

GY

中华人民共和国广播电影电视行业标准

GY/T 304—2016

高性能流化音频在 IP 网络上的 互操作性规范

High-performance streaming audio-over-IP interoperability

2017 - 01 - 04 发布

2017 - 01 - 04 实施

国家新闻出版广电总局

发布

目 次

前言	II
引言	III
1 范围	1
2 规范性引用文件	1
3 术语、定义和缩略语	2
4 同步	6
5 媒体时钟	6
6 传输	7
7 编码与成流	9
8 会话描述	12
9 发现服务	14
10 连接管理	14
附录 A (规范性附录) 媒体类别	16
附录 B (资料性附录) IEEE 802.1AS 时钟域接口	19
附录 C (资料性附录) 网络 QoS 配置建议	21
附录 D (资料性附录) AVB 网络传输	23
附录 E (资料性附录) 发现系统	26
参考文献	28

前 言

本标准按照GB/T 1.1—2009给出的规则起草。

请注意本标准的某些内容可能涉及专利。本标准的发布机构不承担识别这些专利的责任。

本标准由全国广播电影电视标准化技术委员会（SAC/TC 239）归口。

本标准起草单位：中央人民广播电台、中央电视台、国家新闻出版广电总局广播科学研究院、国家新闻出版广电总局广播电视规划院、北京英夫美迪科技股份有限公司、苏州市福川科技有限公司、北京众和传新科技有限公司、杭州联汇科技股份有限公司、上海佰贝科技发展有限公司。

本标准主要起草人：钱岳林、朱峰、罗攀、潘宇、张磊、王兰岚、庞超、张伟、邓向冬、韦安明、何晶、董晓坡、陈武、陈沁、唐卫平、练文杰。

引 言

高性能媒体网络支持以低延迟（小于10ms）传输专业质量的音频信号（采用线性PCM编码，采样率不低于44.1kHz，量化精度不低于16位），并适用于现场扩声。局域网及企业级网络的性能都可以满足作为高性能媒体网络的要求，但广域网或公共互联网的性能通常无法满足要求。

尽管截至本标准颁布前出现的使用各种专有和标准协议的此类媒体网络都基于IP协议，但它们之间无法互操作。

本标准是参照AES67—2015《High-performance streaming audio-over-IP interoperability》编制的。

本标准提出了高性能媒体网络互操作性的具体规定。本标准重点规定如何使用现有的协议来创建可互操作的系统。基于这一点，并未开发其他新的协议。

高性能流化音频在 IP 网络上的互操作性规范

1 范围

本标准规定了在IP网络上传输全频带和低噪声的高性能音频的互操作模式。

本标准适用于广播、音乐制作和影视后期制作设备间信号的交换，也可用于商业音频应用，如固定和流动的现场扩声。

2 规范性引用文件

下列文件对于本标准的应用是必不可少的。凡是注日期的引用文件，仅所注日期的版本适用于本标准。凡是不注日期的引用文件，其最新版本(包括所有的修改单)适用于本标准。

GB/T 25931—2010 网络测量和控制系统的精确时钟同步协议 (IEC 61588—2009, IDT)

IETF RFC 768 用户数据报协议 (User Datagram Protocol)

IETF RFC 791 互联网协议 (Internet Protocol)

IETF RFC 1112 IP组播的主机扩展 (Host Extensions for IP Multicasting)

IETF RFC 2236 互联网组管理协议，第二版 (Internet Group Management Protocol, Version 2)

IETF RFC 2474 IPv4和IPv6包头中的区分服务字段 (DS字段) 的定义 (Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers)

IETF RFC 2616 超文本传输协议-HTTP/1.1 (Hypertext Transfer Protocol - HTTP/1.1)

IETF RFC 2974 会话公告协议 (Session Announcement Protocol)

IETF RFC 3190 12位DAT音频和20位、24位线性采样音频的RTP有效载荷格式 (RTP Payload Format for 12-bit DAT Audio and 20- and 24-bit Linear Sampled Audio)

IETF RFC 3261 SIP:会话发起协议 (SIP:Session Initiation Protocol)

IETF RFC 3264 使用会话描述协议 (SDP) 实现会话提议/应答的模型 (An Offer/Answer Model with the Session Description Protocol (SDP))

IETF RFC 3376 互联网组管理协议，第三版 (Internet Group Management Protocol, Version 3)

IETF RFC 3550 RTP: 实时传输协议 (RTP: A Transport Protocol for Real-Time Applications)

IETF RFC 3551 使用最小控制的音视频会议RTP 类别 (RTP Profile for Audio and Video Conferences with Minimal Control)

IETF RFC 4566 会话描述协议 (Session Description Protocol)

IETF RFC 7273 RTP时钟源信令 (RTP Clock Source Signalling)

AES5—2008 脉冲编码调制采用的优先采样频率 (AES recommended practice for professional digital audio—Preferred sampling frequencies for applications employing pulse-code modulation)

AES11—2009 (R2014) 演播室数字音频系统同步 (AES recommended practice for digital audio engineering —Synchronization of digital audio equipment in studio operations)

EBU Tech 3326 音频在IP网络上的互操作要求 (Audio contribution over IP—Requirements for Interoperability)

IEEE 1588—2002 网络测量和控制系统的精确时钟同步协议 (IEEE Standard for a Precision Clock Synchronization Protocol for Networked Measurement and Control Systems)

