据广播或多播网络协议(例如,eMBMS或IP多播)来接收数据。在这些例子中,获取单元52可以提交加入与特定的媒体内容相关联的多播网络组的请求。在加入该多播组之后,获取单元52可以接收该多播组的数据,而无需向服务器设备60或内容准备设备20发出另外的请求。当不再需要该多播组的数据时(例如,停止播放或者将信道改变至不同的多播组),获取单元52可以提交离开该多播组的请求。

[0077] 网络接口54可以接收选定的表示的分段的数据,并向获取单元52进行提供该选定的表示的分段的数据,其中获取单元52可以转而将这些分段提供给解封装单元50。解封装单元50可以将视频文件的元素解封装成组成PES流,对PES流进行解分组以获取经编码的数据,并且根据该经编码的数据是音频流的一部分还是视频流的一部分(例如,如该流的PES分组报头所指示的),将该经编码的数据发送给音频解码器46或者视频解码器48。音频解码器46对编码的音频数据进行解码,向音频输出42发送该解码的音频数据,同时视频解码器48对编码的视频数据进行解码,并向视频输出44发送该解码的视频数据(其可以包括流的多个视图)。

[0078] 视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元30、获取单元52和解封装单元50均可以根据需要,被实施为多种适当的处理电路中的任何一种,例如,一个或多个微处理器、数字信号处理器(DSP)、专用集成电路(ASIC)、现场可编程门阵列(FPGA)、分立逻辑电路、软件、硬件、固件或者其任意组合。视频编码器28和视频解码器48中的每一个可以包括在一个或多个编码器或解码器中,它们中的任意一个可以被集成为组合的视频编码器/解码器(CODEC)的一部分。同样,音频编码器26和音频解码器46中的每一个可以包括在一个或多个编码器或解码器中,它们中的任意一个可以被集成为组合的CODEC的一部分。包括有视频编码器28、视频解码器48、音频编码器26、音频解码器46、封装单元

30、获取单元52和/或解封单元50的装置,可以包括集成电路、微处理器和/或诸如蜂窝电话之类的无线通信设备。

[0079] 客户端设备40、服务器设备60和/或内容准备设备20可以被配置为根据本公开内容的技术进行操作。为了举例起见,本公开内容针对客户端设备40和服务器设备60来描述这些技术。但是,应当理解的是,内容准备设备20可以被配置为替代服务器设备60(或者除了服务器设备60之外),来执行这些技术。

[0080] 封装单元30可以形成NAL单元,其包括用于标识该NAL单元所属于的程序的报头以及有效载荷(例如,音频数据、视频数据、或者用于描述与该NAL单元相对应的传输或程序流的数据)。例如,在H.264/AVC中,NAL单元包括1字节报头和可变大小的有效载荷。将视频数据包括在其有效载荷中的NAL单元,可以包括各种粒度水平的视频数据。例如,NAL单元可以包括视频数据块、多个块、视频数据片、或者整个图片的视频数据。封装单元30可以从视频编码器28接收处于基本流的PES分组的形式的经编码的视频数据。封装单元30可以将每一个基本流与相对应的程序进行关联。

[0081] 封装单元30还可以对来自多个NAL单元的访问单元进行组合。通常,访问单元可以包括:用于表示视频数据的帧的一个或多个NAL单元、以及与该帧相对应的音频数据(当该音频数据可用时)。访问单元通常包括一个输出时间实例的所有NAL单元,例如,一个时间实例的所有音频和视频数据。例如,如果每一个视图均具有每秒20帧的帧速率(fps),则每一个时间实例可以对应于0.05秒的时间间隔。在该时间间隔期间,可以同时渲染相同访问单元(相同的时间实例)的所有视图的特定帧。在一个例子中,访问单元可以包括一个时间实例中的编码的图像,其可以被呈现成基本编码图片。

[0082] 因此,访问单元可以包括共同时间实例的所有音频帧和视频帧,例如,与时间X相对应的所有视图。此外,本公开内容还将特定视图的编码图像称为“视图组分”。也就是说,视图组分可以包括在特定的时间针对特定的视图的编码图像(或者帧)。因此,可以将访问单元规定成包括共同时间实例的所有视图组分。访问单元的解码顺序并不必需与输出或者显示顺序相同。

[0083] 媒体呈现可以包括媒体呈现描述(MPD),后者可以包含不同的替代表示的描述(例如,具有不同质量的视频服务),该描述可以包括例如编解码器信息、简档值和级别值。MPD是清单文件(例如,清单文件66)的一个例子。客户端设备40可以获取媒体呈现的MPD,以确定如何访问各个呈现的电影片段。电影片段可以位于视频文件的电影片段盒(moof盒)中。

[0084] 清单文件66(例如,其可以包括MPD)可以通告表示68的分段的可用性。也就是说,MPD可以包括:用于指示表示68中的一个的第一分段变得可用时的挂钟时间的信息,以及用于指示表示68内的分段的持续时间的信息。用此方式,客户端设备40的获取单元52可以基于每个分段的起始时间以及特定分段之前的分段的持续时间,来确定每个分段可用的时间。

[0085] 在封装单元30基于接收的数据而将NAL单元和/或访问单元组合到视频文件中之后,封装单元30将该视频文件传送到输出接口32以进行输出。在一些例子中,封装单元30可以本地存储该视频文件,或者经由输出接口32向远程服务器发送该视频文件,而不是直接将该视频文件发送给客户端设备40。例如,输出接口32可以包括发射机、收发机、用于将数据写入到诸如光驱动器、磁介质驱动器(例如,软盘驱动器)之类的计算机可读介质、通用串

行总线(USB)端口、网络接口或其它输出接口的设备。输出接口32向诸如传输信号、磁介质、光介质、存储器、闪存驱动器或其它计算机可读介质之类的计算机可读介质输出视频文件。

[0086] 网络接口54可以经由网络74来接收NAL单元或访问单元,经由获取单元52将NAL单元或访问单元提供给解封装单元50。解封装单元50可以将视频文件的元素解封装成组成PES流,对PES流进行解分组以获取经编码的数据,根据该经编码的数据是音频流的一部分还是视频流的一部分(例如,如该流的PES分组报头所指示的),将该经编码的数据发送给音频解码器46或者视频解码器48。音频解码器46对编码的音频数据进行解码,并且向音频输出42发送该解码的音频数据,而视频解码器48对编码的视频数据进行解码,并且向视频输出44发送该解码的视频数据(其可以包括一个流的多个视图)。

[0087] 系统10的各个元素(例如,客户端设备40、内容准备设备20和/或服务器设备60)可以被配置为执行本公开内容的各种技术。通常,系统10的例子包括可以被配置为执行以下功能的各种元素:根据ISO基本媒体文件格式(ISO BMFF)或者分段的ISO BMFF文件,来开始DASH媒体呈现的消费,其中,通过分层编码传输(LCT)对象传送协议(例如,ROUTE协议)来传送分段,而无需MPD(例如,清单文件66)且无需在传输对象标识符(TOI)和URL(例如,FDT、EFDT或GFD或ROUTE中的实体报头)之间传送相关联的元数据,并通过其它方式来提供MPD信令(例如,ROUTE中的LCT会话实例描述(LSID)、FLUTE中的SDP、LCT报头、ISO BMFF IS等等。也就是说,客户端设备40可以通过执行服务启动(其可以包括访问服务入口点)来开始消费。

[0088] 此外,系统10的元素(例如,客户端设备40)还可以(另外地或替代地)被配置为在上述的传送假定下完全地消费单向流(即,广播分发)。

[0089] 此外,系统10的元素还可以(另外地或替代地)被配置为在无需MPD(例如,清单文件66)的情况下开始DASH媒体呈现,但在初始启动和播放之后,可以对MPD进行接收和处理(例如,由服务器设备60进行传送,并由客户端设备40进行接收)以便获得更丰富的信息并与宽带/单播传送相组合。

[0090] MPD可以只包含用于广播调谐或频道更改的绝对必要信息,其中下面中的一项或多项可以适用:

[0091] ●当根本不需要MPD信息时,MPD可以是空的。

[0092] ●在一些随机访问点(RAP)处稀疏地包含常规MPD副本(其中一些信息不是绝对必要的,例如,仅仅是一些增强或优化需要),而在两个常规的MPD副本之间,在每个RAP处包含一个轻量级MPD(只具有绝对必要的信息)。

[0093] 此外,系统10的元素还可以(另外地或替代地)被配置为:使用被添加到ROUTE分组的每个分组的目标传输时间,其反映了接收器消费(解码和渲染)数据相对于相同ROUTE会话中的其它数据的时间,其中下面中的一项或多项适用于目标传输时间:

[0094] ●在LCT报头中示意该时间。

[0095] ●在CC报头中示意该时间。

[0096] ●在扩展报头中示意该时间。

[0097] ●其中,该时间是相对时间。

[0098] ●其中,该时间是诸如网络时间协议(NTP)时间之类的绝对挂钟时间。

[0099] ●其中,该时间是相对时间,在分组中示意发布时间,并且仅在发布时间小于或等

于目标传输时间时,才发布该分组以用于消费。

[0100] 此外,系统10的元素还可以(另外地或替代地)被配置为:使用被添加到ROUTE分组的标志,其反映接收器是否在消费(解码和渲染)该分组中包含的数据(当其存在时),其中,下面中的一项或多项可以适用:

[0101] ●当该标志等于1时,该分组由发送方保存,而不会被传输给接收器以用于消费。

[0102] ●当该标志等于0时,发送器发布该分组以便接收器进行消费。

[0103] ●针对目标的最后分组,需要该标志的值等于1。

[0104] 因此,系统10可以实现一种设计,其出于带宽效率、初始启动延迟、简单性、稳健性、可扩展性和复杂性原因是一种有利的方案,而没有任何显著的缺点。

[0105] 用此方式,并且如下面所更详细地解释的,服务器设备60表示一种用于发送媒体数据的设备的例子,其包括:用于输出多个分层编码传输(LCT)会话的数据的网络接口;以及处理器,其被配置为构造LCT会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括有基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示这些LCT会话和所述表示之间的对应关系,经由网络接口输出LSID,并且经由网络接口在相对应的LCT会话中输出所述表示的数据。

[0106] 同样,客户端设备40表示一种用于接收媒体数据的客户端设备的例子,其包括:被配置为对媒体数据进行解码的一个或多个媒体解码器;被配置为接收分层编码传输(LCT)会话实例描述(LSID)的网络接口,其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据、以及这些LCT会话中的一个或多个的数据;以及处理器,其配置为使用LSID来开始消费该DASH媒体呈现的所述表示中的一个或多个,且不使用用于DASH媒体呈现的清单文件,其中,为了开始消费,所述处理器被配置为:经由网络接口来接收所述LCT会话的分组,其中,这些LCT会话的分组包括所述表示中的一个或多个表示的数据的部分;将这些分组的数据提供给一个或多个媒体解码器。

[0107] 图2是示出示例性多媒体内容102的元素的示意图。多媒体内容102可以对应于多媒体内容64(图1)、或者存储介质62中所存储的另一多媒体内容。在图2的例子中,多媒体内容102包括媒体呈现描述(MPD) 104和多个表示110A-110N。表示110A包括可选的报头数据112和分段114A-114N(分段114),而表示110N包括可选的报头数据122和分段124A-124N(分段124)。为方便起见,使用字母N来指代表示110A、110N的每一个中的最后的电影片段。在一些例子中,在表示110A、110N之间可能存在不同数量的电影片段。

[0108] MPD 104可以包括与表示110A-110N不同的数据结构。MPD 104可以对应于图1的清单文件66。同样,表示110A-110N可以对应于图1的表示68。通常,MPD 104可以包括通常用于描述表示110A-110N的特性的数据,例如,编码和渲染特性、适配集、MPD 104所对应的简档、文本类型信息、相机角度信息、评级信息、技巧模式信息(例如,用于指示包括时间子序列的表示的信息)、和/或用于获取远程时段的信息(例如,在播放期间插入到媒体内容中的定向广告)。

[0109] 当存在报头数据112时,其可以描述分段114的特性,例如,随机访问点(RAP,其还被称为流访问点(SAP))的时间位置,其中,分段114的特性包括随机访问点、与随机访问点在分段114内的字节偏移、分段114的统一资源定位符(URL)、或者分段114的其它方面。当存

在报头数据122时,其可以描述分段124的特性。另外地或替代地,可以将这些特性全部都包括在MPD 104内。

[0110] 分段114、124包括一个或多个编码的视频样本,其每一个都可以包括视频数据的帧或者片。分段114的每一个编码的视频样本可以具有类似的特性,例如,高度、宽度和带宽要求。这些特性可以通过MPD 104的数据来描述,但在图2的例子中未示出该数据。MPD 104可以包括如3GPP规范所描述的特性,另外加上本公开内容中所描述的示意信息中的任何或者全部信息。

[0111] 分段114、124中的每一个可以与唯一的统一资源定位符(URL)相关联。因此,分段114、124中的每一个可以使用诸如DASH之类的流式传输网络协议来独立地获取。用此方式,诸如客户端设备40之类的目标设备可以使用HTTP GET请求来获取分段114或者124。在一些例子中,客户端设备40可以使用HTTP部分GET请求,来获取分段114或124的特定字节范围。

[0112] 图3是示出一种示例性视频文件150的元素的框图,其可以对应于表示的一个分段(例如,图2的分段114、124中的一个)。分段114、124中的每一个可以包括:基本遵循图3的例子中所示出的数据布置的数据。视频文件150可以说是对分段进行封装。如上所述,根据ISO基本媒体文件格式和其扩展的视频文件将数据存储在一系列对象(其被称为“盒”)中。在图3的例子中,视频文件150包括文件类型(FTYP)盒152、电影(MOOV)盒154、分段索引(sidx)盒162、电影片段(MOOF)盒164和电影片段随机访问(MFRA)盒166。虽然图3表示视频文件的例子,但应当理解的是,根据ISO基本媒体文件格式及其扩展,其它媒体文件可以包括结构上类似于视频文件150的数据的其它类型的媒体数据(例如,音频数据、定时文本数据等等)。

[0113] 文件类型(FTYP)盒152通常描述用于视频文件150的文件类型。文件类型盒152可以包括:用于标识对视频文件150的最佳使用进行描述的规范的数据。可以替代地将文件类型盒152放置在MOOV盒154、电影片段盒164和/或MFRA盒166之前。

[0114] 在一些例子中,诸如视频文件150之类的分段可以包括在FTYP盒152之前的MPD更新盒(未示出)。MPD更新盒可以包括:用于指示与包括视频文件150的表示相对应的MPD将被更新的信息,以及用于更新MPD的信息。例如,该MPD更新盒可以提供:用于更新MPD的资源的URI或URL。再举一个例子,MPD更新盒可以包括用于对该MPD进行更新的数据。在一些例子中,MPD更新盒可以紧跟在视频文件150的分段类型(STYP)盒(未示出)之后,其中STYP盒可以规定用于视频文件150的分段类型。图7(将在下面所进一步详细讨论)提供了关于MPD更新盒的附加信息。

[0115] 在图3的例子中,MOOV盒154包括电影报头(MVHD)盒156、轨道(TRAK)盒158和一个或多个电影扩展(MVEX)盒160。通常,MVHD盒156可以描述视频文件150的一般特性。例如,MVHD盒156可以包括:用于描述原始生成视频文件150的时间、上一次修改视频文件150的时间、用于视频文件150的时间尺度、用于视频文件150的播放的持续时间的数据、或者通常描述视频文件150的其它数据。

[0116] TRAX盒158可以包括针对视频文件150的轨道的数据。TRAX盒158可以包括用于描述与TRAX盒158相对应的轨道的特性的轨道头(TKHD)盒。在一些例子中,TRAX盒158可以包括编码的视频图像,而在其它例子中,可以将轨道的编码视频图像包括在电影片段164中,其中其可以由TRAK盒158和/或sidx盒162的数据来引用。

[0117] 在一些例子中,视频文件150可以包括一个以上的轨道。因此,MOOV盒154可以包

括:等于视频文件150中的轨道的数量的TRAK盒的数量。TRAK盒158可以描述视频文件150的相应轨道的特性。例如,TRAK盒158可以描述用于相应轨道的时间和/或空间信息。当封装单元30(图2)包括视频文件(例如,视频文件150)中的参数集轨道时,类似于MOOV盒154的TRAK盒158的TRAK盒可以描述参数集轨道的特性。封装单元30可以在用于描述参数集轨道的TRAK盒内的参数集轨道中示意序列级别SEI消息的存在性。

[0118] MVEX盒160可以描述相应的电影片段164的特性,例如,以便示意除了MOOV盒154(如果有的话)内包括的视频数据之外,视频文件150包括电影片段164。在流式传输视频数据的背景下,可以将编码的视频图像包括在电影片段164而不是MOOV盒154中。因此,可以将所有编码的视频样本包括在电影片段164而不是MOOV盒154中。

[0119] MOOV盒154可以包括:等于视频文件150中的电影片段164的数量的MVEX盒160的数量。MVEX盒160中的每一个可以描述电影片段164中的相应一个的特性。例如,每一个MVEX盒可以包括用于描述电影片段164中的相应一个的持续时间的电影扩展报头盒(MEHD)盒。

[0120] 如上所述,封装单元30可以将序列数据集包括在不包含实际编码的视频数据的视频样本中。通常,视频样本可以对应于访问单元,其中访问单元是处于特定的时间实例的编码图像的表示。在AVC的背景下,编码图像包括一个或多个VCL NAL单元以及其它相关联的非VCL NAL单元(例如,SEI消息),其中这些VCL NAL单元包含用于构造该访问单元的所有像素的信息。因此,封装单元30可以将序列数据集(其可以包括序列级别SEI消息)包括在电影片段164中的一个里。封装单元30还可以将序列数据集和/或序列级别SEI消息的存在性示意为存在于与电影片段164中的一个相对应的MVEX盒160中的一个内的电影片段164中的所述一个中。

[0121] SIDX盒162是视频文件150的可选元素。也就是说,遵循3GPP文件格式或者其它这种文件格式的视频文件并不需要包括SIDX盒162。根据3GPP文件格式的例子,可以使用SIDX盒来识别分段的子分段(例如,视频文件150内包含的分段)。3GPP文件格式将子分段规定成“具有相应的媒体数据盒的一个或多个连续电影片段盒的自包含集合,包含有电影片段盒所引用的数据的媒体数据盒必须在该电影片段盒之后,并在包含关于相同轨道的信息的下一个电影片段盒之前”。此外,3GPP文件格式还指示SIDX盒“包含针对该盒所记录的(子)分段的子分段的引用序列。所引用的子分段在呈现时间上是连续的。类似地,分段索引盒所指代的字节在该分段内始终是连续的。所引用的大小给出所引用的材料中的字节数量的计数”。