IEEE 802.1AS—2011 桥接局域网中时间敏感应用的定时与同步 (Timing and Synchronization for Time-Sensitive Applications in Bridged Local Area Networks)

IEEE 802.1Q—2011 媒体接入控制桥和虚拟桥局域网 (Media Access Control (MAC) Bridges and Virtual Bridged Local Area Networks)

3 术语、定义和缩略语

3.1 术语和定义

下列术语和定义适用于本标准。

3.1.1

精确时间协议 precision time protocol; PTP

由IEEE 1588—2002、GB/T 25931—2010和IEEE 802.1AS—2011定义的通用时钟分发协议。

3.1.2

边界时钟 boundary clock

在一个域中具有多个精确时间协议 (PTP) 端口, 并维护该域中所用时标的时钟。它可作为时间源, 即为主时钟; 也可与另一个时钟同步, 即为从时钟。

[GB/T 25931—2010, 定义3.1.3]

3.1.3

数字音频参考信号 digital audio reference signal; DARS

AES11—2009 (R2014) 中定义的音频时钟信号。

3.1.4

汇入源 contributing source; CSRC

为RTP混合器生成组合流起到贡献作用的RTP流输入源。

3.1.5

区分服务 differentiated services

对IP网络流量分级并提供不同QoS保障的系统。

3.1.6

区分服务代码点 differentiated services code point; DSCP

位于IP包头中用于分级的一个6位字段, 是区分服务架构中的一部分。

3.1.7

64位扩展的唯一标识符 64 bit extended unique identifier; EUI-64

由一个24位或36位的公司注册标识符和一个公司唯一设备标识符组合构成的64位全球唯一标识符。

3.1.8

最高级时钟 grandmaster

通过PTP进行时钟分发所需的同步主时钟源。最高级时钟是用64位扩展的唯一标识符（EUI-64）标识的一种网络设备。

3.1.9

最高级时钟标识符 grandmaster identifier; GMID

一种EUI-64唯一标识符，用于标识为同步域提供服务的最高级时钟，在GB/T 25931—2010和IEEE 802.1AS—2011同步标准中规定。

3.1.10

互联网组管理协议 internet group management protocol; IGMP

主机用来向IPv4路由器报告其组播组成员的通信协议。

3.1.11

链路偏移量 link offset

媒体消耗在网络、发送器的缓存和接收器的缓存中的总时间。

3.1.12

媒体时钟 media clock

发送器用于采样和接收器用于播放数字媒体流的时钟。音频流的媒体时钟以音频样值数标注。

3.1.13

媒体包 media packet

媒体流的一部分，承载媒体数据的数据包。每个媒体包包含一个或多个音频通道的一个或多个样值。

3.1.14

最大传输单元 maximum transmission unit; MTU

特定数据链接中能传输的最大IP数据包大小，以字节为单位。以太网数据链路的MTU为1500字节。

3.1.15

网络时钟 network clock

由第4章定义的网络同步机制提供的时间，以秒为单位。

3.1.16

开放式系统互联模型 open systems interconnection model; OSI model

以抽象层的方式描述和规范了通信系统的功能。

3.1.17

网络层 network layer

OSI model的第3层，负责将可变长度数据序列从源转发和路由到目的地。

3.1.18

包时间 packet time

媒体包中的媒体数据的实际持续时长。

3.1.19

服务质量 quality of service; QoS

描述依据性能要求，在网络中对流量分级、标记和传输的系统。

3.1.20

接收器 receiver

一种可以从网络接收媒体流的网络设备。

3.1.21

请求注释 request for comment; RFC

由IETF发布的、与Internet及Internet互联系统的操作相关的规范文档，以编号作为索引。

3.1.22

实时传输协议 real-time transport protocol; RTP

一种由RFC 3550定义并为应用通过UDP/IP网络构建、标记和传输媒体数据包的协议。

3.1.23

实时传输控制协议 real-time transport control protocol; RTCP

实时传输协议（RTP）的伴生协议，为RTP媒体数据包提供统计分析和控制信息。

3.1.24

RTP 时钟 RTP clock

在包含流数据的RTP包中携带的时间戳。每个流都有自己的RTP时钟。

3.1.25

RTP 会话 RTP session

一种发送器与接收器之间的、基于RTP协议的媒体连接。RTP会话可以是单播或组播形式。

3.1.26

RTP 流 RTP stream

由已规定的时间间隔发送的媒体数据组成的RTP包串。一个流可以包含多个通道。每个RTP会话可以由多个媒体流构成。

3.1.27

音频流 audio stream

媒体数据为音频的RTP流。

3.1.28

会话描述协议 session description protocol; SDP

一种用于描述RTP会话和操作属性的格式，包括网络寻址、编码格式和其他元数据属性。

3.1.29

发送器 sender

一种可以将媒体流发送到网络上的网络设备。

3.1.30

SIP URI

一种在SIP协议中用于识别用户代理URI的字段。SIP URI采用sip:<user>@<domain>或sips:<user>@<domain>的描述形式，见10.2.1。