[0122] SIDX盒162通常提供用于表示视频文件150中包括的分段的一个或多个子分段的信息。例如,该信息可以包括:子分段开始和/或结束时的播放时间、用于这些子分段的字节偏移、这些子分段是否包括流访问点(SAP)(例如,开始于此)、用于该SAP的类型(例如,该SAP是否是瞬时解码器刷新(IDR)图像、干净随机访问(CRA)图像、断开链接访问(BLA)图像等等)、SAP在子分段中的位置(在播放时间和/或字节偏移方面)等等。

[0123] 电影片段164可以包括一个或多个编码的视频图像。在一些例子中,电影片段164可以包括一个或多个图片组(GOP),每一个GOP可以包括多个编码的视频图像(例如,帧或者图片)。此外,如上所述,在一些例子中,电影片段164可以包括序列数据集。电影片段164中的每一个可以包括电影片段报头盒(MFHD,图3中未示出)。MFHD盒可以描述相应的电影片段的特性,例如,用于该电影片段的序列号。可以按照视频文件150中的序列号的顺序包括电

影片段164。

[0124] MFRA盒166可以描述视频文件150的电影片段164内的随机访问点。这可以有助于执行技巧模式,例如,执行寻求视频文件150所封装的分段内的特定时间位置(即,播放时间)。在一些例子中,MFRA盒166通常是可选的,其不需要包括在视频文件中。同样,客户端设备(例如,客户端设备40)并不必需参考MFRA盒166来正确地解码和显示视频文件150的视频数据。MFRA盒166可以包括:等于视频文件150的轨道的数量的轨道片段随机访问(TFRA)盒(未示出)的数量,或者在一些例子中,其可以等于视频文件150的媒体轨道(例如,非提示轨道)的数量。

[0125] 在一些例子中,电影片段164可以包括一个或多个流访问点(SAP),例如,IDR图片。同样,MFRA盒166可以提供SAP的视频文件150内的位置的指示。因此,MFRA盒166可以提供SAP的视频文件150中的位置的指示。因此,可以根据视频文件150的SAP来形成视频文件150的时间子序列。此外,该时间子序列还可以包括其它图片,例如,取决于SAP的P帧和/或B帧。可以在分段内对时间子序列的帧和/或片进行布置,以使得可以适当地对取决于时间子序列的其它帧/片的该子序列的帧/片进行解码。例如,在数据的分层布置中,用于其它数据的预测的数据也可以包括在该时间子序列中。

[0126] 图4是示出在广播传送中经常出现的示例场景的概念图。可以如下描述该场景:

[0127] ●分发多个服务(通常在不同的信道上)。

[0128] ●服务具有广播组分,并且可以具有附加的单播/宽带组分。

[0129] ●通常,这些服务包括不同的组分,例如,通过元数据来描述的视频、音频、子标题、替代音频、替代视频等等。

[0130] ●在广播中,可以在不同的传输会话中发送组分,以使得可以对这些组分进行选择/过滤。

[0131] ●在分组级别上,对组分和服务进行复用。

[0132] 在图4的例子中,存在两个单独的媒体广播站180A、180B。站180A提供引导数据182,其包括LSID 184和相关的信令片段186A-186N(信令片段186)。此外,站180A还提供音频数据188(其包括图4中所示出的音频数据188A和188B)、视频数据190(其包括图4中的视频数据190A和190B)、定时文本数据192、视频增强层数据194(其包括图4中所示出的视频增强层数据194A和194B)、以及替代的音频语言数据196(其包括图4中所示出的替代的音频语言数据196A和196B)。在一个例子中,站180A使用传输会话标识符(TSI) 0经由用户数据报协议(UDP)来提供引导数据182,使用TSI 1经由UDP来提供音频数据188,使用TSI 2经由UDP来提供视频数据190,经由宽带来提供视频增强层数据194,以及经由HTTP来提供替代的音频语言数据196。

[0133] 在该例子中,站180B提供引导数据200,其包括LSID 202和相关的信令片段204A-204N(信令片段204)。此外,站180B还提供音频数据206(其包括图4中所示出的音频数据206A和206B)、视频数据208(其包括图4中的视频数据208A和208B)、以及可访问音频数据210(其包括图4中所示出的可访问音频语言数据210A和210B)。在一个例子中,站180A使用传输会话标识符(TSI) 0经由用户数据报协议(UDP)来提供引导数据182,使用TSI 1经由UDP来提供音频数据188,使用TSI 2经由UDP来提供视频数据190,经由宽带来提供视频增强层数据194,以及经由单播(例如,根据HTTP)来提供替代的音频语言数据196。

[0134] 通常,站180A、180B可以包括类似于图1的内容准备设备20和服务器设备60的设备。具体而言,站180A、180B可以包括彼此分离的设备,它们执行类似于归因于内容准备设备20和/或服务器设备60的功能。另外地或替代地,站180A、180B可以准备媒体数据,并将准备的媒体数据提供给可以经由网络广播或单播或者空中传输来分发媒体数据的内容传送网络(CDN)、本地站分支机构、有线电视公司或者其它内容分发者。

[0135] 虽然图4中未示出,但可以将数据流(即,引导数据182、200、音频数据188、206、视频数据190、208、定时文本数据192、视频增强层数据194、替代的音频语言数据196和可访问音频210)组织在分段中,例如,包括静态配置数据的初始化分段(IS)和包含编码的媒体数据样本的媒体分段序列。这些分段可以包含用于定时的足够的信息,以使得可以对该数据进行解码,并与其它表示进行联合地呈现以实现同步播放。例如,来自音频数据188和视频数据190的分段可以包括同步数据,以使得在播放期间音频和视频被同步。

[0136] 分段遵循根据例如DASH的特定规则。也就是说,分段是数据对象,内容准备设备20或服务器设备60可以独立于其他分段来生成分段。分段按照解码顺序,即,包含在具有较低编号的分段中的所有数据与具有较高分段号的分段相比具有更小(即,更早)的解码时间。此外,可以在某些位置对内容进行拼接。这将导致一个新的时段,对于该新的时段每个表示将接收到一个新的IS。存在着其它情况,其中可以提供新的IS,但时间轴是连续的。这可能是这种情况,例如,如果IS包括新的元数据。

[0137] 图5是示出图1的客户端设备40的获取单元52的示例性组件的框图。在该例子中,获取单元52包括IP多播单元220、静态配置单元222、ROUTE接收单元224、高速缓存226、代理服务器单元228和DASH客户端230。通常,DASH客户端230可以使用单播请求经由网络74(例如,从服务器设备60)获取一个或多个表示的分段。另外地或替代地,获取单元52可以使用IP多播单元220来订阅IP多播,接收包括一个或多个表示的分段的IP多播数据。IP多播单元220可以将接收的分段232传送给ROUTE接收单元224,其中ROUTE接收单元224可以将所接收的分段存储到高速缓存226。随后,DASH客户端230可以从代理服务器单元228请求分段。用此方式,DASH客户端230可以从服务器设备60的IP地址请求分段,或者从包括获取单元52的客户端设备40的本地主机地址请求分段。在任一情况下,DASH客户端230可以使用HTTP Get或者部分Get请求来获取分段。因此,获取单元52通常可以经由单播或者IP多播来接收分段。

[0138] 如果数据经由单播而可用,则使用诸如媒体呈现、时段、适配集、表示和分段之类的DASH概念在清单文件(例如,媒体呈现描述(MPD))中描述该可用数据。在DASH媒体呈现的单向广播传送的基本部署中,建议使用诸如ROUTE、FLUTE或FCAST之类的对象传送协议将MPD和分段作为常规数据对象进行分发。随后,对象接收器可以使用所恢复的对象,以便从本地服务器向如图1和图5中所示出的DASH客户端提供媒体呈现(使用ROUTE作为示例,但这些技术也可适用于FLUTE或FCAST)。对这些分段进行分发,以使得将MPD中的URL分配给对象传送协议中传输的对象,例如,通过文件传送表(FDT)、扩展FDT(EFDT)、或者使用HTTP实体报头作为对象的一部分。

[0139] 发送器(例如,服务器设备60或类似的设备)传送这些对象,以使得在MPD中宣布的相应的分段可用时间之前在客户端设备40处接收这些段。服务器设备60(或者内容准备设备20)对这些分段进行分组化,并相应地对这些分组进行复用。可以通过不同的传输对象标

识符 (TOI) 来分隔对象,并提供针对URL的分配,如上面所提及的。在更复杂的发送设置中,一个表示的所有分段都在专用会话中进行传送,例如,使用专用的用户数据报协议 (UDP) 端口或者由唯一的传输会话标识符 (TSI) 标识的专用分层编码传输 (LCT) 会话。这种分配准许接收器基于DASH客户端230对表示的选择来选择或者忽略整个会话。

[0140] 但是,通过广播提供媒体呈现可能会对发送器造成一些挑战,这是因为发送器可能需要预测传送和处理时间,以便在MPD中正确地设置分段可用时间。如果由于某些原因,分发系统具有比所预期的或多或少的延迟,则DASH客户端230或者其它播放器的性能可能受到不利影响。播放启动可能会太早或者太迟。播放启动太早可能会导致媒体播放的停滞,而播放启动太迟可能会导致与可能的其它方式相比较慢的信道改变及增加的端到端延迟。另一个问题在于:在DASH客户端230使用分段和所包含的媒体之前,DASH客户端230可能需要接收FDT (或者等同的数据结构) 和MPD。对于频繁的随机访问点来说,这可能增加了来自元数据的开销。

[0141] 如上文和下文所进一步分别详细讨论的,图1和图6示出了示例性发送基础设施。具体而言,图6是根据本公开内容的技术,示出提供不同的发送选项的基本发送器架构240的例子的框图。在该例子中,发送器架构240包括多比特率离线媒体编码器244、多比特率实况媒体编码器248、广告 (ad) 插入单元252、DASH离线编写单元250、DASH实况编写单元256、HTTP实况仿真服务器258、HTTP实况服务器260和ROUTE服务器262。在该例子中,多比特率离线媒体编码器244对一组或多组音频/视频 (A/V) 源内容242进行编码,而多比特率实况媒体编码器248对A/V源内容246A和A/V实况内容246B中的任意一个或二者进行编码。A/V源内容246A可以表示预先录制的内容,而A/V实况内容246B可以表示即时捕获的实况内容,例如,实况新闻事件、体育赛事或者其它这样的实况事件。

[0142] 多比特率离线媒体编码器244向DASH离线编写单元250提供编码的媒体内容。DASH离线编写单元250通常可以准备编码的媒体数据,以便例如经由DASH来传输。例如,对于视频数据而言,DASH离线编写单元250可以准备包括一组编码的视频帧的分段,其中这些编码的视频帧可以包括随机访问点 (RAP) 视频帧。Ad插入单元252可以准备用于在由HTTP实况仿真服务器258准备的媒体流的适当点处进行插入的广告内容。同样,DASH离线编写单元250可以向HTTP实况仿真服务器258提供一个或多个媒体表示的准备好的分段。

[0143] 通常,HTTP实况仿真服务器258仿真诸如HTTP实况服务器260之类的实况媒体服务器。例如,HTTP实况仿真服务器258可以示意各个表示 (例如,音频和视频表示,以及一些例子中的定时文本表示) 的分段的分段可用时间。HTTP实况仿真服务器258可以向ROUTE服务器262提供这些分段,ROUTE服务器262可以基于提示和定时信息254使用IP多播经由ROUTE来在特定的时间输出这些分段。另外地或替代地,ROUTE服务器262可以从HTTP实况服务器260接收实况DASH内容,其中HTTP实况服务器260从DASH实况编写单元256接收DASH内容,其中DASH实况编写单元256可以根据从多比特率实况媒体编码器248接收的编码的媒体数据来准备实况DASH内容。虽然在该例子中示出了单个多比特率实况媒体编码器248,但应当理解的是,在一些例子中,可以采用多个编码器对实况媒体数据进行编码。

[0144] 发送器架构240可以将点播内容中的一个或两者提供成伪实况服务和/或使用DASH生成的实况服务。此外,发送器架构240可以包括用于支持广告 (ad) 插入的工具,例如,ad插入单元252。另一选项是向DASH增加一些稳健性,例如,通过发送多个时段,以便准许使

用MPD级别和/或分段级别进行重新同步。可以通过HTTP服务来获得内容,以使得经由单播(例如,通过网络74)连接的客户端可以通过单播来访问该服务的全部或者一部分。一种示例性用例是通过文件传送协议的分发。也就是说,发送器架构240可以通过文件传送协议(例如,FLUTE、FCAST、MPEG媒体传输(MMT)通用文件传送(GFD)或者ROUTE)来传送所生成的DASH对象(MPD、分段或者其它数据对象)。在实况流式传输的情况下,ROUTE的使用可以是优先的,这是由于ROUTE支持实时能力,并将对象适当地映射到通过IP多播以及然后潜在地通过广播物理层(例如,具有通用流封装(GSE)的DVB-T2或者ATSC 3.0规定的技术)传送的分组。

[0145] 通常,对于多媒体广播多播服务(MBMS)、增强型MBMS(eMBMS)和ATSC 3.0来说,考虑服务条目的自上而下的方法:

[0146] ●可以存在服务入口点,例如,包含MPD的HTML-5页面或者MPD本身。

[0147] ●可以通过包含多个媒体组分的MPD来描述流式传输服务,每个媒体组分记录在一个或多个适配集中以进行适当选择,包括元数据、评级、可访问性等。

[0148] ●每个适配集可以包含一个或多个表示。

[0149] ●某些表示可以通过广播获得,并建立一个基本服务。

[0150] ●其它表示可以通过单播获得,以便丰富和增强呈现。

[0151] ●为了引导服务,MPD是必需的,且MPD控制时间和播放。

[0152] 以下被认为是用于随机访问的数据:

[0153] ●对于如MBMS中所规定的基于FLUTE的DASH:

[0154] ○假设某些元数据片段,即用户服务描述绑定(USDB)和会话描述协议(SDP),是静态的并且在缓存中。

[0155] ○需要实时获取的数据如下:FDT实例、MPD、IS和以封闭式图片组(GOP)随机访问点(RAP)开始的媒体分段。封闭

[0156] 式GOP RAP要求是对当今部署的限制。必要数据的集合称为一个组。

[0157] ○FDT的第一分组是组的第一分组,即,该分组需要被接收,然后是与该组相关联的后面剩余的分组。

[0158] ○为了潜在地将数据传送到DASH客户端以进行播放,需要接收至少一个完整的分段,即,仅在完全接收之后,该分段才“可用”并且可以被调度以供DASH客户端进行请求。然后,DASH客户端将通过生成适当的缓冲或者通过遵循MPD中建议的呈现延迟的指示,来进行和调度分段的播放。

[0159] ○应当注意,该过程等同于基于分段的基于ROUTE的DASH。

[0160] 图7是示出在基于FLUTE的DASH的例子中,服务条目的不同方面的例子的概念图。具体而言,图7示出了一种示例性服务条目280,其中,互联网协议(IP)地址和端口以及MPD URL引用可以用于获取数据。具体而言,服务条目280示出:诸如客户端设备40之类的客户端设备可以使用文件传送表(FDT) 288来获取MPD 284,随后使用FDT 288和MPD 284来获取初始化分段282和媒体分段286。

[0161] 此外,图7还示出了这些各个元素之间的处理依存关系290。具体而言,FDT 288是独立的,MPD 284取决于FDT 288,初始化分段282取决于FDT 288和MPD 284,媒体分段286取决于FDT 288、MPD 284和初始化分段282中的每一个。

[0162] 此外,图7还示出了从左到右的典型发送顺序292。由于处理依存关系290,根据典型的发送顺序292,在MPD 284之前发送FDT 288,其中在初始化分段282之前发送MPD 284,在一个或多个媒体分段分组286A、286B(其形成媒体分段286的全部或者一部分)之前发送初始化分段282。

[0163] ●对于基于ROUTE的DASH而言:

[0164] ○在该情况下,假设LCT会话实例描述(LSID)在接收器设备的高速缓存中可用,即,接收器具有关于调度的会话的信息以及分配给每个LCT传输会话的信息。

[0165] ○要实时获取的数据可以包括以下内容:EFDT实例、MPD、IS和以封闭式GOP RAP开始的媒体分段。封闭式GOP RAP要求是对当今部署的限制。必要数据的集合被称为组。在本公开内容的讨论中,提出了去除EFDT和MPD。

[0166] ○EFDT是组的第一分组,即,该分组需要被接收,然后是与该组相关联的后面剩余的分组。

[0167] ○在一种操作模式下,如上所讨论的基于分段的接收模式将是可行的。然而,由于ROUTE支持基于媒体传送事件(MDE)的传送,因此当接收到足够大的媒体分段前缀时,DASH客户端和/或ISO BMFF处理引擎可以更早地发起播放。在这种情况下,为了潜在地将数据传送到DASH客户端进行播放,MDE需要是可用的并且可以被调度以供DASH客户端进行请求。随后,DASH客户端将进行并调度媒体分段的前缀的播放,以加快初始延迟。相关的并且需要考虑的是适当地得到解码时间(通常相同)和可用性的方法。

[0168] 图8是示出在基于ROUTE的DASH数据的传送的例子中,服务条目300的另一组示例性方面的概念图。在该例子中,扩展FDT(EFDT) 302包含TOI到URL的映射、以及针对MPD 306、初始化分段304和媒体分段310(其由媒体分段媒体传送事件(MDE) 308A和媒体分段剩余部分308B来形成)的内容类型。从处理的角度来看(即,如处理依存关系312所示出的),具有媒体数据的媒体分段310的访问取决于初始化分段304、MPD306和EFDT 302。初始化分段304取决于处理MPD 306和EFDT 302。MPD306的处理取决于EFDT 302的处理。为了实现这种处理,分组的发送顺序314可以如图7中所示,即,EFDT 302、接着MPD 306、接着初始化分段304、以及接着媒体分段分组(例如,媒体分段MDE 308A,跟着的是媒体分段剩余部分308B)。

[0169] 针对ROUTE的具体接收已记录如下:

[0170] ●LSID描述了IP/端口组合和不同的LCT会话。

[0171] ●将分段的ISO BMFF文件中表示的不同媒体组分关联到并唯一地分配给IP/端口/TSI的组合。

[0172] ●通过MPD URL来描述服务。

[0173] ●ROUTE接收器从信道获取针对组分/服务的对象。

[0174] ●在这些对象的接收中,ROUTE接收器使用FDT/EFDT(或者可能的实体模式)向每个对象分配URL。

[0175] ●DASH客户端获得MPD URL,从本地高速缓存提取,并基于MPD中的URL开始消费这些分段。

[0176] ●定时由MPD和DASH客户端进行控制:这些分段不具有分配的定时。

[0177] 在替代的方法中,以下可以适用:

[0178] ●下层信令提供足够的信息,以便在没有MPD的情况下启动服务。

[0179] ●如果增加了单播,或者如果需要更丰富的选择,并且查询了MPD,则MPD仍然可以是统一的MPD,但被视为是在启动时不需要的。甚至可以对MPD进行配置,以使得仅仅用于广播,不使用MPD,并且也不需要URL绑定。用此方式,这些技术可以避免使用FDT、EFD T或者用于将URL绑定至对象的其它方式。

[0180] 本公开内容描述了类似于该方法的示例性设计,其可能提供带宽效率、初始启动、简单性、稳健性、可扩展性和复杂性原因的优势,而没有显著的缺点。为了保持与现有服务的兼容性并应对不同的用例,可以考虑下面的方法:

[0181] ●服务信令提供可以采用以下哪些方法的指示:

[0182] 1、即使启动,接收器也需要MPD,即,在没有MPD的情况下,则不能启动DASH客户端,不能进行服务启动。这将基本上重复当前的方法。

[0183] 2、服务提供商提供足够的信令,以使得在没有MPD的情况下启动是可能的和允许的,但也提供MPD并且描述更丰富的内容提供和替代方案。

[0184] 3、服务提供商根本不提供MPD,服务由较低层信令完全描述,没有实现丰富的信息。这将阻止混合广播/宽带服务。

[0185] ●用于不同情况的示例性服务:

[0186] 1、使用第一种方法,在MPD中具有详细信息的加密服务。

[0187] 2、使用第二种方法,可以丰富服务的单播组分的免费A/V服务。

[0188] 3、使用第三种方法,不需要MPD的简单的免费播放的A/V服务。

[0189] 以分配MPD作为入口点的自顶向下方法,可能发生以下问题:

[0190] ●DASH媒体格式通过广播分发来分发。

[0191] ●为了启动该会话,需要接收若干数据对象(较低层信令、会话描述、FDT/EFD T/等同方法、MPD、IS和媒体分段的至少一部分),以便随机地访问数据并安排播放。

[0192] ●时间控制和媒体的选择处于MPD中,所以需要在启动之前接收MPD,并接收元数据以识别MPD。

[0193] ●另一个问题在于:需要在发送器处以可以预测接收器的时序的方式来产生MPD(如果未修改的话)。

[0194] ●另一个问题在于:需要在每个组分中的每个随机访问点发送MPD,以便快速地访问视频或音频。

[0195] ●另一个问题在于:需要FDT或EFD T以便能够映射URL。

[0196] ●最后,MPD和元数据通常根据XML进行格式化,统一的MPD描述整个服务,其包括所有组分以及所有的单播/宽带传送的数据。这可能使MPD不必要地大,但在开始时可能只需要广播信息。

[0197] 本公开内容的技术可以克服上面所讨论的任何问题或者全部问题。

[0198] 例如,接收设备(例如,客户端设备40)可以被配置为在单向/广播模式下进行操作(也就是说,对媒体数据进行接收、处理和解码)而不需要EFD T、FDT和/或MPD。特别地,客户端设备40可以从这些数据结构接收信息的替代版本,否则这些信息对于通过其它方式的启动和/或永久操作来说是必需的。

[0199] 图9是根据本公开内容的技术,示出可以用于携带数据的根据RFC5651的LCT报头320的字段的概念图。具体而言,LCT报头320包括版本字段322、拥塞控制标志324、协议特定

指示 (PSI) 字段326、传输会话标识符标志 (S) 328、传输对象标识符 (O) 标志330、半字 (H) 标志332、保留 (RES) 字段334、关闭会话 (A) 标志336、关闭对象 (B) 标志338、LCT报头长度 (HDR_LEN) 字段340、代码点字段342、拥塞控制信息344、传输会话标识符346、传输对象标识符348和报头扩展字段350。

[0200] 如下面所讨论的,根据RFC 5651并根据本公开内容的技术来设置以下各项的值:版本字段322、拥塞控制标志324、协议特定指示 (PSI) 字段326、传输会话标识符标志 (S) 328、传输对象标识符 (O) 标志330、半字 (H) 标志332、保留 (RES) 字段334、关闭会话 (A) 标志336、关闭对象 (B) 标志338、LCT报头长度 (HDR_LEN) 字段340、代码点字段342、拥塞控制信息344、传输会话标识符346和传输对象标识符348。可以根据本公开内容的技术来设置报头扩展字段350。

[0201] 为了支持以其它方式在MPD和/或EFD T中提供的信息的传送,以下功能是可用的并且可能被使用。根据本公开内容的技术,可用于携带数据的字段在下文中以斜体来标记:

[0202] ●LSID (如ATSC 3.0的ROUTE规范草案中所规定的)

- *使用LSID 中的内容描述符来示意针对传送会话的组分属性的分配, 以便选择用于被传递到媒体层的传输会话。*
- *代码点分配的使用*

[0203] ●ROUTE/LCT报头字段 (参见RFC 5651和图9)

[0204] ○ *拥塞控制信息 (CCI) 344: 32、64、96 或128 比特。用于携*

带拥塞控制信息。例如，拥塞控制信息可以包括层数量、逻辑信道数量和序列号。该字段用于本规范的目的是不透明的，因此可以任意使用。

- TSI：传输会话标识符（TSI）346：0、16、32 或 48 比特。TSI 唯一地标识来自特定发送器的所有会话之中的会话。TSI 由发送器的 IP 地址进行限定，因此 TSI 和发送器的 IP 地址一起唯一地标识会话。虽然 TSI 结合发送器的 IP 地址始终唯一地标识会话，但 TSI 是否包含在 LCT 报头中取决于什么被用作 TSI 值。如果底层传输是 UDP，则 16 位 UDP 源端口号可以作为用于该会话的 TSI。
- 传输对象标识符（TOI）348：0、16、32、48、64、80、96 或 112 比特。该字段指示该分组属于会话中的哪个对象。例如，发送器可以在同一会话中发送多个文件，对于第一个文件使用 TOI = 0，对于第二个文件使用 TOI = 1 等等。再举一个例子，TOI 可以从若干发送器同时发送的对象的唯一全局标识符，TOI 值可以是应用于对象的哈希函数的输出。
- 代码点 342：8 比特可用于示意所包含对象的不同属性以及与当前分组的关系。被传递给分组有效载荷的不透明标识符被解码以传送关于该分组有效载荷所使用的编解码器的信息。代码点和实际编解码器之间的映射在每个会话的基础上进行定义，并作为会话描述信息的一部分进行带外传送。CP 字段的使用类似于如[RFC3550]中所描述的 RTP 报头中的有效载荷类型（PT）字段。
- LCT 报头中的特定报头标志
 - PSI 326：这些比特的使用（如果有的话）特定于使用 LCT 构建块的每个协议实例化。如果没有定义这些比特的协议实例化特定的使用，则发送器必须将其设置为零，并且接收器必须忽略这些比特。

[0205]

[0206] ■ RES 334：由LCT拥有，所以目前尚未使用。

[0207] ● ROUTE/LCT扩展报头（参见RFC 5651和图9）

- 可以在 LCT 中使用报头扩展 350, 以适应并不是始终使用或者具有可变大小的可选报头字段。报头扩展的使用的示例包括:
 - [0208] 包括:
 - [0209] ■已经存在的报头字段的扩展大小版本。
 - [0210] ■发送器和接收器认证信息。
 - [0211] ■定时信息的传输
 - [0212] ● *ISO BMFF 的初始化分段 (电影报头) (参见 ISO/IEC 14496-12)*
 - [0213] ○用于呈现的元数据存储发生在文件顶层的单个电影盒中。电影报头允许添加与组分的媒体特定方面相关的附加信息, 例如:
 - [0214] ■媒体报头、关于该媒体的整体信息
 - [0215] ■处理机, 声明媒体 (处理机) 类型
 - [0216] ■媒体信息容器
 - [0217] ●在ROUTE、LCT和ISO BMFF之外
 - [0218] ○关于物理层 (FIT) 的信息
 - [0219] ○运行该服务的呈现层/应用中的信息
 - [0220] MPD的相关功能总结如下, 并且下面将更详细地讨论如何根据本公开内容的技术来解决它们。
 - [0221] 可用时间和呈现时间锚定: 示意可用时间, 以使得在报头中示意对象的可用时间或者对象的MDE部分。呈现时间锚定使得当数据可用于ISO BMFF客户端时, 客户端立即开始解码和执行播放。下面将讨论该定时的细节。
 - [0222] 呈现的类型-简档、静态、动态等等: 对于广播来说, 并非必须设置这些参数, 但数据可以遵循某些规则。可以定义某个简档。该类型认为是静态的。下面将提供详细的简档和媒体格式定义。
 - [0223] 带宽和缓冲器描述: 这些参数与广播分发无关, 因为定时是由数据的到达决定的。但是, 为了启动较短的最小缓冲时间, 应考虑这一方面。下面将讨论这个问题。
 - [0224] 时间偏移缓冲器深度: 该参数与广播分发无关, 因为时间偏移缓冲器是由接收信息来决定的。但是, 可以考虑来自发送器的关于在客户端设备 (即, 接收器设备) 中应该存储多长时间的数据的一些指令。
 - [0225] 使用时段的拼接和复位信息: 为此, 信令可能是必需的。时段更改初始化分段并重置时间轴。必须传送该信息, 并相对于先前的数据来正确地调度该时段开始的播放。下面将提供关于如何实现这一点的细节。
 - [0226] 用于选择/拒绝的适配集和表示元数据: 对于基于ROUTE的分发而言, 需要在LCT传输会话级别上进行适配集和表示的选择。一旦选择了传输会话, 就对该数据进行转发以解码和处理到ISO BMFF层。基于初始化分段中的信息, ISO BMFF层仍然可以选择或者拒绝或者过滤一些数据, 但这与这里的讨论无关。为了选择正确的会话, 使用LSID中的信息。下面将更详细地讨论这个问题。
 - [0227] 表示的关系 (可切换、依存关系等等) 和无缝切换信息: 在大多数应用中, 每个适配集只有一个表示在广播中分发。然而, 如果广播多个表示的情况, 则可以如下所述地在LSID

中提供每个表示的各个方面。如果需要在不同的表示之间进行无缝切换,则需要使用MPD,但是只有在执行表示切换时才需要MPD,并且也可以在初始化分段中示意在不同的表示之间进行无缝切换所需的信息。在某些情况下,例如,当使用分层编码时,该信息也在初始化分段中进行提供。在LSID中示意LSID会话的依存关系,下面将描述另外的细节。

[0228] 呈现时间偏移:呈现时间偏移示意在该时段中的表示的第一呈现时间,即它提供一个锚点。必须通过使用ROUTE报头中的信息来复制该信息。下面更详细地讨论这个问题。

[0229] 初始化分段和媒体分段的位置和URL:MPD通过解决下面的问题来描述对象的位置以及与MPD结构的绑定。它说明其是哪个对象(初始化分段、媒体分段、其它类型),并将其放入上下文中,例如,描述媒体分段的序列,并向DASH客户端提供有关如何使用数据对象的信息。MPD指向URL,通过MPD信息与文件之间的绑定,建立媒体流式传输。此外,(E)FDT提供从TOI到URL的映射。该关系和类型指示必须通过ROUTE传送来复制,为此,使用ROUTE报头并且必须严格使用TSI、TOI和ROUTE报头,以便实现该目的。此外,还可以考虑某些发送顺序。可以考虑的另一个方面是IS如何与媒体分段有关。下面讨论这些限制和细节。

[0230] 媒体分段的持续时间:通常,使用媒体分段的持续时间来计算分段可用时间,并用于寻找目的。该持续时间否则是不相关的。所以一般来说,该信息对于广播来说不是必要的。

[0231] 关于内容保护的细节信息:可以在初始化分段中示意内容保护。如果示意复杂的问题,则可以认为MPD是必要的。但一般来说,初始化分段中的信息被认为是足够的。

[0232] 事件流信令:表示可以携带可能必须由DASH客户端进行解析的事件流。如果事件流被认为是相关的,则可以在LSID内容描述符中示意其。或者,可以使用MPD来传送事件流,但是这样的事件流与启动无关。下面将讨论关于在LSID中发送带内事件流的另外的细节。

[0233] 辅助数据-程序信息、度量:MPD可以携带与某些操作相关的信息,如程序信息、度量收集启动或其它方式。该信息不是实时关键的,如果需要的话,可以在以较慢的频率传送或者通过单播提供的MPD中提供该信息。

[0234] 考虑以下方面以建立定时和缓冲模型:

[0235] ●每个分组的目标传输时间(TTT)类似于RFC 2250中针对MPEG-2TS的基于RTP传送的定义。目标传输时间可以是:

[0236] ○仅发送器提示信息,或者

[0237] ○在LCT分组的拥塞控制报头中示意成NTP时间戳(中间32位)或者发送成类似于MPEG-2TS的90kHz时钟,或者发送成具有LSID中定义的时间尺度的任何其它时钟,例如,所包含的媒体的时钟。

[0238] ●如果在接收器处存在并已知,则TTT提供关于确切地何时将数据发布到下一层(即,发布到ISO BMFF层)的信息。

[0239] ●假设ISO BMFF处理机立即启动解码,并一旦从ROUTE级别发布之后,立即渲染数据。

[0240] ●考虑与如何向接收器示意其可以传送数据的时间有关的不同方式:

[0241] ○RAP组的第一分组(通常为IS)和媒体分段的第一分组中的扩展报头,示意该数据需要在ROUTE接收器中保持多长时间,直到被发布到下一级别为止。这个时间称为发布时间(RT)。RT与TTT的时间尺度相同,可以与TTT时间戳(而不是实际时间)进行比较。随着TTT

的增加,RT一定不会减小。如果扩展报头中的RT时间超过当前对象/分段的最大TTT时间,则该分组中包含的数据应该被保持,直到在下一个分段或任何未来的分段中达到TTT时间为止。具体来说,接收器可以如下操作:

[0242] ■如果在RT时间内接收到分组,则可以保持该数据,直到在超过RT的TTT内接收到分组为止。

[0243] ■如果没有在RT时间内接收到分组,则可以将所包含的对象数据与前面的数据一起发布,其中“前面的”是增加start_offset和增加TOI编号的对象的顺序。

[0244] 如果TTT正在使用,并且接收器在TTT与实际接收时间之间观察到显著的抖动,则应该补偿该抖动以避免缓冲器下溢。

[0245] ■例如,设备可以增加一些附加延迟,或者发送基础设施增加一些附加启动延迟。

[0246] ○选项2:在该情况下,使用绝对挂钟时间来调度解码。给对象分配一个挂钟时间,在该时间,将对象移出到ISO BMFF引擎。当在不同用户之间进行同步播放很重要时,这种情况是主要相关的。随着挂钟时间的增加,RT必须增加。具体来说,预期接收器可以如下操作:

[0247] ■如果在RT时间内接收到分组,则可以保持该数据,直到达到在RT中记录的挂钟时间为止。

[0248] ■如果没有在RT时间内接收到分组,则可以将所包含的对象数据与前面的数据一起发布,其中前面的是增加start_offset和增加TOI编号的对象的顺序。

[0249] ○选项3:ROUTE报头中的报头位被设置为10,以便示意所包含的数据尚无法被发布到下一个级别。只有该比特未被设置、设置(即,等于1),才可以传送该数据并向前推送。如果对象的最后一个分组(由LCT报头中设置的B标志来指示)仍然将该比特设置为10,则只有当该对象完成时才能对其进行发布。但是,应当注意,始终可以传送一个完整的对象。或者,对于每个对象的最后一个分组,将该报头比特强制为等于1。

[0250] ●如果TTT正在使用,并且接收器在TTT与实际接收时间之间观察到显著的抖动,则应该补偿该抖动以避免缓冲器下溢。

[0251] ○例如,设备可以增加一些附加延迟,或者发送基础设施增加一些附加启动延迟。

[0252] 考虑三种不同类型的接收:

[0253] 1、基于MDE的接收:在第一情形下,基于ISO BMFF文件的前缀,接收器处的播放尽可能快地发生。

[0254] 2、在第二情形下,只有完整的分段才必须被发布到下一个级别。完整的分段由DASH客户端请求,或者根据RT和TTT或挂钟时钟比较来发布。通常,ROUTE接收器基于RT信息来使该分段可用。DASH客户端基于MPD信息来发送请求。应当对内容进行编写和传送,以使得分段可用性开始时间不早于发布时间。

[0255] 3、挂钟锚定:在第三情形下,将数据的呈现保持一段时间,以便与其它媒体(例如,在DASH系统之外传送的媒体)实现同步。必须根据该数据来发布RAT。

[0256] 但是,所有这三种情形的原理是一样的;正发布的数据单元可能正好是MDE或者完整的分段。

[0257] 图10是示出当对象的前缀可以发布到下一层以进行解码时,用于信令的各个选项的概念图。下面的示例性选项(其对应于下面将进一步详细讨论的各个选项中的一些)将在下文被简短地解释,并在图10中进行说明:

[0258] 1、第一示例性选项360:对于该例子,图10示出了包括RAP分组366和分组368A、368B和374的一系列分组。这些分组中的每一个包含相应的目标传输时间(TTT),它们基本与系统级别的解码时间相匹配。在该例子中,RAP分组366和分组368A、368B中的每一个包含TTT1362,而分组374包括TTT2 370。调度器使用TTT信息来调度传送。此外,RAP分组366包括与TTT信息处于相同域的发布时间(RT)364,RT 364确定何时可以将该组的第一分组(通常,初始化分段(IS))或者任何其它分组发布到下一层以进行立即处理。可以为对象添加一次RT,也可以为对象添加多次。RT可以随着一个对象内的发送时间的增加而增加,以便以更加流式的方式来发布对象。在该例子中,当 $TTT1 < RT \ 364 < TTT2$ 时,发布RAP分组366、分组368A和分组368B。

[0259] 2、第二示例性选项380:对于该例子,图10示出了包括RAP分组384和分组386、388、392的一系列分组。RAP分组384与RT 382相关联。假定在时间NTP1接收到RAP分组384,在NTP2接收到分组386,在NTP3接收到分组388,在NTPX接收到分组392。在该情况下,只为对象增加NTP中的发布时间。这类似于MPD中的分段可用时间,即,准许在该时间之后转发该对象。因此,当 $NTP \geq RT \ 382$ 时,可以在时间390发布RAP分组384。有益的方面是,这不需要另外的定时信令,但发送器需要在NTP中设置该发布时间时考虑传送延迟和抖动,或者存在以下问题:在传送时任何意外的延迟将导致一些问题,如不能及时地呈现对象(其导致启动延迟)或者对象被延迟接收,以至于错过时间轴。该选项有利于使用挂钟同步时的目的。

[0260] 3、第三示例性选项400:对于该例子,图10示出了包括RAP分组404和分组406A、406B、412的一系列分组。在该情形下,RAP分组404包含发布标志(RF)402(或者“保持标志”),其包括用于指示在何时之后可将数据发布到下一层的信息。在该例子中,对于RAP分组404和分组406A、406B, $RF = 0$ 。因此,在NTP时间410之后,可以发布RAP分组404和分组406A、406B,这是由于在NTP时间410处的时间是 $RF = 1 \ 408$ (其与分组412相关联)。这种方法很简单,但是可能会遇到一个问题,因为它可能不允许跨越分段边界的信令。然而,仍然应该考虑选项400,因为它简化了ROUTE接收器的操作。

[0261] 图11是示出用于接收数据和对接收的数据进行缓存的两个示例性模型的概念图。客户端设备40的获取单元52可以包括根据这些例子中的任意一个进行配置的一个或多个组件。在第一示例性模型432中,ROUTE接收器和输出缓冲器426接收分组(例如,RAP分组420和分组422、424)。ROUTE接收器和输出缓冲器发起和调度借助于ISO BMFF解码器430的解码和呈现。ROUTE接收器和输出缓冲器将ISO BMFF流形式的所接收数据传送到ISO BMFF缓冲器428。基于ROUTE接收器和输出缓冲器426所建立的调度,ISO BMFF解码器430从ISO BMFF缓冲器428获取数据,随后对所获取的数据进行解码和呈现。

[0262] 在第二示例性模型452中,ROUTE接收器446接收分组(例如,RAP分组440和分组442、444)。ROUTE接收器将所接收的分组缓冲在ROUTE输出和ISO BMFF缓冲器448中,并发起和调度借助于ISO BMFF解码器450的解码和呈现。ISO BMFF解码器450从ROUTE输出和ISO BMFF缓冲器448获取数据,并对所获取的数据进行解码和呈现。

[0263] 基于上面的描述,以下方面可以是有关的:

[0264] ●提供可用时间,例如,根据上面图10中的三种选项中的一种,例如,选项360、380或400中的一种。这是分段的第一部分被发布到ISO BMFF客户端的时间。

[0265] ●呈现时间是使得基于提供的信息在接收器处立即发起解码和渲染。应当注意,

这只是一个数据模型,并且可以共享ROUTE接收器和ISO BMFF接收器之间的缓冲器。图11中示出了两个示例性模型(432、452)。RAP分组(例如,RAP分组420或者RAP分组440)可以包含用于补偿时段开始和呈现时间偏移的MDE中的信息,即,媒体分段中的最早呈现的开始可以是瞬时的。

[0266] ●可以通过上述三个示例性选项(360、380、400)中的一个来示意缓冲器验证和初始播放延迟。

[0267] ●在下面的操作模式中可以考虑不同的情况(下面提供了不同类型的示例信令)。要考虑的不同情况包括:

[0268] ○常规媒体分段的分组只被传递到ISO BMFF缓冲器,而无需在ROUTE接收器中进行任何调度。发送器保证初始播放延迟提供无缝播放而无缓冲器下溢。这些分组也可以包含用于调度发布的RT,但通常这不是必需的,因为随着ISO BMFF层调度播放,可以立即对其进行发布。

[0269] ○接收器忽略冗余初始化分段,并进行丢弃,而不会传送给ISO BMFF解码器。

[0270] ○增大的初始化分段被提供给ISO BMFF解码器。但是,应当向解码器通知媒体是时间连续的,并且只有一些非基本信息发生改变。然而,如果不存在这样的API,则可以进行重置,并且可以重新调度播放。

[0271] ○内容时段边界:在这种情况下,必须转发IS,并执行完整的新调度。应当注意,这可能会导致具有一些样本的ISO BMFF缓冲器被清空或者在需要由解码器处理的较小时间内没有数据可用的情况。这与常规DASH操作相同。

[0272] ●应当注意,上述操作是可能的,因为发送器必须遵守下面更详细地讨论的发送要求。

[0273] 如果发送器和接收器被配置为使用基于分段的传送/接收,则可以向定时信令增加信息,如下面所讨论的。但是,在根据基于MDE的传送来执行基于分段的接收的情况下,那么问题是棘手的,因为接收器需要了解如何利用该定时信息。应当考虑下面的选项:

[0274] ●使MPD可用,以调度和示意分段可用时间。

[0275] ●针对每个组分向LSID增加一个属性,其指示媒体时间(和TTT)的另外延迟时间,以支持基于分段的接收。通常,这样的时间是最大分段持续时间。应当注意,该时间并非以挂钟时间,而是以传输时间,因此如果完成了基于突发的传送,则分段传送的另外延迟可能比基于MDE的接收之一稍微更多。

[0276] 如果发送器和接收器被配置为使用挂钟锚定(也就是说,基于挂钟的解码),则可以将信息添加到下面所讨论的定时信令信息中。但是,可能存在以下情况:示意基于MDE的传送,但是接收器需要将解码和播放同步到挂钟时间。可以考虑下面的选项:

[0277] ●使MPD可用,以调度和示意分段可用时间以及所建议的呈现延迟。

[0278] ●针对每个组分增加一个属性,其指示TTT或者解码时间与挂钟时间的映射。

[0279] ●创建用于LCT的扩展报头,其提供这种信息,即,与挂钟播放调度的映射。

[0280] 为了实现基于LSID中的信息来选择在LCT传输会话中传送的媒体,除了表示之外,还可以向LSID增加被分配给适配集的很多元素和属性。具体而言,这可以包括以下信息中的任何一个或全部(详见ISO/IEC 23009-1,第5.3.4节(适配集)和第5.3.7节(公共属性)):

[0281] ●使用@id元素的标识符。

- [0282] ●使用@group属性的组关联。
- [0283] ●如@lang属性所描述的语言。
- [0284] ●@contentType属性所描述的媒体组分类类型。
- [0285] ●如@par属性所描述的图像宽高比。
- [0286] ●如角色元素所描述的角色属性。
- [0287] ●如可访问性元素所描述的可访问性属性。
- [0288] ●如视角元素所描述的视角属性。
- [0289] ●如评级元素所描述的评级属性。
- [0290] ●关于分段属性的信息(随机访问等等):参见SISSI核心实验。
- [0291] ●用于适配集的@profiles属性。
- [0292] ●@width和@height提供指定在由@sar属性所确定的网格上视频媒体类型的水平和垂直视觉呈现大小。
- [0293] ●@sar指定视频媒体组分类类型的样本宽高比。
- [0294] ●@framerate:指定视频媒体类型的输出帧速率(或者在交织的情况下,输出字段速率的一半)。
- [0295] ●@audiosamplingRate:音频媒体组分类类型的最大采样速率。
- [0296] ●@mimeType:指定初始化分段(如果存在)和表示中所有连续的媒体分段的级联的MIME类型。
- [0297] ●@codecs:指定表示中存在的编解码器。该编解码器参数还可以包括简档和级别信息(如果适用的话)。
- [0298] ●@scanType:指定视频媒体组分类类型的源材料的扫描类型。
- [0299] ●FramePacking(帧封装):指定视频媒体组分类类型的帧封装布置信息。
- [0300] ●AudioChannelConfiguration(音频信道配置):指定音频媒体组分类类型的音频信道配置。
- [0301] ●ContentProtection(内容保护):指定关于用于相关联的表示的内容保护方案的信息。
- [0302] ●EssentialProperty(基本属性):指定与包含元素有关的信息,选择该组分的媒体呈现作者认为该包含元素是基本的。
- [0303] ●SupplementalProperty(补充属性):指定与可以用于处理该组分的包含元素有关的补充信息。
- [0304] ●InbandEventStream(带内事件流):指定相关联的表示中的带内事件流的存在性。
- [0305] 所有这种信息都可以用于LCT传输会话级别的选择。如果不提供选择,则可以将所有流转发给ISO BMFF处理机以进行处理。可以提供该信息用于早期过滤、以及对传输会话级别的选择。
- [0306] 图12是示出接收器设备460的一组示例性组件的概念图,其中该接收器设备460包括选择和过滤单元462、ROUTE接收器464、其它对象处理器466、ISO BMFF媒体处理器468和ISO BMFF预处理器470。接收器设备460可以对应于客户端设备40,其中获取单元52可以包括选择和过滤单元462以及ROUTE接收器464,而解封装单元54可以对应于ISO BMFF预处理

器470、ISO BMFF媒体处理器468和其它对象处理器466。

[0307] 在该例子中,ROUTE接收器464向选择和过滤单元462提供具有内容描述符的LSID。基于该信息,选择和过滤单元462可以选择应当向处理器(例如,ISO BMFF预处理器470、ISO BMFF媒体处理器468和其它对象处理器466)转发的适当组分。

[0308] 通过应用上面的原理,可以通过新的方案和扩展来确保可扩展性。系统保持与基于DASH的部署兼容。此外,ISO BMFF媒体处理器468还可以在其级别上选择或者拒绝某些媒体流。

[0309] 可以灵活地使用LCT报头,以便示意某些属性。但是,对于ROUTE需要强制使用这些报头。这将在下面进行讨论。

[0310] 在携带媒体流(内容类型xxx/mp4所指示的)的LSID中,表1提供了代码点字段的设置。这允许发送器和接收器在不具有FDT和MPD的情况下进行操作。

[0311] 表1:用于具有内容类型xxx/mp4的媒体流的代码点分配

| 代码点 | 语义 |
|-----------|--|
| 0 | 在 FDT 中描述的对象类型,未片段化的(整个分组包含 IS,无 ROUTE 报头) |
| 1 | 32 比特 ROUTE 报头、FDT 中描述的对象类型 |
| 2 | 新 IS、时间轴不连续,未片段化的(即,整个分组包含 IS,无 ROUTE 报头) |
| 3 | 新 IS、时间轴改变的、片段化的、32 比特 ROUTE 报头 |
| 4 | 新 IS、时间轴不连续,未片段化的(即,整个分组包含 IS,无 ROUTE 报头) |
| 5 | 新 IS、时间轴继续、片段化的、32 比特 ROUTE 报头 |
| 6 | 冗余 IS,未片段化的(即,整个分组包含 IS,无 ROUTE 报头) |
| 7 | 冗余 IS,片段化的,32 比特 ROUTE 报头 |
| 8 | 未片段化的媒体分段 |
| 9 | 片段化的媒体分段 |
| 10 | 未片段化的实体模式中的对象 |
| 11 | 片段化的实体模式中的对象 |
| 12 - 127 | 保留用于静态对象关联 |
| 128 - 255 | 通过 LSID 的动态代码点分配 |

[0313] 在无MPD的ROUTE中,可以使用下面的LCT报头:

[0314] ●TSI示意单个基于ISO BMFF的表示

[0315] ●使用TOI来示意对象。必须以增加的TOI编号(即,对象增加1)来发送一个内容时段内的媒体分段。仅在内容时段处,这可能会发生改变。媒体分段TOI可以只使用TOI的第一比特被设置为1的空间。非媒体分段可以使用第一比特被设置为0的TOI空间。

[0316] ●代码点的使用如第3.6.2节所述。

[0317] ●如下所述地使用PSI比特:

[0318] ○如果第一比特被设置为0,则应用基于时间的发布信令:

[0319] ■如果第二比特被设置为0,则拥塞控制字段是未用的。

[0320] ■如果第二比特被设置为1,则拥塞控制字段包含90kHz时钟频率的时间戳。

[0321] ○如果第一比特被设置为1,则:

- [0322] ■第二比特示意用于对象的发布标志。
- [0323] ●可以使用拥塞控制字段来示意90kHz时钟频率的时间戳。规定下面的示例性LCT扩展报头：
- [0324] ●TTT中的初始发布时间：以TTT时间指定该分组中包含的数据的最早发布时间。大小32比特的扩展报头是适当的。
- [0325] ●NTP中的初始发布时间：以NTP时间指定对象的当前初始字节的确切发布时间。同样，32比特可以是适当的，其提供NTP时间戳的中间32比特。
- [0326] 假设下面选择来自选项1的发送过程。进一步研究针对选项2和选项3的变型，并且可以结合下面所描述的技术或者其它技术来使用。考虑下面的发送行为，而无需使用MPD进行某种媒体呈现：
- [0327] ●可以通过引导来提供LSID片段。
- [0328] ○可以更新LSID片段，以示意服务提供的任何改变。
- [0329] ○LSID片段描述该服务的所有广播组分。
- [0330] ○可以增加内容描述符，以描述每个媒体组分的属性。内容描述符应当提供足够的信息，以使得ROUTE接收器可以选择或者拒绝某些组分。
- [0331] ●通常，仅分发每个组分(适配集)的单个表示。但是，如果针对相同适配集分发不同表示，则LSID可以包含足够的信息来区分这些组分，例如，通过质量排名、空间分辨率、必要的编解码器参数等等。
- [0332] ●(通过TSI标识的)一个LCT会话可以用于每个表示。在分组级别上对LCT会话进行复用。应当在复用的LCT会话之间，一致性地使用TTT。
- [0333] ●每个分组可以包含在90kHz基础上的目标传输时间。该时间表达了目标传送时间。优选地，使用以ISO BMFF的所包含的媒体的解码时间。
- [0334] ●下面的发送过程可以应用于单一LCT会话中的表示：
- [0335] ○可以在LCT会话中发送任何对象。如果该对象不是媒体分段，则TOI中的最高有效位可以是0。如果该对象是媒体分段，则TOI中的最高有效位可以是1。可以通过表1中的静态分配来标识对象，或者可以在LSID中规定代码点。
- [0336] ○假定媒体分段具有特定的发布时间。与该媒体分段相对应的IS可以示意相同的发布时间。应当注意，如果媒体分段紧随其后，则可以只在IS中发送发布时间。
- [0337] ○使用ROUTE协议，将初始化分段作为对象进行发送。对于IS而言，可以仅仅使用以最高有效位为0开始的TOI编号。
- [0338] ■根据表1，以代码点通告初始化分段的类型。
- [0339] ■如果IS适于单个ROUTE分组，则可以使用代码点2、4或6。如果IS不适于单个ROUTE分组，则可以使用代码点3、5或7。
- [0340] ■如果首先发送IS，或者时间轴是不连续的，则示意代码点2或者3。IS必须使用新的TOI。IS的第一分组可以提供用于TTT的随机数。但是，在后一情况下，在持续流中，TTT应该继续表示播放安排。
- [0341] ■如果对IS进行重复以支持随机访问，则可以示意代码点6或7。TOI可以保持不变。TTT可以是连续的，而不减小。
- [0342] ■如果对IS进行更新，但时间轴是连续的，则使用代码点4或5。IS必须使用新的

TOI。TTT可以是连续的，而不减小。

[0343] ■连同IS一起，可以发送扩展报头，其中该扩展报头以与TTT相同的尺度指示向下一层发布IS的最早时间。如果不存在的话，则发布时间是即时的。

[0344] ○使用ROUTE协议，将媒体分段(MS)作为对象进行发送。对于MS而言，可以仅仅使用以最高有效位为1开始的TOI编号。

[0345] ■根据表1，以代码点通告MS的类型。

[0346] ■如果MS适于单个ROUTE分组，则可以使用代码点8。如果MS不适于单个ROUTE分组，则可以使用代码点9。

[0347] ○可能需要一些附加规则。

[0348] 图13是示出用于示例性的一系列分组的示例性发送过程的概念图。在该例子中，第一表示包括初始化分段480和两个媒体分段482、484，在该第一表示之后发送具有新的初始化分段486和后续媒体分段488的第二表示。这些分段中的每一个使用TSI=1。也就是说，使用该LCT会话。在图13中，示出了按顺序对这些分组的发送。

[0349] ●待发送的第一分组(初始化分段分组490A)对应于初始化分段480，其中TTT被设置成MS中的第一样本的解码时间为A。选择TOI用于初始化分段分组490A，并将CP设置为2以指示新的时间轴。将初始化分段480的发布时间设置为比A稍大的值，以指示何时开始呈现。

[0350] ●第二分组(媒体分段分组492A)是媒体分段482的第一分组，且媒体分段482是片段化的。因此，CP=9。使用新的TOI来指示新对象。如针对初始化分段480所使用的来使用相同的TTT。由于数据与初始化分段480一起被发布，因此不发送新的RT。

[0351] ●对于媒体分段482的第二分组(媒体分段分组492B)，设置新的传输时间(假设其包含较晚的解码时间)，其中 $C > A$ 。如果 $B < C$ ，则这将导致将初始化分段480和媒体分段分组492A发布到下一层。

[0352] ●为了实现随机访问，以与媒体分段484的媒体分段分组494A的TTT相匹配的TTT，来冗余地(CP=4)发送初始化分段480。以TTT的尺度来设置发布时间。

[0353] ●随后，以与媒体分段482相同的方式，在两个分组(媒体分段分组494A、494B)中发送媒体分段484。这些时间用于发布这些分组。使用新的TOI，其比媒体分段482的TOI多一个。

[0354] ●在新的时段开始时，将发送新的初始化分段486(以初始化分段分组496的形式)，因此标记为CP=2。新的TOI用于初始化分段486。新的表示的TTT和RT应该是先前TTT的延续，以便指示播放顺序和时序。

[0355] ●随后，可以将媒体分段488作为一个或多个分组进行发送，其包括媒体分段分组498。

[0356] 在替代的例子中，可以使用轻量级MPD。在该例子中，定时和缓冲信令和操作与上述相同，而通过其它方式对非定时MPD信息的信令的上述其它变化未被采用。代之，对MPD进行改变，以使得其只包含不同场景的绝对必要信息(包括用于在广播中进行调谐和信道更改)，而不存在其它信息。在极端情况下，如果完全不需要MPD信息，则其可以是空的。当仅仅该信息的一个子集是必要的并存在时，MPD是轻量级的，因此避免了不必要的开销。发送器可以选择在一些RAP处并在两个常规的MPD复本之间稀疏地放置常规MPD复本，在每一个RAP处放置轻量级MPD。

[0357] 这可以导致下面中的任何一种或全部:

[0358] ●一旦接收到完整的MPD,启动简单的但冗余的MPD。

[0359] ●这些MPD将只包含该组分的信息。

[0360] ●由于由较低层驱动,因此忽略定时。

[0361] ●由于通过对象协议来传送分段,因此不指定用于分段的URL。

[0362] ●基本上,上面所提及的内容标识符将被映射成单独的对象。

[0363] 在该时间点,可以使用遵循ISO BMFF实时简档的分段。在以下中讨论了针对这些应用的改进的媒体格式的优化:Stockhammer等人于2015年2月10提交的标题为“LOW LATENCY VIDEO STREAMING”的美国临时申请62/114,423,该申请的全部内容以应用的方式并入本文中。

[0364] 对于混合服务提供而言,MPD可以是重要的。可以使用用于描述广播和宽带分发内容的统一MPD,据此:

[0365] ●将MPD作为ROUTE会话中的对象进行传送,或者只通过链路经由宽带来提供该MPD。

[0366] ●将EFDT作为ROUTE会话中的对象进行传送,或者只通过链路经由宽带来提供该EFDT。

[0367] ●EFDT记录对象传送中的TOI与该MPD中可见的URL之间的映射。

[0368] ●随后,定时由该MPD进行控制,即,广播分段不再推送到ISO BMFF客户端,但DASH客户端控制定时。但是,这些细节是特定于实现的。

[0369] 图14是示出一种示例性混合DASH客户端模型的概念图。图1的客户端设备40的获取单元52可以根据图14的例子进行配置。在该例子中,混合DASH客户端模型包括ROUTE接收器和输出缓冲器506和DASH客户端512。根据该混合DASH客户端模型,客户端设备可以使用ROUTE接收器和输出缓冲器506经由ROUTE传输来接收分段(例如,RAP分组500和分组502、504),或者使用DASH客户端512经由单播传送从网络514接收分段。在接收到分段之后,ROUTE接收器和输出缓冲器506向ISO BMFF缓冲器508提供这些分段,或者DASH客户端512向ISO BMFF缓冲器508提供这些分段。

[0370] 针对广播的接收类型,存在两种示例类型的实现:

[0371] ●无MPD接收:ROUTE接收器和输出缓冲器506直接将数据转发给ISO BMFF缓冲器508,以便由ISO BMFF解码器510进行解码并播放。

[0372] ●基于MPD的接收:ROUTE接收器和输出缓冲器506生成包括所有信息的MPD,以使得DASH客户端512从ROUTE接收器和输出缓冲器506(例如,其充当代理服务器)获取数据,将该数据存储到ISO BMFF缓冲器508以进行解码和播放。

[0373] 这两个版本都是可能的,并且都是实现选择。此外,可以以如图14中所示出的方式,使用广播和宽带来实现混合客户端。在该情况下,在ROUTE接收器和输出缓冲器506从广播流接收到数据之后,向ISO BMFF缓冲器508提供所接收的数据。一旦增加了宽带,则通过MPD来产生关系。DASH客户端512向TOI分配URL,因此可以基于MPD中的信息来切换到宽带。

[0374] 图15是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的示例性方法的流程图。参照图1的服务器设备60和客户端设备40来解释图15的例子。但是,应当理解的是,其它设备也可以被配置为执行这些技术。例如,图4、图5、图6、图11、图12

或图14的设备可以被配置为执行这些技术。

[0375] 初始时,服务器设备60可以获得媒体呈现的多个表示的数据(550)。例如,服务器设备60可以从内容呈现设备20(图1)接收准备的内容。替代地,服务器设备60可以包括媒体编码器(例如,音频编码器26和视频编码器28)和/或封装单元(例如,封装单元30),以准备表示用于分发。所述多个表示可以与图1的多媒体内容64的表示68相对应。此外,服务器设备60可以获得针对多个表示的清单文件(例如,MPD)(例如,图1的清单文件66)。

[0376] 然后,服务器设备60可以向LCT会话分配这些表示(552)。通常,这些LCT会话中的每一个LCT会话可以携带这些表示中的一个表示的数据,以使得在LCT会话与表示之间存在一对一关系。此外,服务器设备60可以构造针对这些LCT会话的LSID,其指示表示与LCT会话之间的对应关系(554)。例如,LSID可以示意用于LCT会话的TSI与表示标识符之间的关系(例如,这些表示的带宽)。如上所述,服务器设备60还可以构造LSID,来描述用于各个LCT会话的IP和端口组合。

[0377] 此外,服务器设备60还可以构造LSID以包括传统上包含在清单文件中的数据,例如,包括下面中的任何一项或全部的属性:@id、@group、@lang、@contentType、@par(图像宽高比)、角色元素、可访问性元素、视角元素、评级元素、表示的分段的分段属性、@profiles、@width和@height、@sar、@framerate、@audiosamplingRate、@mimeType、@codecs、@scanType、FramePacking、AudioChannelConfiguration、ContentProtection、EssentialProperty、SupplementalProperty和/或InbandEventStream,如上所述。

[0378] 此外,服务器设备60可以根据例如图9的例子来构造LCT会话的分组以包括LCT报头。这些分组可以包括表示的分段的数据。例如,报头信息可以指示相应的表示和分段的TSI和/或TOI。

[0379] 然后,服务器设备60可以经由相对应的LCT会话来发送这些表示的LSID和数据(556)。客户端设备40可以接收针对LCT会话的LSID(558)。虽然在图15的例子中未示出,但服务器设备60还可以定期地发送清单文件,例如,使用表示的某些随机访问点。具体而言,在该例子中,假定客户端设备40在接收到清单文件之前(例如,在清单文件之间)接收到LSID。但是,根据本公开内容的技术,客户端设备40可以使用LSID来访问LCT会话中的一个或多个LCT会话(以及,因此,相应的表示)的媒体数据,而无需接收清单文件(或者在接收清单文件之前)。

[0380] 例如,客户端设备40可以确定LCT会话与表示之间的对应关系。此外,客户端设备40还可以使用在LSID中示意的数据来确定表示的特性。例如,客户端设备40可以确定哪些表示与该客户端设备40的元件(例如,音频解码器46、视频解码器48、音频输出42和视频输出44)所支持的编码和渲染特性相匹配。基于所支持的特性以及这些表示的编码和渲染特性,客户端设备40可以使用LSID来选择表示(560)。例如,客户端设备40可以选择具有客户端设备40的元件所支持的编码和渲染特性的那些表示。

[0381] 此外,客户端设备40还可以从所选定的表示的LCT会话中提取分组(562)。每个分组可以与表示的分段相对应。可以以一个或多个分组的形式来发送表示的每个分段。如上面参照图10的各个示例所讨论的,这些分组可以包括用于指示何时可以发布这些分组的信息。因此,在接收到一个分段的所有分组之后,客户端设备40的获取单元52可以向解封装单元50发送重建的分段,该解封装单元50可以最终对媒体数据进行解封装,并向适当的解码

器(例如,音频解码器46或者视频解码器48)发送媒体数据。用此方式,客户端设备40可以向适当的解码器发送来自这些分组的媒体数据(564)以进行解码,并最终进行呈现。

[0382] 用此方式,图15的示例性方法表示一种方法的示例,其中该方法包括:根据分层编码传输(LCT)会话实例描述(LSID)来确定基于HTTP的动态自适应流(DASH)媒体呈现的多个表示,其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括所述表示中的相应表示的数据;使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的表示中的一个或多个,其中,所述开始消费包括:接收LCT会话的分组,其中这些LCT会话的分组包括所述表示中的所述一个或多个表示的数据中的部分;以及将这些分组的数据提供给媒体解码器。

[0383] 此外,图15的示例性方法还表示一种方法的例子,其中该方法包括:构造分层编码传输(LCT)会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示所述LCT会话与所述表示之间的对应关系;输出该LSID;以及在相对应的LCT会话中输出所述表示的数据。

[0384] 图16是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的另一种示例性方法的流程图。在该例子中,客户端设备40初始时根据分层编码传输(LCT)会话实例描述(LSID)来确定基于HTTP的动态自适应流(DASH)媒体呈现的多个表示,其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括所述表示中的相应表示的数据(580)。然后,客户端设备40使用LSID且不使用用于DASH媒体呈现的清单文件来开始消费该DASH媒体呈现的所述表示中的一个或多个(582)。具体而言,当开始消费时,客户端设备40接收所述LCT会话的分组,其中,这些LCT会话的分组包括所述表示中的所述一个或多个表示的一部分数据(584),并且将这些分组的数据提供给媒体解码器(586)。

[0385] 图17是根据本公开内容的技术,示出用于经由LCT会话来传输媒体呈现的媒体数据的另一种示例性方法的流程图。在该例子中,服务器设备60构造分层编码传输(LCT)会话实例描述(LSID),其中,该LSID包括表示多个LCT会话的信息,这些LCT会话中的每个LCT会话包括基于HTTP的动态自适应流(DASH)媒体呈现的多个表示中的相应表示的数据,其中,该LSID指示所述LCT会话与所述表示之间的对应关系(590)。然后,服务器设备60输出该LSID(592),并且在相对应的LCT会话中输出所述表示的数据(594)。

[0386] 在一个或多个例子中,所描述的功能可以利用硬件、软件、固件或者其任意组合来实现。当利用软件来实现时,可以将这些功能存储在计算机可读介质上,或者作为计算机可读介质上的一个或多个指令或代码进行传输,并由基于硬件的处理单元来执行。计算机可读介质可以包括计算机可读存储介质,该计算机可读存储介质对应于诸如数据存储介质或通信介质之类的有形介质,其中通信介质包括有助于例如根据通信协议来将计算机程序从一个地方传送到另一个地方的任何介质。用此方式,计算机可读介质通常可以对应于:(1)非暂时性的有形计算机可读存储介质;或者(2)诸如信号或载波之类的通信介质。数据存储介质可以是一个或多个计算机或者一个或多个处理器能够进行访问以获取用于实现本公开内容中描述的技术的指令、代码和/或数据结构的任何可用介质。计算机程序产品可以包括计算机可读介质。

[0387] 举例而言,但非进行限制,这种计算机可读存储介质可以包括RAM、ROM、EEPROM、

CD-ROM或者其它光盘存储器、磁盘存储器或其它磁存储设备、闪存或者能够用于存储以计算机能够访问的指令或数据结构的形式期望的程序代码的任何其它介质。此外,可以将任何连接适当地称作计算机可读介质。举例而言,如果指令是使用同轴电缆、光纤光缆、双绞线、数字用户线路(DSL)或者诸如红外线、无线电和微波之类的无线技术从网站、服务器或其它远程源传输的,那么所述同轴电缆、光纤光缆、双绞线、DSL或者诸如红外线、无线电和微波之类的无线技术包括在介质的定义中。但是,应当理解的是,计算机可读存储介质和数据存储介质并不包括连接、载波、信号或者其它暂时性介质,而是针对非暂时性有形存储介质。如本文所使用的,磁盘和光盘包括压缩光盘(CD)、激光光盘、光盘、数字通用光盘(DVD)、软盘和蓝光光盘,其中磁盘通常磁性地复制数据,而光盘则用激光来光学地复制数据。上述的组合也应当包括在计算机可读介质的保护范围之内。

[0388] 指令可以由诸如一个或多个数字信号处理器(DSP)、通用微处理器、专用集成电路(ASIC)、现场可编程门阵列(FPGA)之类的一个或多个处理器或者其它等同的集成或分立逻辑电路来执行。因此,如本文所使用的,术语“处理器”可以指代前述的结构或者适于实现本文所描述的技术的任何其它结构中的任何一种。此外,在一些方面,本文所描述的功能可以提供在被配置用于实现编码和解码的专用硬件和/或软件模块中,或者包含在组合的编解码器中。此外,这些技术可以在一个或多个电路或逻辑元件中完全实现。

[0389] 本公开内容的技术可以使用多种多样的设备或装置来实现,包括使用无线手持装置、集成电路(IC)或者一组IC(例如,芯片集)。在本公开内容中描述了各种组件、模块或单元,以强调被配置用于执行所公开的技术的设备的功能方面,但不一定需要由不同的硬件单元来实现。相反,如上所述,各个单元可以组合在编解码器硬件单元中,或者通过协作的硬件单元集合(其包括如上所述的一个或多个处理器)结合适当的软件和/或固件来提供。