3.1.31

从时钟 slave clock

一种使用精确时间协议（PTP）与主时钟（即时钟提供者）保持同步的时钟。从时钟可以作为其他时钟的主时钟，也可以作为边界时钟。

3.1.32

传输层安全协议 transport layer security; TLS

一种IP网络安全通信加密协议。

3.1.33

透明时钟 transparent clock

测量精确时间协议（PTP）事件报文通过该设备的时间，并向接收该PTP事件报文的时钟提供该信息的设备。

3.1.34

传输层 transport layer

OSI model的第4层，为网络应用提供端到端的通信服务。

3.1.35

用户代理 user agent

一种SIP终端设备，例如VoIP电话机。

3.1.36

历元 epoch

时标的原点。

[GB/T 25931—2010，定义3.1.9]

3.1.37

快速转发 expedited forwarding

区分服务的其中一种分级，具有低延时、低丢包率和低抖动的特点，适用于语音、视频和其他实时服务。快速转发流量相较于别的流量，通常拥有严格优先级队列。

3.1.38

普通时钟 ordinary clock

在一个域中具有单个精确时间协议（PTP）端口，并维护该域中所用时标的时钟。它可作为时间源，即为主时钟；或与另一个时钟同步，即为从时钟。

[GB/T 25931—2010，定义3.1.23]

3.2 缩略语

下列缩略语适用于本标准。

ARB 任意 (Arbitrary)

- AVB 音视频桥接 (Audio Video Bridging)
- BE 尽力转发 (Best Effort)
- DNS 域名服务系统 (Domain Name System)
- IEEE 电气和电子工程师协会 (Institute of Electrical and Electronics Engineers)
- IETF 互联网工程任务组 (Internet Engineering Task Force)
- IP 互联网协议 (Internet Protocol)
- IPv4 互联网协议版本4 (Internet Protocol version 4)
- IPv6 互联网协议版本6 (Internet Protocol version 6)
- SAP 会话公告协议 (Session Announcement Protocol)
- SIP 会话发起协议 (Session Initiation Protocol)
- UDP 用户数据报协议 (User Datagram Protocol)
- URI 统一资源标识符 (Uniform Resource Identifier)

4 同步

4.1 概述

高性能流化音频区别于低性能流化音频 (如互联网广播和IP电话) 之处是运行前者的设备或应用可共用一套精确的统一时钟。通过时钟的统一, 网络中的任何一台接收器都能与其他的接收器同步回放。通过时钟的统一, 发送器和接收器之间的延时固定并可测量。统一时钟确保所有流的采样速率和还原速率相同。接收器更易于合成具有相同采样率的音频数据流。这一特性对网络音频设备的高效运行特别关键, 如数字调音台。

时钟的同步应通过GB/T 25931—2010中规定的精确时间协议 (PTP) 来实现。

GB/T 25931—2010定义了各种同步传输应用类别 (profile)。这些类别描述了协议的属性、可选项和设备的性能需求。GB/T 25931—2010为延时请求-响应机制 (GB/T 25931—2010附录J中J.3) 和点到点延时机制 (GB/T 25931—2010中J.4) 定义了缺省的类别。

除了部分AVB设备 (见下段) 外, 其他设备均应支持GB/T 25931—2010缺省的类别。支持缺省类别的设备应使用GB/T 25931—2010附录D规定的IPv4封装。

唯一的例外是, 两种设备不需要实现GB/T 25931—2010缺省的类别, 一种是4.4中描述的使用AVB同步机制的设备, 另一种是为完成媒体流传输而连接到AVB网络的设备。

4.2 IP 网络同步

标准IP网络中的设备宜使用附录A中定义的媒体类别, 以确保为各种应用提供所需的性能, 也可使用缺省类别, 但锁定时间会延长、精度会降低。

4.3 GB/T 25931—2010 网络同步

在GB/T 25931—2010 (边界时钟或透明时钟) 交换机构建的网络上恰当地使用缺省类别, 可满足音频传输所需的性能。

注: 由于性能的限制, 一些GB/T 25931—2010网络设备可能不支持媒体类别。

4.4 AVB 网络同步

IEEE 802.1Q—2011中定义的增强型以太网 (即音视频桥接, AVB) 按照IEEE 802.1AS—2011的规定分发同步信息。IEEE 802.1AS—2011定义了一个GB/T 25931—2010的类别。相对于缺省类别或媒体类别, AVB网络可优先使用自有的IEEE 802.1AS—2011同步类别。使用GB/T 25931—2010和IEEE 802.1AS—2011搭建异构同步网络的方法参见附录B。

5 媒体时钟

发送器使用媒体时钟进行采样，接收器使用媒体时钟播放数字媒体流。媒体时钟与网络时钟存在确定的对应关系。媒体时钟与网络时钟应共用GB/T 25931—2010中7.2.2定义的历元，即1970年1月1日 00:00:00 TAI。通过网络传输的数字音频应根据媒体时钟进行采样，或者根据媒体时钟转换采样频率。

注：由于1972年伊始引入了闰秒，1972年及之后的TAI和UTC时间戳的偏移量变成了整数秒，而1971年及之前该值为非整数秒。这导致出现了另一个可用的历元—1969年12月31日 23:59:51.999918 UTC，见GB/T 25931—2010中7.2.2。这种非整秒的UTC时间偏移量仅在1972年之前的网络时钟时间中存在。