[0390] 描述了各个示例。这些和其它示例落入所附权利要求的保护范围之内。

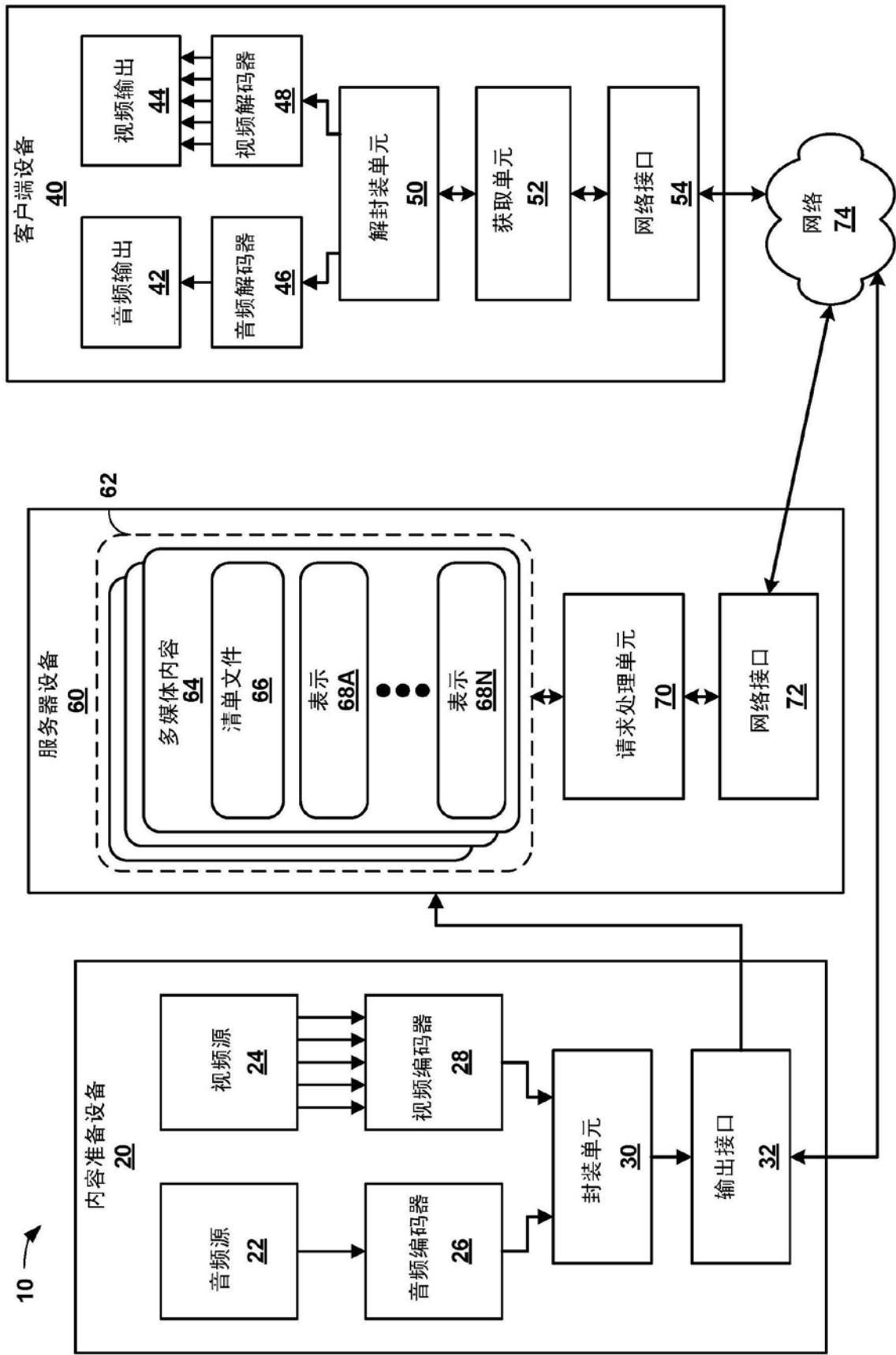


图1

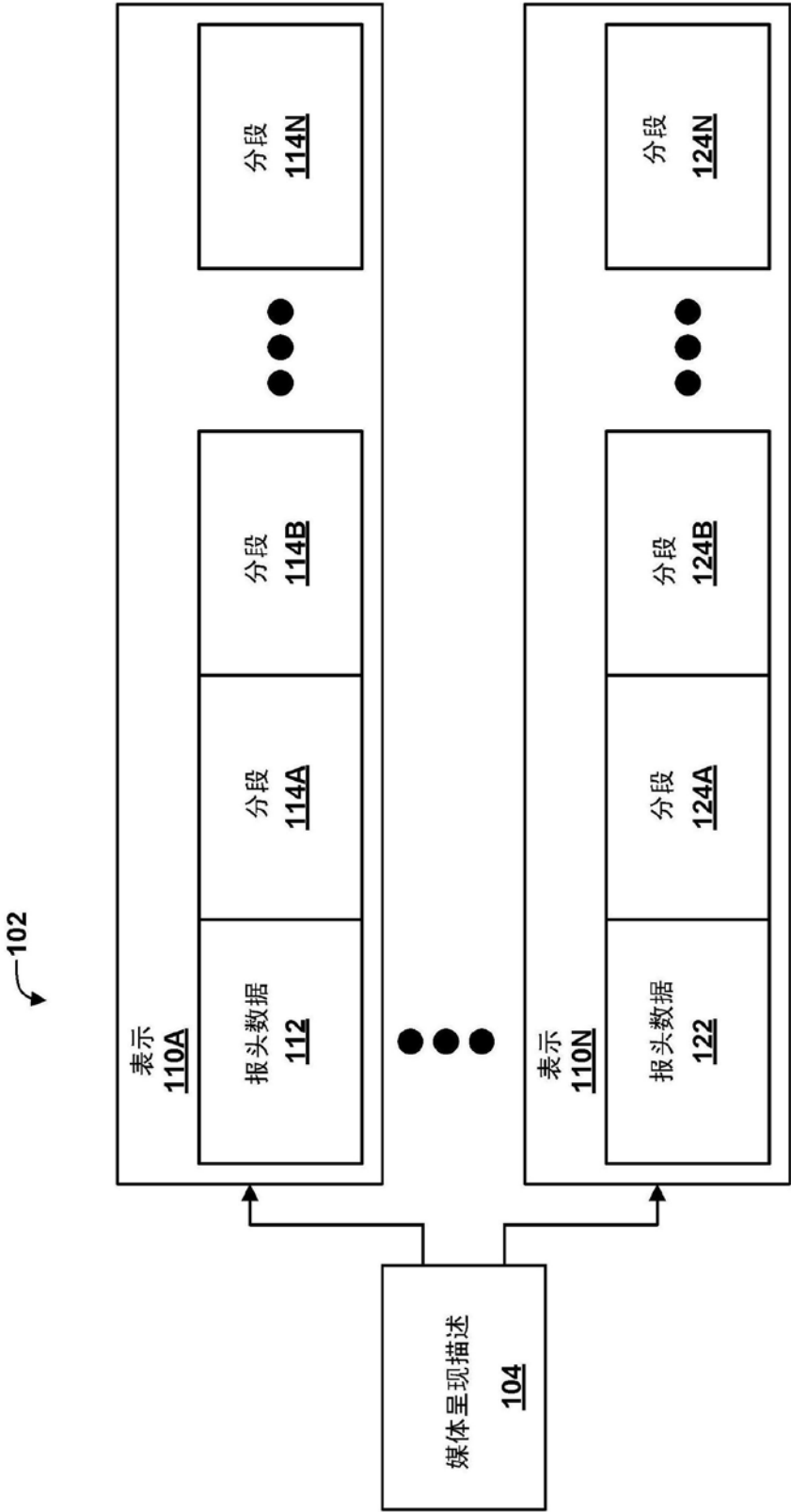


图2

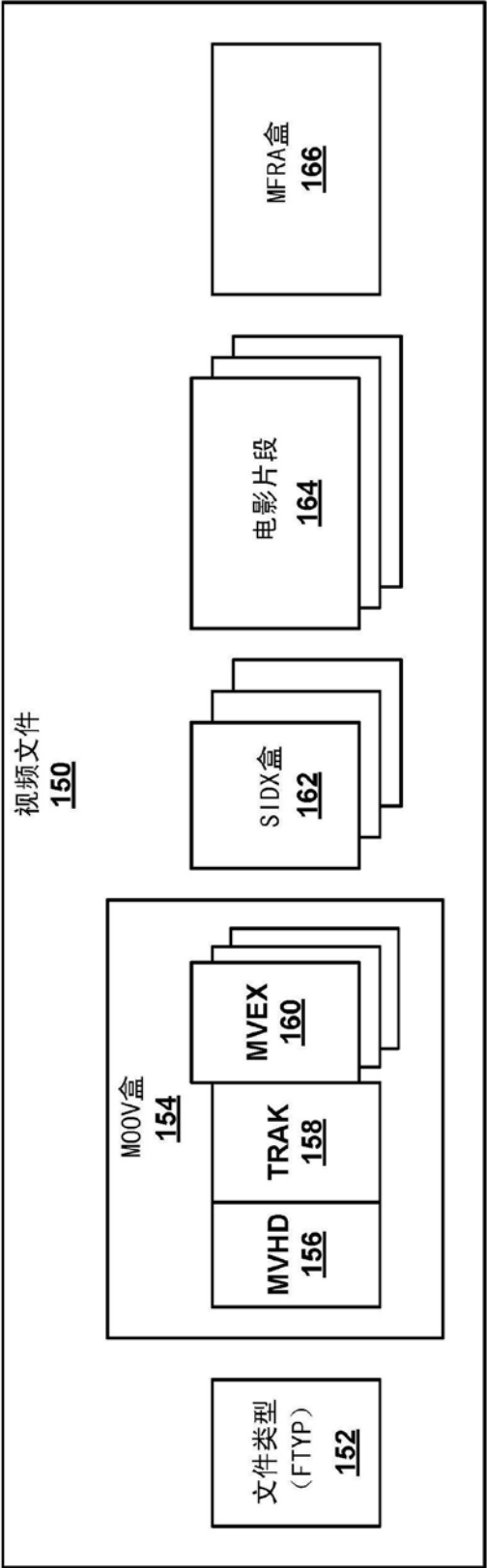


图3

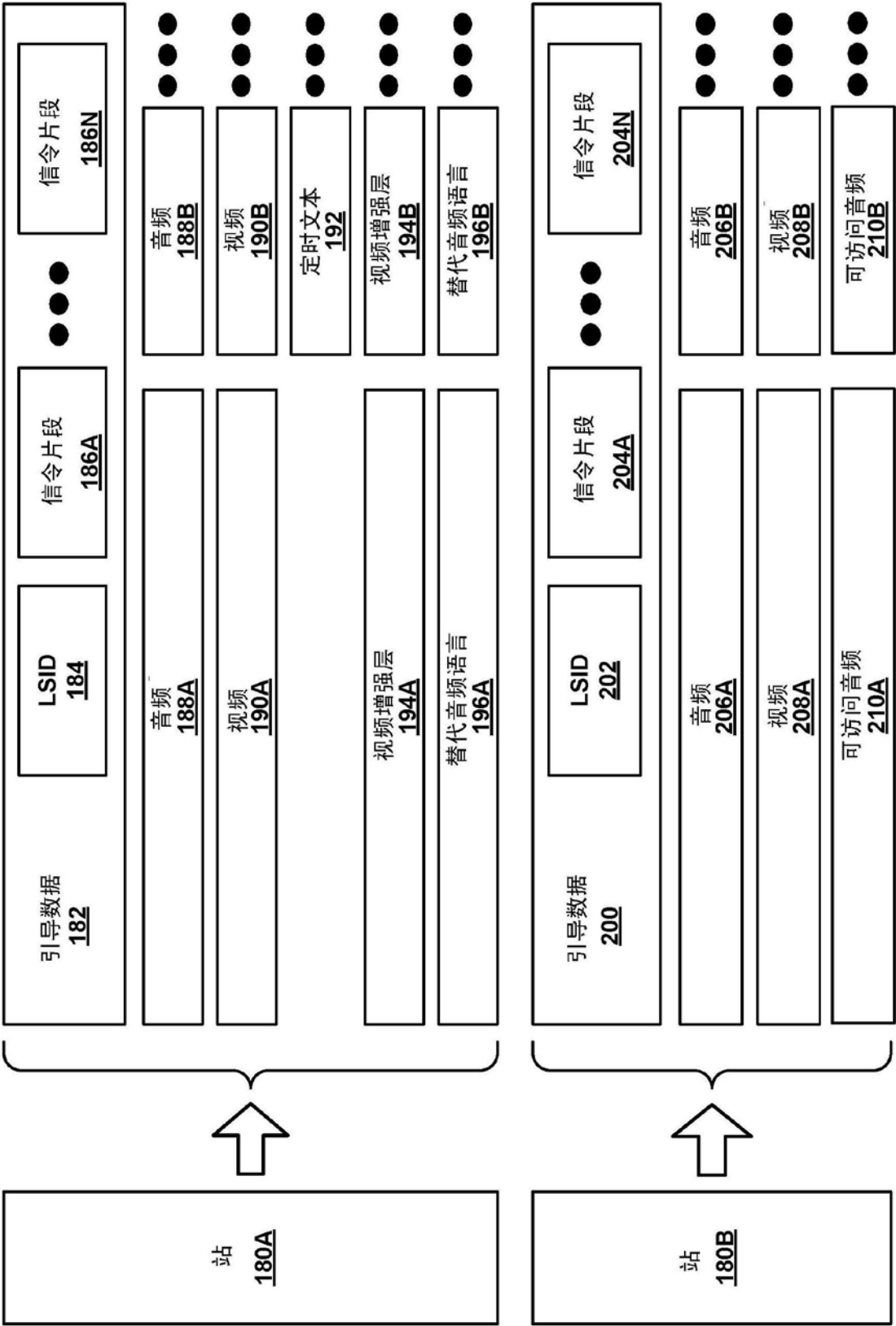


图4

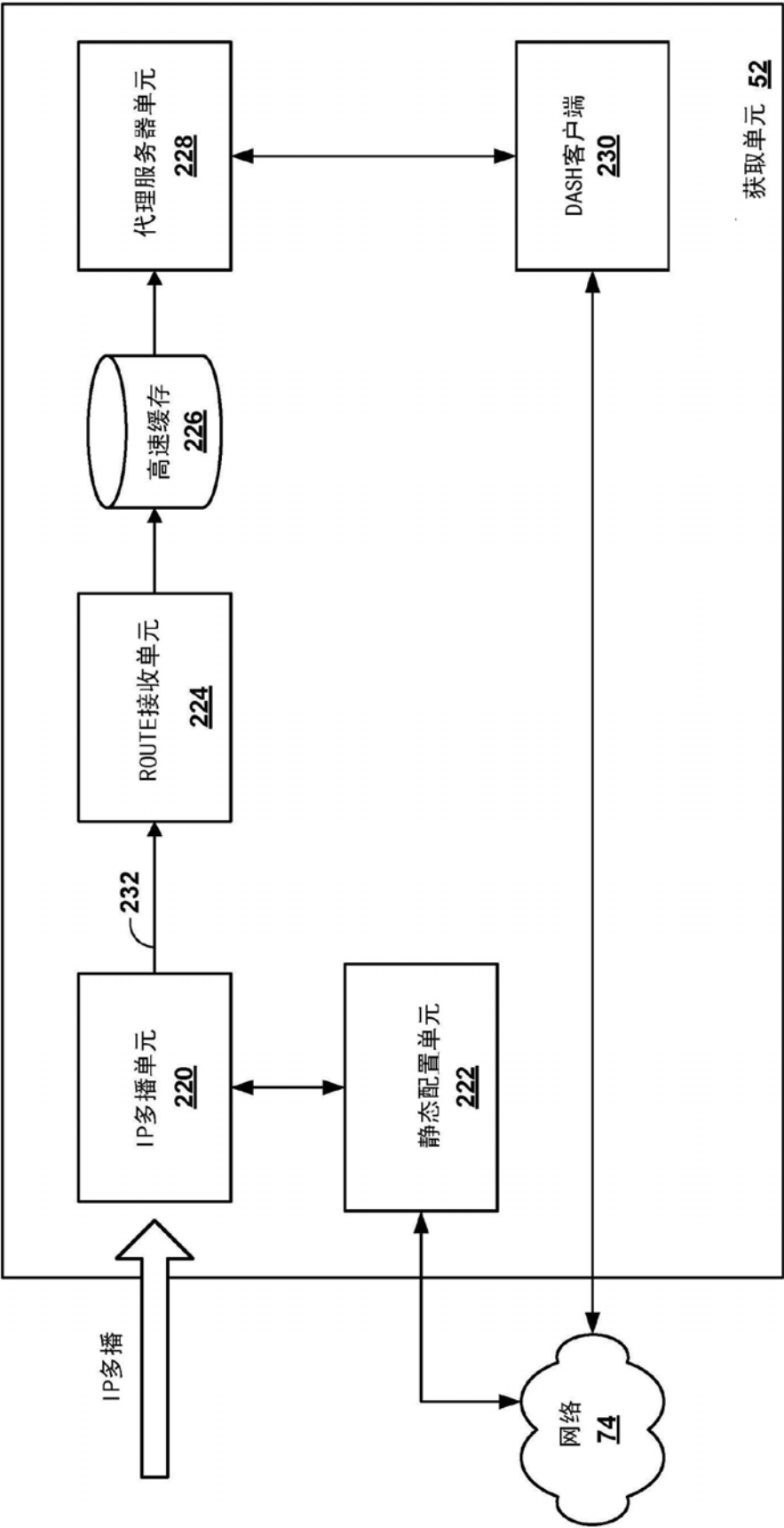


图5