与网络时钟相比，媒体时钟拥有更精准的速率，此速率应与音频采样率相同。

本标准支持三种采样率：44.1kHz、48kHz和96kHz（见7.2）。在一秒网络时钟中，如果音频流的采样率为48kHz，则媒体时钟前进48000个采样时间。在GB/T 25931—2010中7.2.2定义的历元，媒体时钟的值应为0，每一个采样周期过后自动加1。

RTP时钟相对媒体时钟有恒定的偏移量，此偏移量应在每个流的会话描述（见8.4）中表示。

在网络协议和管理接口中，RTP和媒体时钟通常以32位整数表示。48kHz流的媒体时钟大约每24.86小时会产生溢出。为了确保与网络时钟同步，以32位整数表示的媒体时钟应准确地处理所有发生在历元与当前时间之间的溢出（反转）。

6 传输

6.1 概述

本章介绍编码和封包的媒体数据如何在网络中传输。基于开放式系统互联模型（OSI model），本章定义第3层（网络层）和第4层（传输层）上的操作方式。本标准并未规定如何在OSI model的更低层实现互操作。

本标准是基于理想的IP网络传输技术。

注：IP报文在以太网上的传输标准见RFC 894。

6.2 网络层

媒体数据包应使用RFC 791中定义的IPv4传输。

注1：本标准暂不支持IPv6。

尽管RFC 791中要求支持分片数据包重组，但本标准对接收器不作此要求。不支持分片数据包重组的接收器应忽略IP分片数据包。

发送器可在传出的媒体数据包的IP包头中设置禁止拆分标记（DF）位。如果网络要对标记为DF的数据包进行分片，此数据包将被丢弃，发送器将收到ICMP“Too Big”消息。发送器在收到此消息后，宜终止传输该流。

组播消息，例如同步消息，应使用RFC 1112中描述的IP组播实现。

注2：有关IP组播的更多信息见RFC 3170。

所有设备应支持RFC 2236中定义的IGMPv2，也可支持RFC 3376中定义的IGMPv3。

注3：上文提出支持IGMPv2的要求是因为运行于IGMPv2网络中的IGMPv3设备需两分钟启动延时来查找网络中的IGMPv3服务。

注4：RFC 2236和RFC 3376提出了向下兼容的要求。支持IGMPv2的设备可以在IGMPv1或IGMPv2的网络中正常运行。支持IGMPv3的设备可以在IGMPv1、IGMPv2或IGMPv3的网络中正常运行。

设备应使用互联网组管理协议（IGMP）来请求接收所需的组播，包括接收同步信息（见GB/T 25931—2010附录D中D.3）、使用组播寻址的媒体流（见7.7）以及设备上可能使用的其他应用协议信息（例如第9章的发现服务）。

注5：在某些路由重新配置的情况下，IGMP注册数据可能会被网络清除，这可能会导致流数据中断。重传IGMP成员报告，是一种有效的快速恢复服务的方法，可通过关闭并立即重启受影响的组播网络套接字来实现。

在发送组播媒体数据包之前，发送器应使用IGMP协议向接收器发送查询并收到确认报文。

注6：通过发送IGMP请求，发送器实际上不会接收到它发送的数据包，而是为了阻止不必要的组播媒体数据泛滥。有些IGMP探测的实现可能会造成无注册成员的组播组数据包泛滥。

注7：这种需求通常以发送器接收实时传输控制协议（RTCP）消息的方式实现。在 RFC 3550 中，虽然建议但不强求设备发送或接收 RTCP 数据包。

6.3 服务质量

若网络中同时存在无管理的非实时流量和时间敏感的媒体流量时，后者通常具备优先处理权，即 QoS。为了在网络中有效实施适当的 QoS 策略，设备应使用 RFC 2474 中所述的区分服务方法。区分服务会根据流量分级，通过 IP 包头中的 DSCP 字段来标记数据包，这样网络就能轻易识别需要优先处理的数据包。

所支持的流量等级划分最少应包含表 1 中所示的三种。设备应使用合适的 DSCP 值标记输出的流量。对于设备，DSCP 字段宜使用表 1 所示的缺省值，但网络管理员或用户也可以通过管理接口，以其他的 DSCP 值标记流量。发送器可对多个等级使用同一 DSCP 值，以达到等级合并的效果。本标准不要求设备具有管理接口，设备可仅使用缺省值。

当有极低传输延时需求的媒体流与使用较长包时间的媒体流以同一 QoS 等级传输时，前者可能无法在网络中得到低延时的传输保障。为了区别对待有不同需求的媒体流，发送器可配置使用表 1 规定之外的服务等级。

表1 QoS 等级与区分服务的对应关系

等级名称	流量类型	缺省区分服务等级 (DSCP 十进制值)
时钟	GB/T 25931—2010 规定的 <i>Announce, Sync, Follow_Up, Delay_Req, Delay_Resp, Pdelay_Req, Pdelay_Resp</i> 和 <i>Pdelay_Resp_Follow_Up</i> 串	EF (46)
媒体	RTP 和 RTCP 媒体流数据	AF41 (34)
尽力转发	GB/T 25931—2010 规定的信令与管理消息、发现和连接管理消息	DF (0)

注1：数据包上的 DSCP 标记并未定义网络设备的任何特定行为，或暗示网络应执行的特定策略。作为一种安全措施，网络甚至可忽略传入的 DSCP 标记，转而通过其他方式（例如：UDP 端口号，IP 寻址）对流量进行识别和分级。此类网络问题不在本标准范围内，但附录 C 为网络管理员提供了指导性建议。

注2：对于本标准范围外的 DSCP 标记，本标准未做规定。大部分系统将其他应用产生的流量标记为 DF (0)，这与本标准的要求一致。

对传出的 RTCP 数据包，发送器宜使用与其对应的 RTP 流数据包相同的 DSCP 值进行标记。接收器如果传输类似的流，对于传出的 RTCP 数据包，宜使用与其对应的 RTP 数据包相同的 DSCP 值进行标记。

接收器不应根据接收到的数据包的 DSCP 标记做任何等级关联假设。

注3：DSCP 标记在网络路由过程中可能会被更改。发送器也可能在不通知接收器的情况下更改 DSCP 标记。

6.4 传输层

传输层为网络设备之间提供端到端的通信。传输层处理数据包丢失和重新排序问题，以及实现多路复用，使用单个网络连接能为终端站点上的多个应用提供服务。

设备应使用 RFC 3550 中定义的实时传输协议。设备应运行与 RFC 3551 中定义的满足最低控制要求的音视频会议一致的 RTP 类别。设备宜使用为 RTP 分配的缺省端口：RTP 使用 5004，RTCP 使用 5005（见 RFC 3551 第 8 章）。发送器可使用其他端口。接收器应支持对应的端口。

设备应使用 RFC 768 中定义的 UDP 协议传输 RTP 数据包。

本标准不支持数据包分片，也不要求接收器重组（见 6.2），同时假设以太网采用 1500 字节的最大传输单元（MTU），为防止数据包在标准的 IPv4 以太网中传输时出现分片，以及确保日后与 IPv6 兼容，最大的 RTP 有效载荷应为 1440 字节。

注：在 MTU 比以太网的 1500 字节更小的网络连接中，发送器可能希望使用更小的最大有效载荷。

尽管RTP协议栈支持使用传输层安全协议（TLS）加密流，但本标准并不支持。

发送器不宜在RTP包头中加入汇入源标识符（CSRC），也不宜添加RTP包头扩展（见RFC 3550中5.1），但是根据RFC 3551，接收器应允许CSRC标识符与包头扩展名的存在。

发送器和接收器均宜传输RFC 3550第6章规定的RTCP消息，且宜按照RFC 3551第2章（RTCP报告间隔）的建议分配RTCP带宽。

7 编码与成流

7.1 简介

编码描述音频信号数字化和构建为数据包序列并形成音频流的方法。

7.2 有效载荷格式和采样率

有效载荷格式定义音频数据样值的编码。本标准支持以下有效载荷格式：

——L16 16位线性编码格式，见RFC 3551中4.5.11；

——L24 24位线性编码格式，见RFC 3190中第4章。

所有设备应支持48kHz采样率。设备宜支持96kHz和44.1kHz采样率，见AES5—2008。

当设备在48kHz采样率下工作时：

a) 接收器应同时支持L16和L24编码；

b) 发送器应至少支持L16和L24中的一种。

当设备在96kHz采样率下工作时，发送器和接收器都应支持L24编码。

当设备在44.1kHz采样率下工作时，发送器和接收器都应支持L16编码。

尽管所有设备应支持48kHz采样率，但并不要求任何时候都能接受48kHz的音频流连接。例如，设备可以设计一个全局的采样率配置，仅当用户选择全局48kHz模式时，设备才接受48kHz的音频流连接。本标准不要求设备同时支持多种采样率。

设备可以支持采样率为96kHz的L16编码，及其他由RFC 3551和RFC 3190定义的音频格式的组合，但超出上述定义的其他采样率和有效载荷的组合不在本标准讨论的范围。

注1：音频流所用的格式和采样率，由RTP数据包头的有效载荷类型字段（见RFC 3550中5.1）和流的描述信息共同定义（见第8章）。

注2：RFC 3190中提到“emphasis”（预加重）参数，一般被认为是一个可追溯到二十世纪八十年代的遗留问题。预加重不再普遍用于数字音频信号，且不希望被用于可互操作的网络流中。

7.3 包时间

7.3.1 概述

包时间是指媒体数据包中包含的媒体数据的实时播放持续时间。如果给定采样率和包时间，就能计算出每个媒体IP数据包中包含的媒体信号采样点个数。

较短的包时间产生的时延较小，但这样会导致越多的网络开销和更高的数据包速率，可能会使某些设备或网络超负荷。越长的包时间传输延时越高，而且需要的缓存越大，这在内存有限的设备上可能无法实现。

包时间是发送器决定的，在会话描述（见8.2）中指定，并通过连接管理（见第10章）协商。在一个会话周期内，发送器不应更改包时间。接收器可默认在一个会话周期内，包时间是固定的。为了能够在标准的RTP实现中实现互操作，接收器不宜依赖于包时间描述信息的有无或精度。接收器宜具备根据所接收到的数据包中的时间戳来确定包时间的能力。

本标准讨论的互操作性，是在要求设备支持1ms包时间的条件下进行的（见7.3.2）。鼓励使用7.3.3中给出的其他包时间提高互操作性。

设备的产品文档中应标明发送和接收方向所支持的包时间。

7.3.2 必要的包时间

1ms的包时间为音频和网络设备的互操作性和兼容性提供了最大可能。

在采样率为48kHz或44.1kHz时，发送器应能够向每个音频数据包中加载48个采样的音频数据；在采样率为96kHz时，发送器能为每个音频数据包加载96个采样的音频数据。在采样率为48kHz或44.1kHz时，接收器应能够接收并解码含有48个采样的音频数据的数据包；采样率为96kHz时，接收器可以接收并解码含有96个采样的音频数据的数据包。