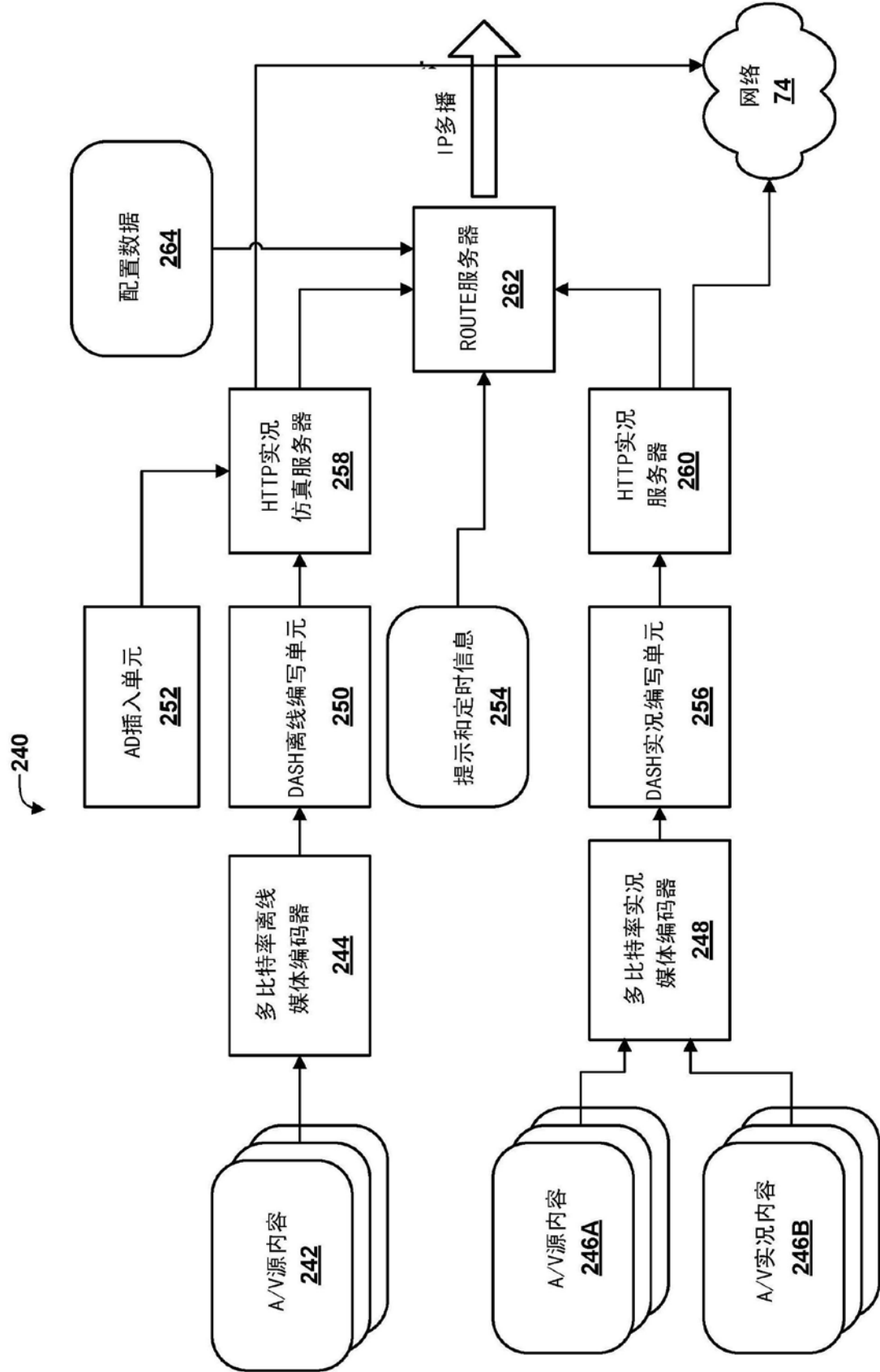


图6

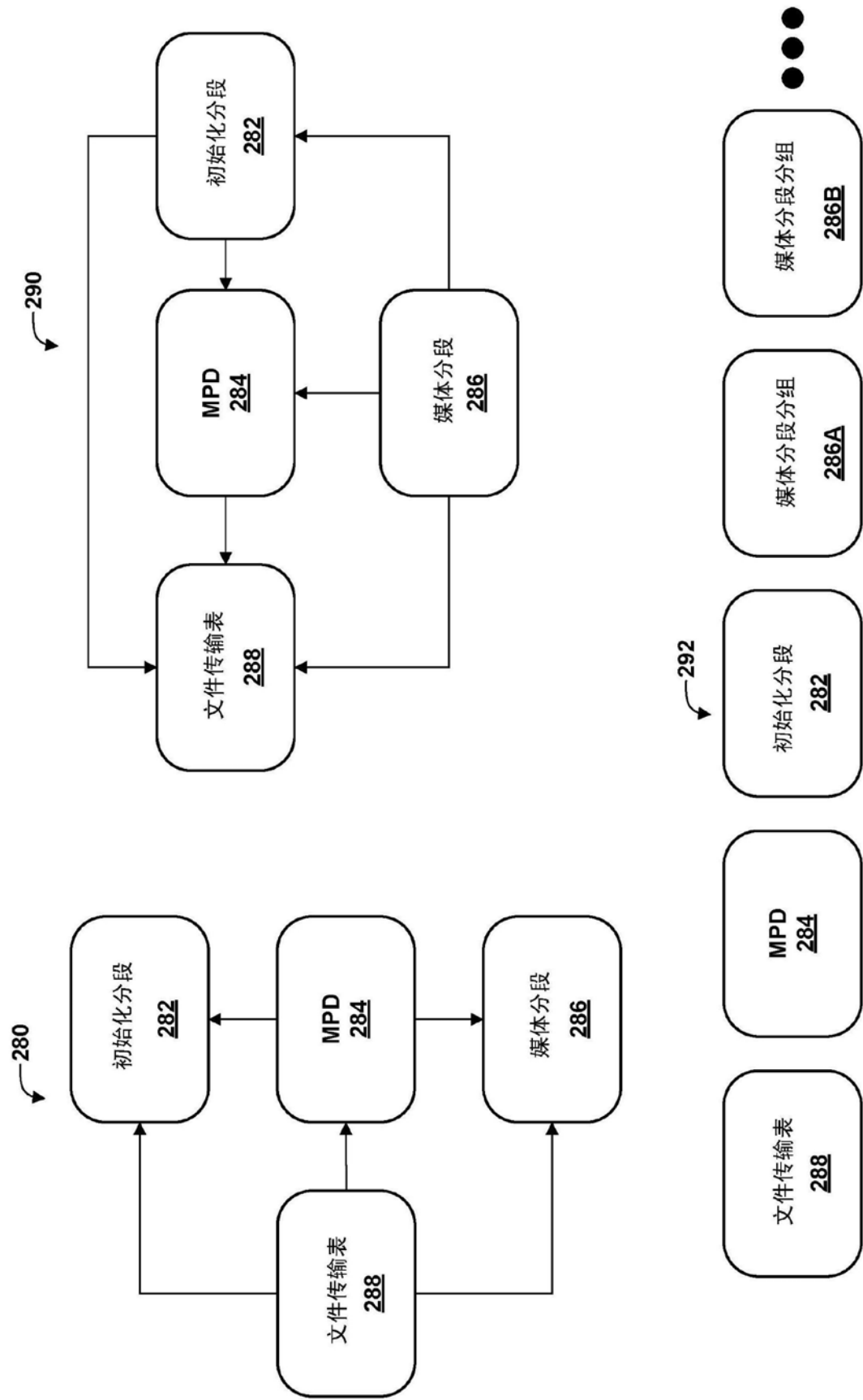


图7

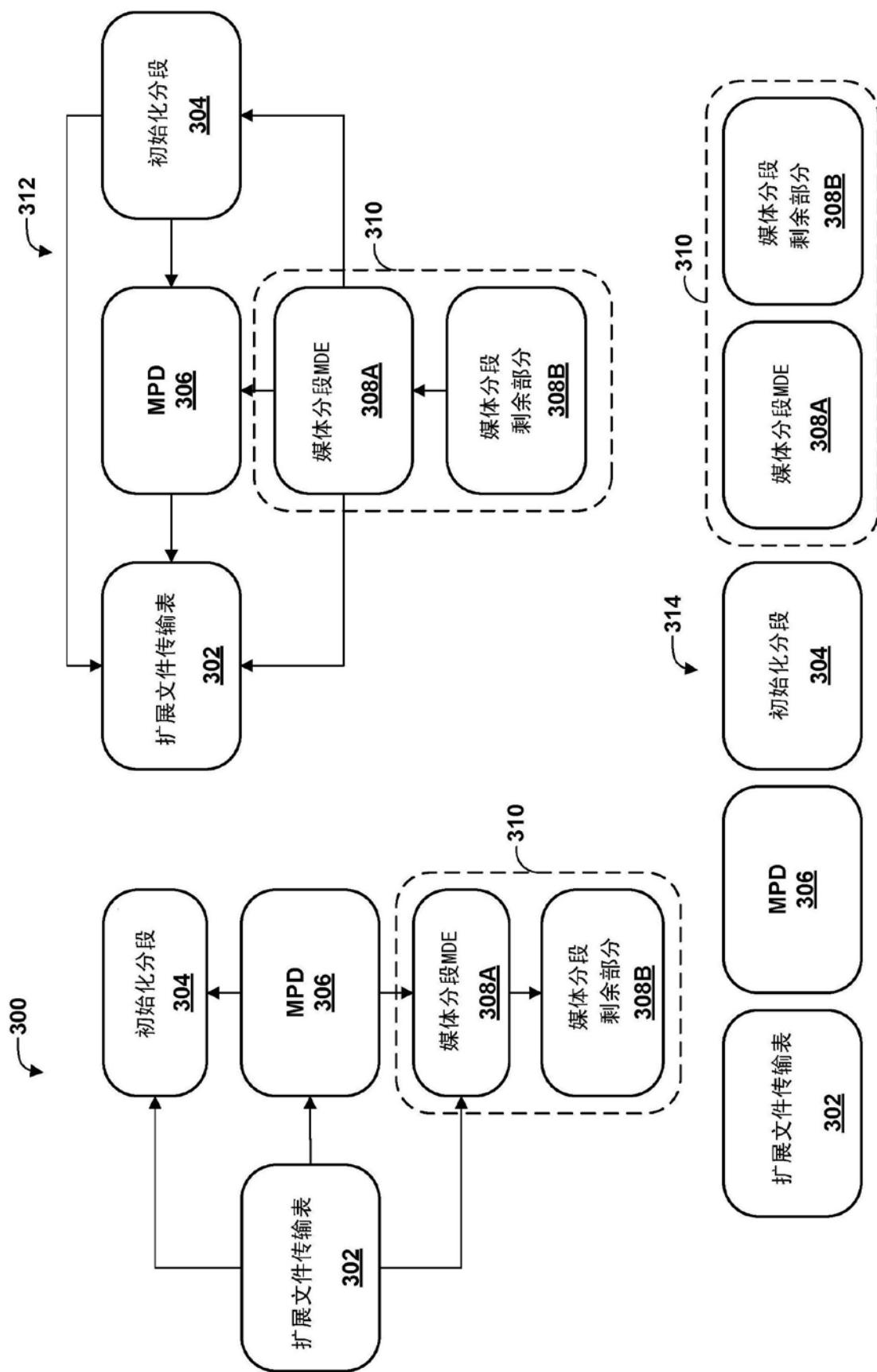


图8

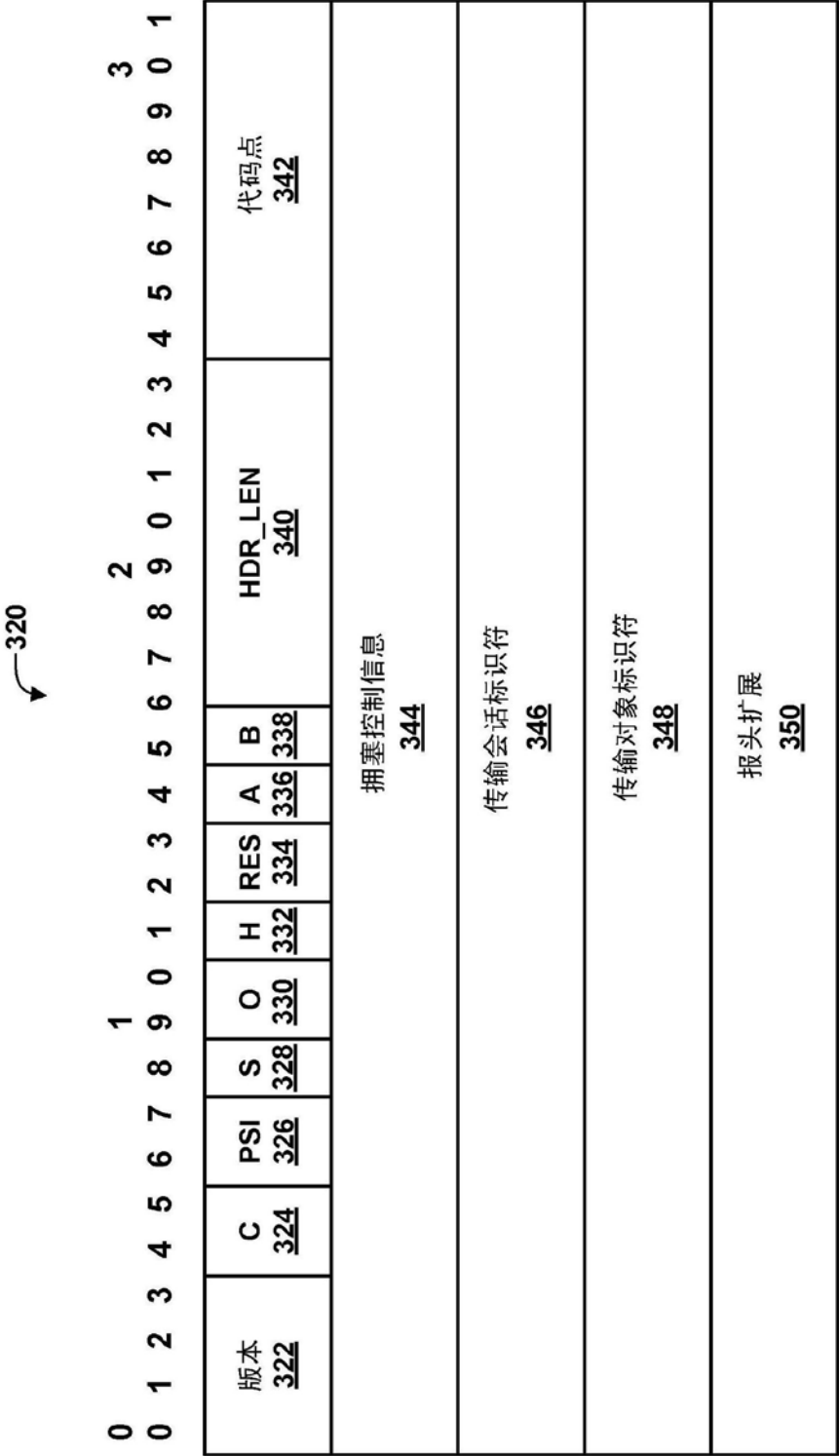


图9

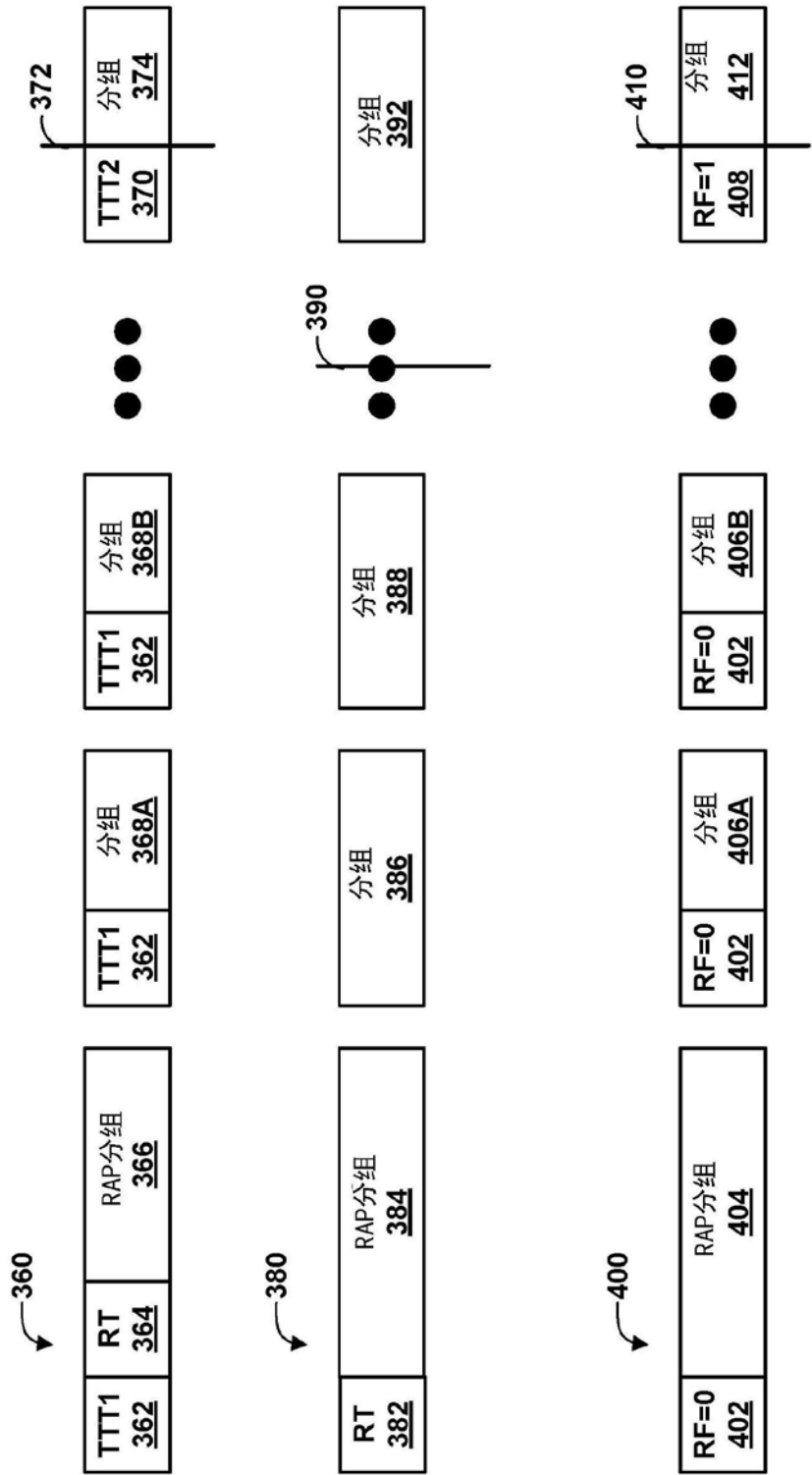


图10

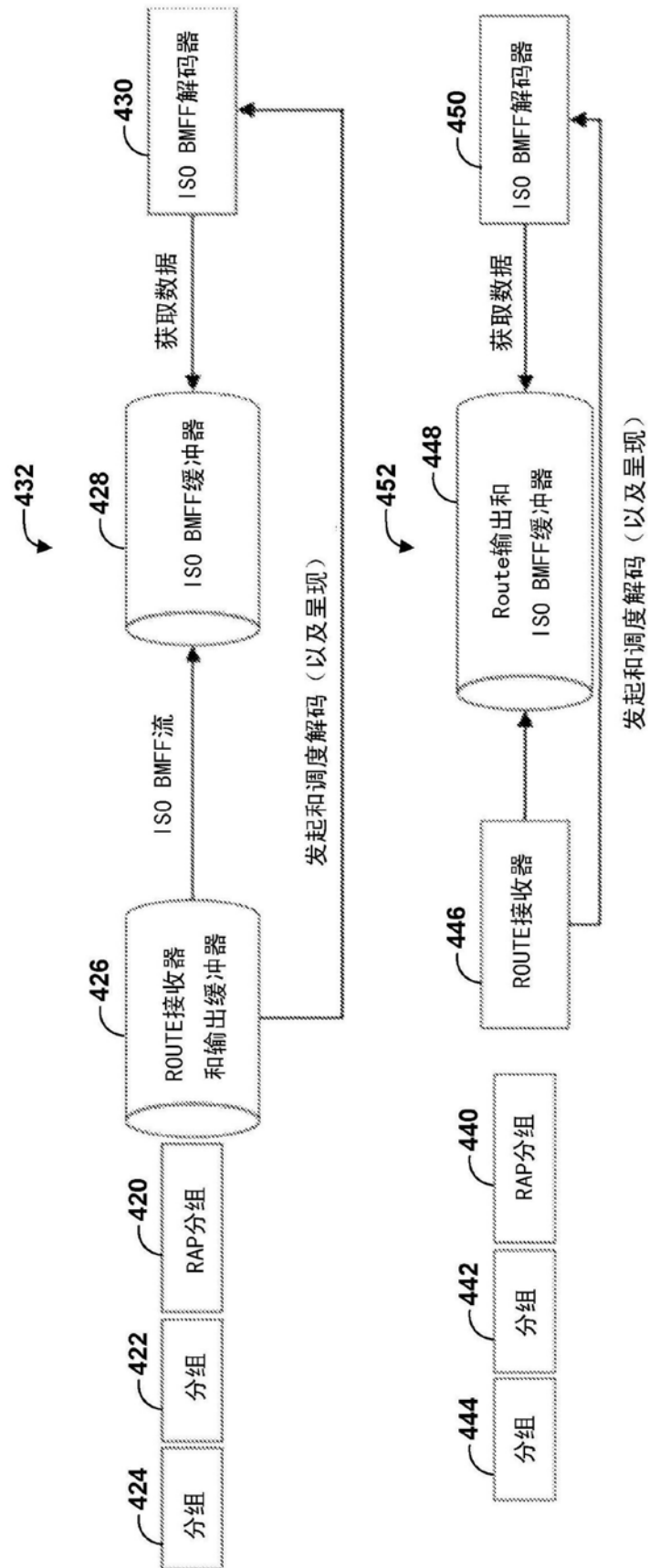


图11