尽管所有设备应支持上述规定的包时间的相关要求，但并不要求它任何时候都能接受这些音频流连接。例如，设备可以设计一个全局的包时间配置，仅当用户选择设备工作在某一包时间模式时，设备才接受该模式下的音频流连接。本标准不要求设备同时支持多种包时间。

7.3.3 推荐的包时间

为了增强应用间的互操作性，发送器和接收器宜支持表2中所列的一种或多种包时间。

发送器和接收器可支持表2以外的其他包时间。包时间上限受限于网络MTU（见6.4）。

表2 必要的包时间和推荐的包时间

包时间	包采样数 (48kHz)	包采样数 (96kHz)	包采样数 (44.1kHz)	备注
125 μs	6	12	6	兼容 A 类 AVB 网络传输(有关 AVB 网络传输参见附录 D)
250 μs	12	24	12	高性能、低延时操作；可与 A 类 AVB 网络互操作，以及兼容 B 类 AVB 网络传输（有关 AVB 网络传输参见附录 D）
333 μs	16	32	16	高效的低延时操作
1 ms	48	96	48	本标准规定的必要的包时间（见 7.3.2）
4 ms	192	n. a.	192	适用于需要与 EBU Tech 3326 互操作、或在更广的区域内传输、或在 QoS 能力有限的网络上传输

注：EBU Tech 3326中未讨论96kHz。受MTU的限制（见6.4），当包时间为4ms时，一个IP数据包只能传输96kHz采样率音频流的一个通道。

7.4 流通道数

每个音频流允许的最大通道数受限于包时间、编码格式和6.4中所述网络MTU。表3给出了每个流允许的最大通道数示例。

接收器应支持接收包含1个~8个通道的音频流。接收器可支持多于8个通道的音频流。发送器应至少提供一个包含不大于8个通道的音频流。

表3 每个流允许的最大通道数示例

编码格式，采样率	包时间	每个流的最大通道数
L24, 48kHz	125 μs	80
L16, 48kHz	250 μs	60
L24, 48kHz	250 μs	40
L24, 48kHz	333 μs - 1/3 μs	30
L24, 96kHz	250 μs	20
L24, 48kHz	1ms	10
L24, 48kHz	4ms	2

注：尽管在一个流中编组多个通道能提高网络效率和处理效率，但是建议主要基于应用服务进行编组。例如将

立体声或环绕立体声等内容相关的通道进行编组是较好的选择。不建议为了减少网络开销，而将发送到不同接收器的没有相关性的通道进行编组，这会使媒体路由配置变得复杂。

7.5 链路偏移量

链路偏移量是指媒体信号通过媒体网络的时间，其定义为音频信号进入发送器（即入口时间）和离开接收器（即出口时间）之间的时间差。

入口时间指信号进入发送器网络系统的时刻。RTP数据包在“timestamp”字段（见RFC 3550中5.1）中将该时刻标记为初始时间戳。出口时间指信号离开接收器网络系统的时刻。链路偏移量即入口时间和出口时间的差值。图1给出了链路偏移量、信号入口和出口时间参考点的示意。

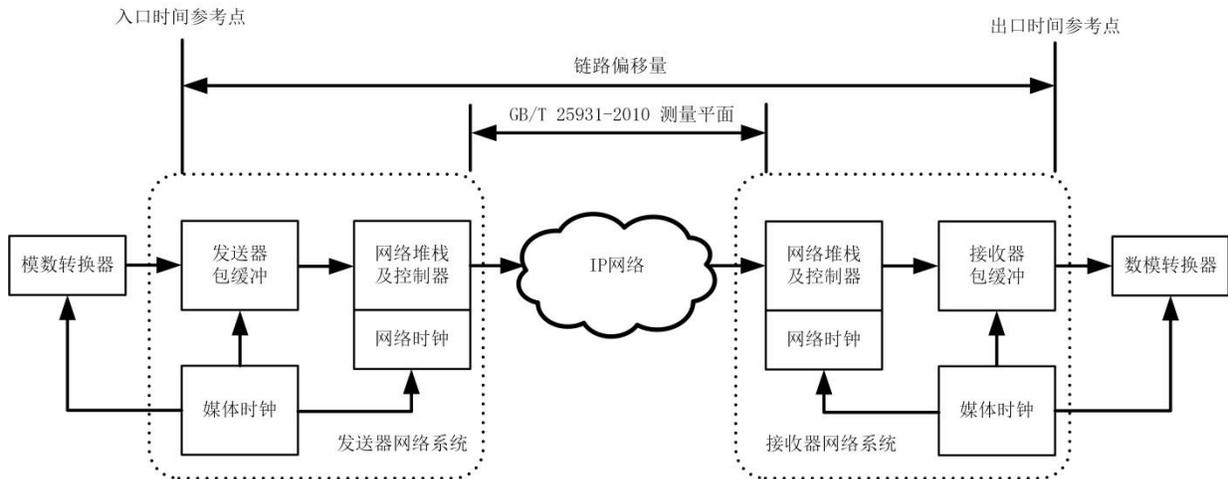


图1 链路偏移量、信号入口和出口时间参考点的图例

链路偏移量在接收器端决定，并受多种因素的影响，包括包时间、传播时间和队列延时，设备端的数据包处理时间和接收器的缓冲时间。接收器宜尽力维持链路偏移量恒定不变。同时，网络状态的意外变化可能会改变接收器的缓冲时间，从而导致链路偏移量的变化。如出现链路偏移量和链路偏移量变化，宜通知管理实体。