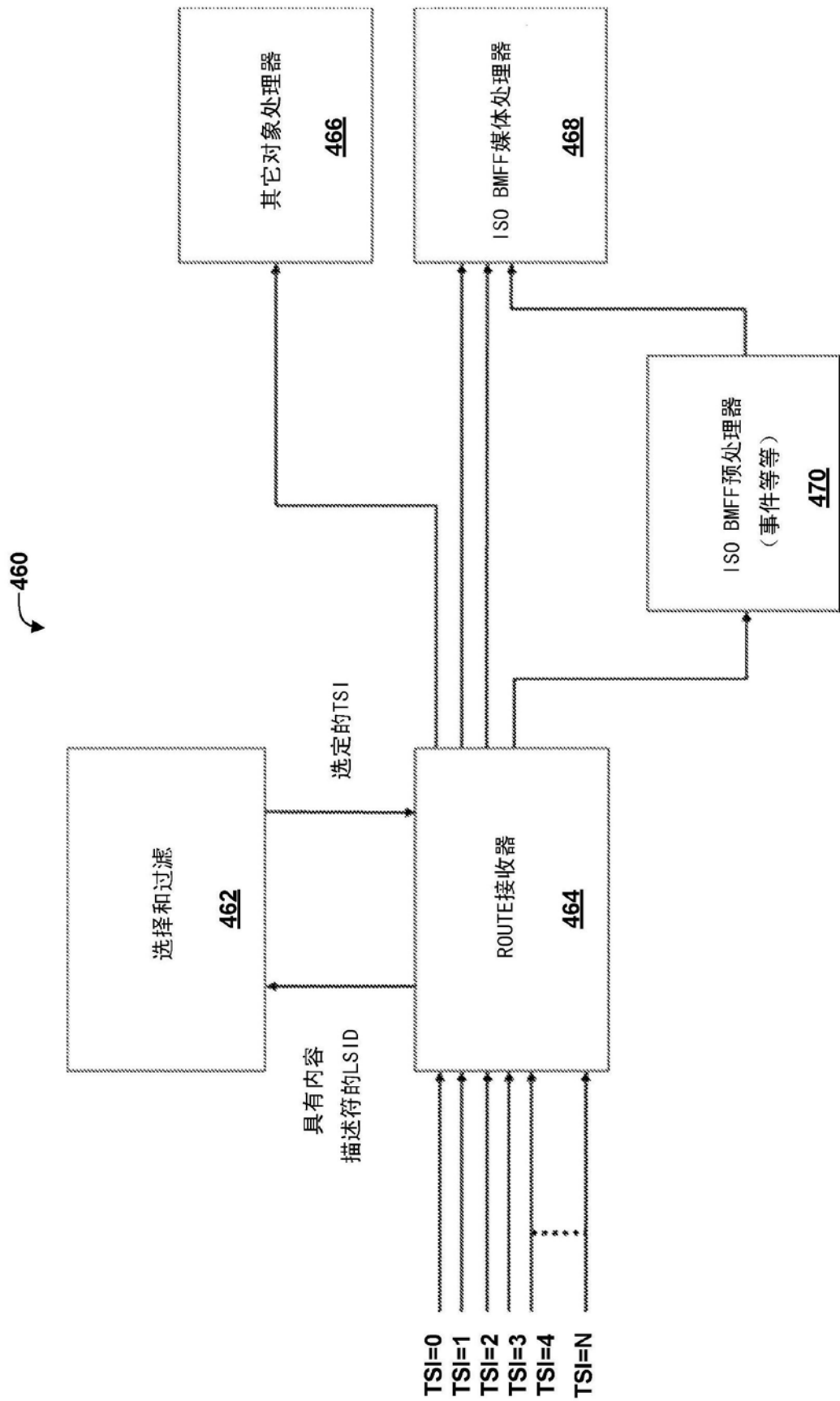


图12

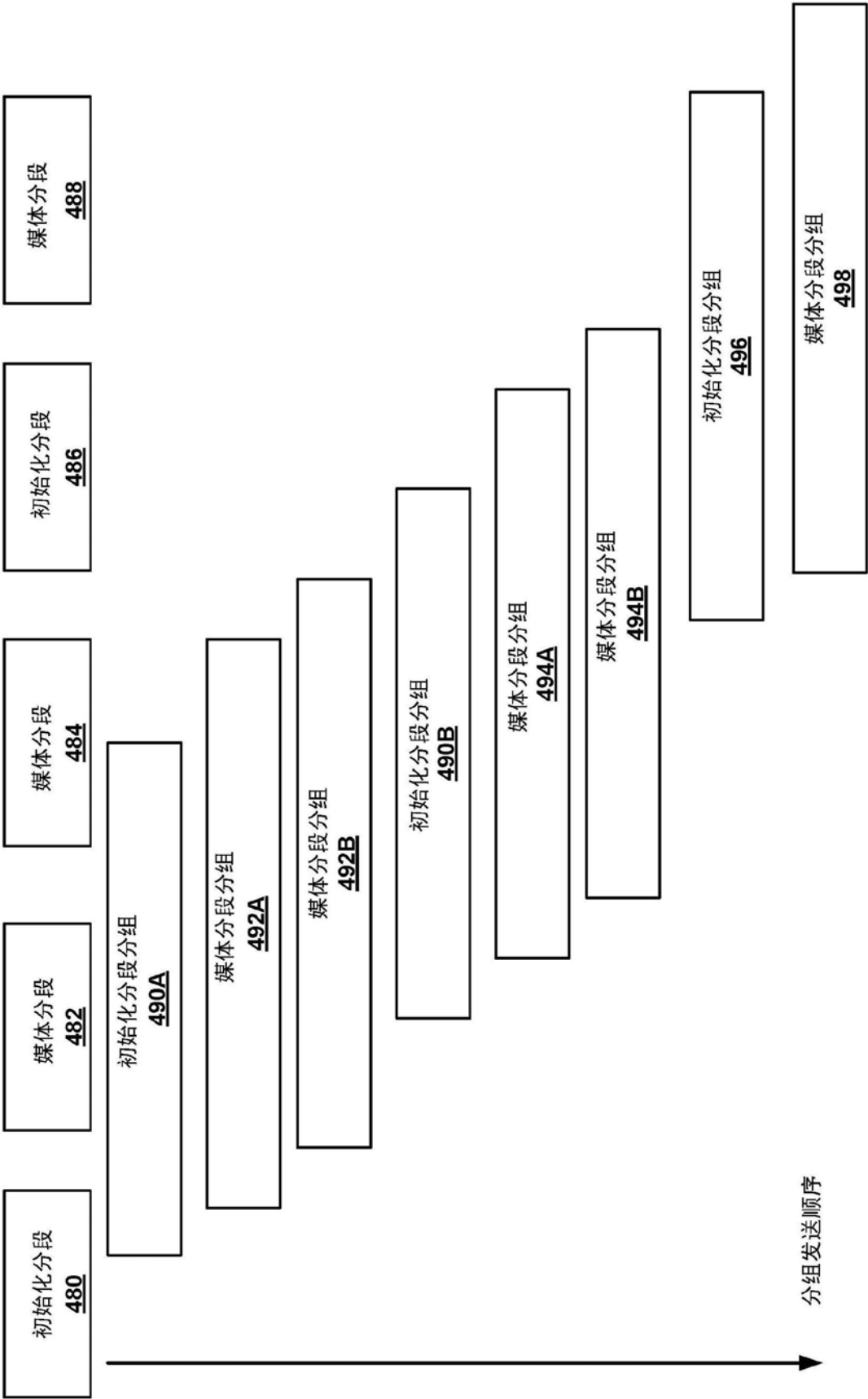


图13

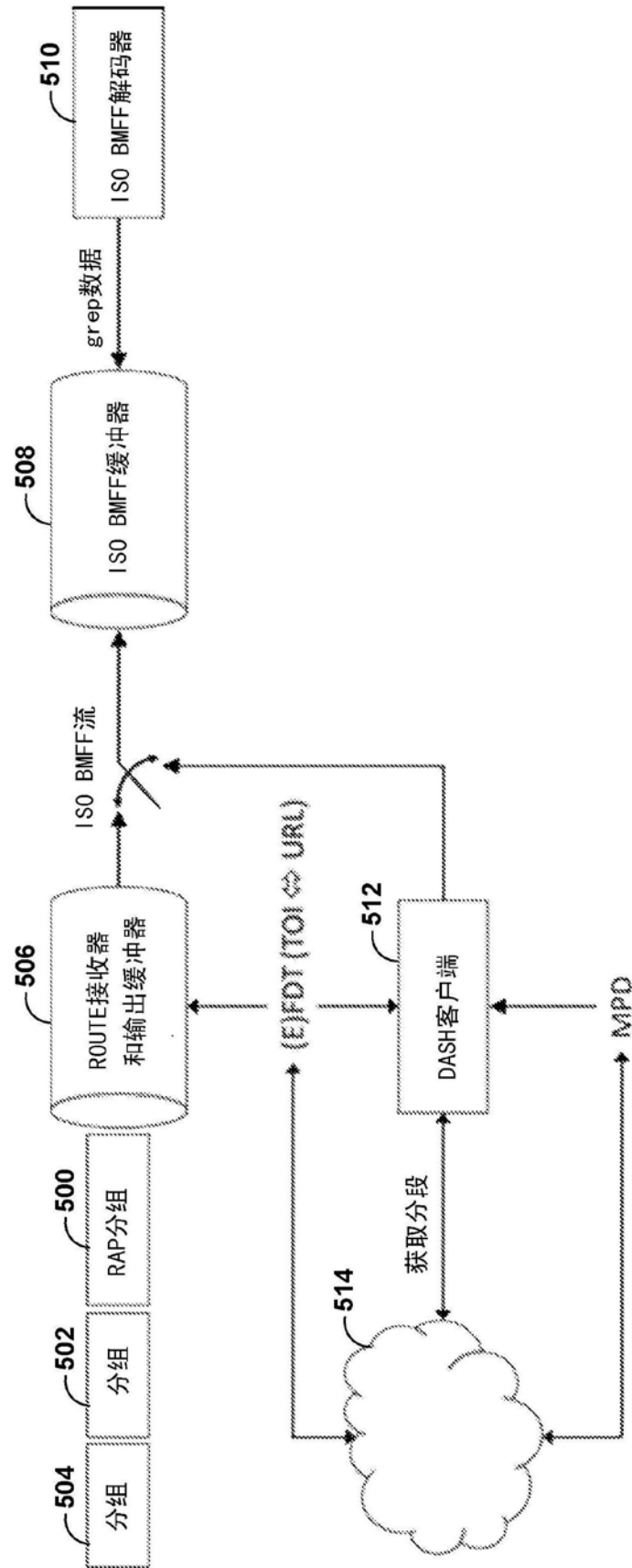


图14

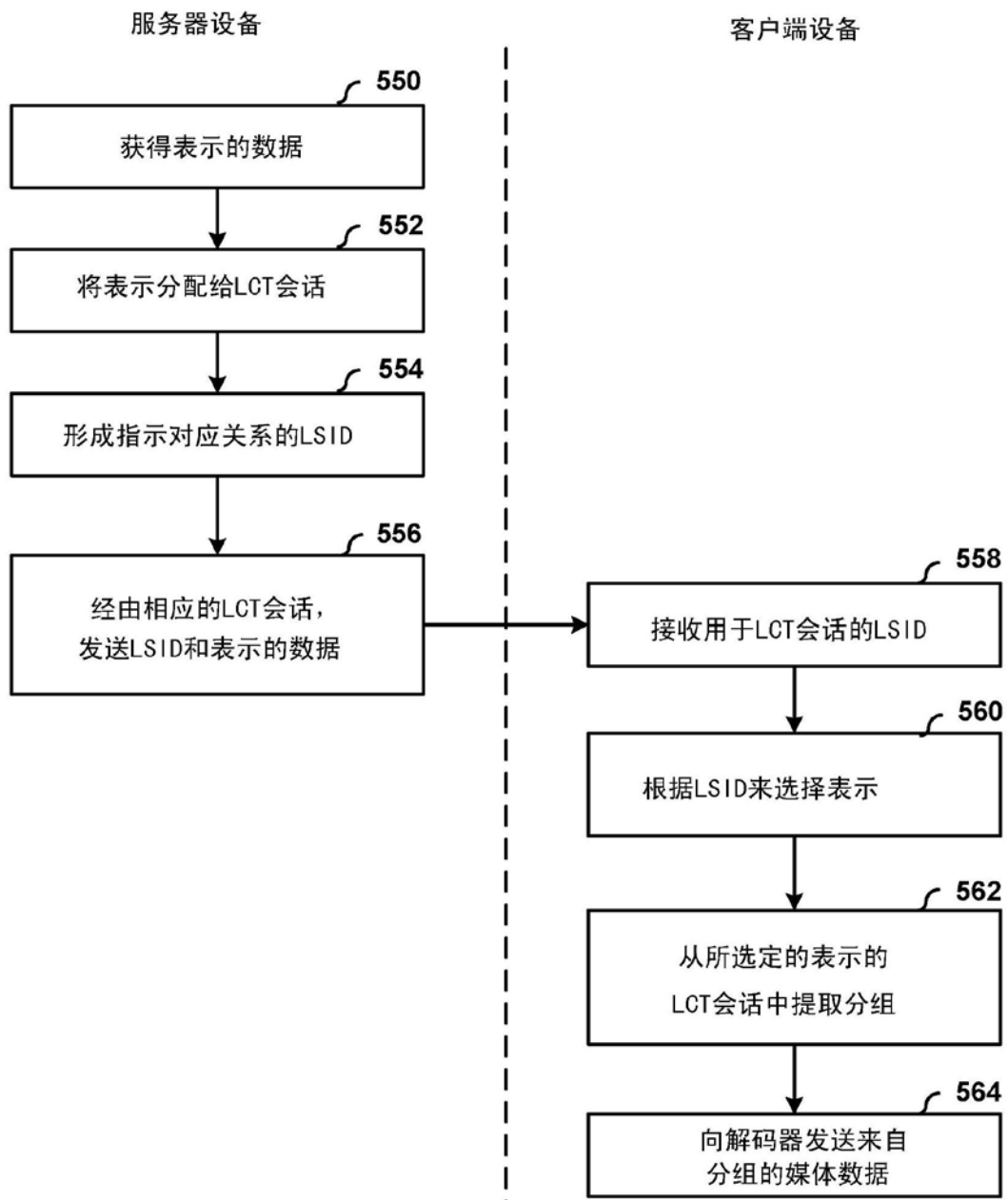


图15

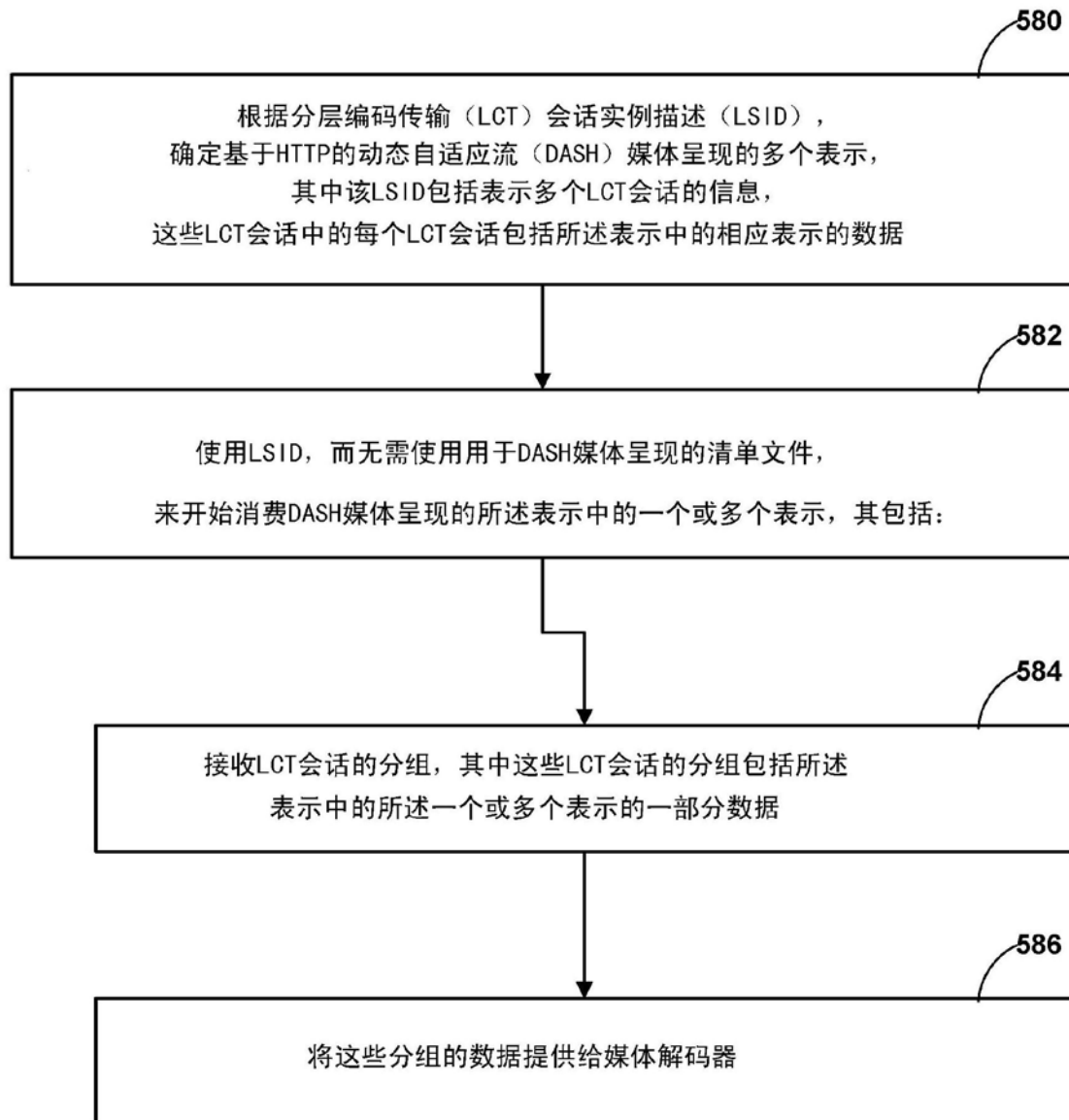


图16

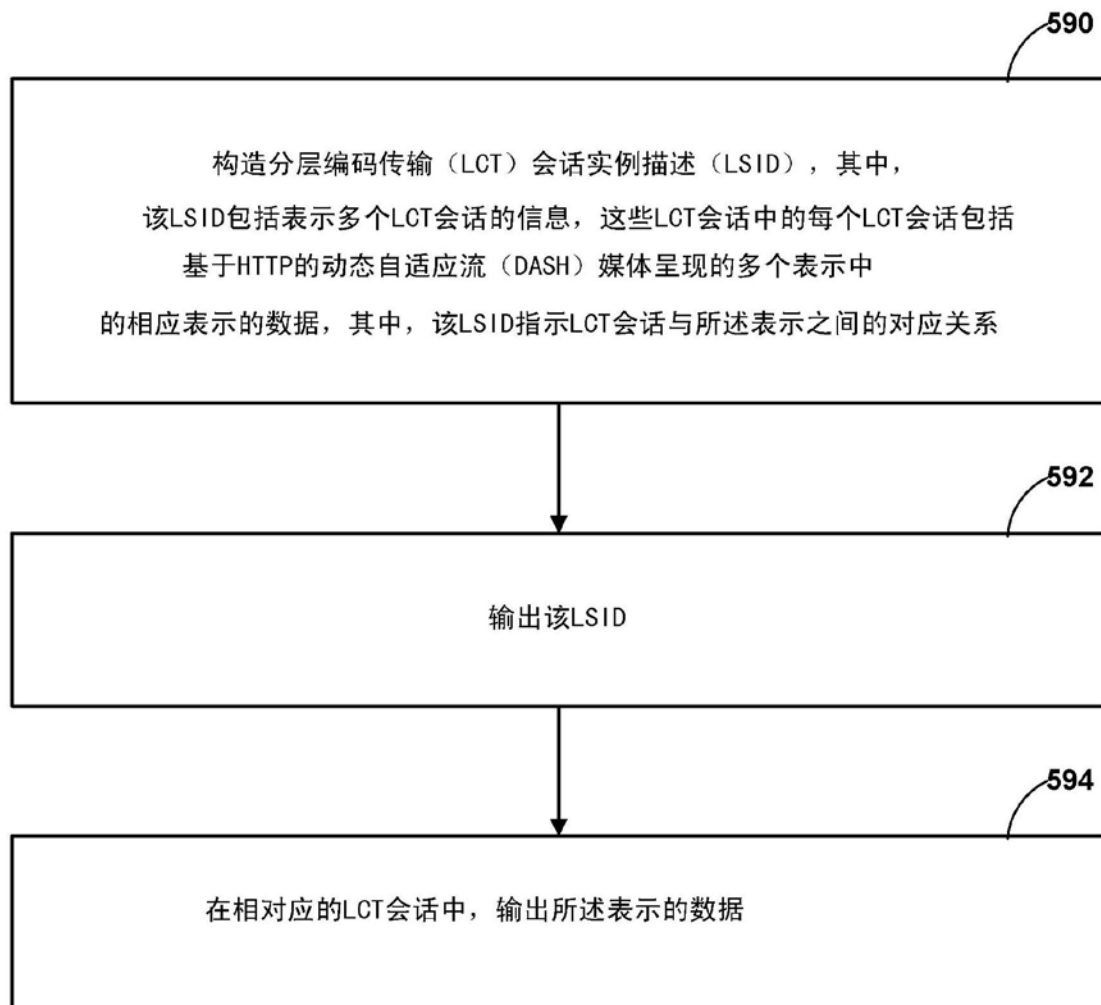


图17