注1：链路偏移量的最小值为包时间（见 7.3）加网络转发时间。在点对点的千兆以太网连接中最小数据包的转发时间可能低至 $0.5 \mu\text{s}$ 。在实际网络环境下链路偏移量的最小值将接近包时间的两倍以上。

注2：下一步可能会建立链路偏移量的管理机制，该机制将要求增加额外的缓存，并将链路偏移量报告给网络中的中央延时管理服务器。接收器根据延时管理服务器的命令调整其链路偏移量。见 IETF draft-ietf-avtcore-idms。

7.6 发送器定时和接收器缓存

接收器需要使用缓存来吸收由打包、网络递送以及发送器接收器中网络堆栈和控制器所产生的抖动。接收器的缓冲区应能容纳在“链路偏移量时间减去发送器与接收器之间的最小传输时间”内收到的媒体数据。发送器的数据包缓冲区应能容纳“包时间加由发送器网络堆栈和控制器所产生的任何时间变化”内所产生的媒体数据。缓冲时间太短，数据可能无法及时送达进行播放，从而导致音频丢失。缓冲时间越长，音频信号的鲁棒性越高，但会产生附加延时。

接收器应至少具备包时间的3倍缓冲能力。接收器的缓冲容量宜为包时间的20倍或20ms，取较小者。

发送器以标称的包时间间隔发送与流关联的数据包。发送器宜在标称的传输时间内进行传输，其波动范围小于等于1个包时间。发送器应在标称的传输定时内进行传输，其波动范围小于等于17个包时间或17ms，取较小者。

上述要求旨在保证从硬件到桌面操作系统应用程序可实现。附加的缓存以及更精确的传输时间会提高鲁棒性及互操作性。

7.7 组播

流数据组播可以有效地从一点到多点分发音频流。组播也是一种简化连接管理（见10.3）的重要方法，在组播情况下，发送器发送一个组播流，接收器只需发现并监听正在接收的流。

接收器应能够接收组播和单播流。

尽管RTP支持多点到多点连接，但是本标准假设每个组播地址只对应一台发送设备。

组播应使用管理的组播地址，地址范围为239.0.0.0至239.255.255.255。此范围内的地址可由网络管理员进一步细分，也可将一个子集分配给媒体网络使用。

发送器应通过管理接口配置特定流的组播地址。本标准认为每个流都将被指定一个唯一的在此范围内的组播地址。管理接口属性和分配机制不在本标准的规定范围。

注：有关IP组播范围管理更多信息见RFC 2365-IP组播范围管理。

8 会话描述

8.1 概述

发现服务（见第9章）和连接管理（见第10章）使用会话描述来表示流的关键信息，包括网络寻址（见第6章）、编码格式（见第7章）和原始信息。

在连接管理中，应使用RFC 4566定义的会话描述协议（SDP）来标识会话。实现互操作性需要借助8.2、8.3、8.4、8.5中讨论的额外的SDP要求和建议。

8.2 包时间

包时间由RFC 4566中定义的两个SDP属性描述。

a=ptime:<milliseconds>[.<milliseconds decimal>]

a=maxptime:<milliseconds>[.<milliseconds decimal>]

由信令中的包时间乘以采样频率再四舍五入后得到的整数值就是每个媒体数据包中包含的样点个数。描述信息中携带的包时间误差应小于半个采样周期，这样计算出的样点个数四舍五入后得到的是期望的整数。在很多情况下，需要把<milliseconds decimal>字段放到描述符中。在不需要用<milliseconds decimal>来精确地表示包时间的地方，则应在信令交互过程中把它忽略掉。

表4给出了与7.3中的包时间对应的描述值示例。描述符中的包时间应以十进制表示。表4中未能列举的值也应得到相同的理解。

表4 信令中的包时间示例

包时间	采样率		
	48kHz	96kHz	44.1kHz
125μs	0.12	0.12	0.13
250μs	0.25	0.25	0.27
333μs	0.33	0.33	0.36
1ms	1	1	1.09
4ms	4	4	4.35

描述信息应包含一个ptime属性来表示所需的包时间。如果支持多个包时间，应使用maxptime属性来表示允许的最大包时间。有关ptime和maxptime可用的<milliseconds>属性值，见表2。

此描述信息的要求表明包时间越短越好。如果需要提供其他优先顺序可选的包时间，可能要用到RFC 5939的能力协商属性，以列出可选的包时间及其优先顺序。

如果支持的包时间包含了表2中规定的标准包时间的其中两个或多个，描述信息宜使用RFC 5939的能力协商属性来列举所支持的包时间及其优先顺序。

注：非整数毫秒描述信息可能无法被未遵循本标准的连接管理系统正确识别。因此，当尝试连接此类系统时，描述信息需要使用整数的<millisecond>值。

8.3 时钟源

*ts-refclk*属性值指定了流所使用的网络时钟参考。*ts-refclk*支持三种版本的PTP规范以及其他的时钟源。应为每个流的网络时钟源指定一个或多个RFC 7273中规定的*ts-refclk*属性值。

以下示例说明了*ts-refclk*属性在本标准支持的同步选项中的用法。

示例1：使用4.2或4.3中讨论的IEEE1588—2008 (GB/T 25931—2010)网络时钟，且本例中的最高级时钟标识符(GMID)为39-A7-94-FF-FE-07-CB-D0，PTP域值为0：

a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:0

示例2：使用4.4中描述的IEEE 802.1AS—2011网络时钟，且本例中的GMID为39-A7-94-FF-FE-07-CB-D0：

a=ts-refclk:ptp=IEEE802.1AS-2011:39-A7-94-FF-FE-07-CB-D0

尽管RFC 7273中提到了PTP域值是可选的，但在本标准中，符合GB/T 25931—2010要求的RTP流信令应标明GMID和PTP域值。

IEEE 802.1AS—2011总是使用域值0，因此不需要标明该值。两台同步到IEEE 802.1AS—2011时钟的设备默认使用了相同的域值。

当与发送器使用相同的GMID时钟参考时，接收器会尝试连接发送器。当两者的时钟参考的PTP域值不同时，接收器不宜尝试连接发送器。

当出现PTP域值匹配但GMID失配时，可能的原因有：网络处于非稳态、发送器与接收器之间缺乏统一的时钟参考、不同的最高级时钟指向同一可溯源时钟（例如GPS）。在这种情形下，接收器可以尝试连接发送器，但也宜做好同步可能会失败的准备。

根据RFC 7273，发送器可指定多个等效的时钟源。接收器宜评估所有被指定的时钟源，并根据本章的建议尝试连接发送器。

发送器和接收器宜在传输过程中监测同步状态是否有变化。如果监测到变化，发送器宜更新其时钟源描述信息。

当同步状态出现变化，或者接收到发送器发来的更新的描述信息时，接收器宜根据本章的建议重新评估其继续工作的可能。在接收过程中，当监测到同步失配时，不要求接收器中断接收过程，但其宜做好同步可能会失败的准备。

8.4 RTP 时钟和媒体时钟

应使用RFC 7273中5.2规定的*a=mediaclock:direct=<offset>*属性描述每个流的媒体时钟与RTP时钟之间的关系。描述信息中应包含*offset*属性值。应为每个被描述的流提供*mediaclock*属性值。

注1：RFC 7273 允许在会话、媒体或音源层面指定 *mediaclock* 属性值，如 RFC 7273 中 5.4 所述。

注2：媒体时钟与网络时钟的关系是固定的，见第5章。

*mediaclock*属性支持众多媒体时钟场景，以下示例说明本标准如何使用该属性值。

示例：媒体时钟描述信息 - 按媒体时钟历元计算，如果RTP时间戳的值为1810024580，则：

a=mediaclock:direct=1810024580

8.5 载荷类型

为了指定可互操作的编码格式（见表2），需要分配动态的有效载荷类型及关联的*rtpmap*属性值。在RFC 3551第6章的表4中，这些格式不会以静态的有效载荷类型的方式被调用。接收器不应认为载荷类型值和载荷类型的关系是固定的。发送器逐个流地使用*rtpmap*属性来更新载荷类型值和载荷类型的关系。

8.6 描述信息示例

8.6.1 组播会话描述信息示例

以下为一段简单的SDP描述信息，用以描述一个以组播流的方式传输8通道、24位、48kHz采样频率的音频流。该音频流的包时间为1ms。

```
v=0
o=- 1311738121 1311738121 IN IP4 192.168.1.1
s=Stage left I/O
c=IN IP4 239.0.0.1/32
t=0 0
m=audio 5004 RTP/AVP 96
i=Channels 1-8
a=rtpmap:96 L24/48000/8
a=recvonly
a=ptime:1
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:domain-nmbr=0
a=mediaclk:direct=963214424
```

8.6.2 单播会话描述信息示例

以下为一段简单的SDP描述信息，用以描述一个以单播流的方式传输8通道、24位、48kHz采样频率的音频流。该音频流的包时间为250μs。

```
v=0
o=audio 1311738121 1311738121 IN IP4 192.168.1.1
s=Stage left I/O
c=IN IP4 192.168.1.1
t=0 0
m=audio 5004 RTP/AVP 96
i=Channels 1-8
a=rtpmap:96 L24/48000/8
a=sendonly
a=ptime:0.250
a=ts-refclk:ptp=IEEE1588-2008:39-A7-94-FF-FE-07-CB-D0:domain-nmbr=0
a=mediaclk:direct=2216659908
```

9 发现服务

发现服务是一项网络服务，可让成员设备建立成员列表或可用的会话列表。此列表能协助用户进行连接管理。连接管理需要SIP URI（见10.2）或SDP描述信息（见第8章）。可通过Bonjour、SAP、静态配置或其他方式实现该功能。

发现服务不是必需的，设备可实现一个或多个发现服务，如Bonjour、SAP等。

有关发现服务系统的详细信息参见附录E。

10 连接管理

10.1 概述

连接管理是用于在一个发送器和一个或多个接收器之间建立一个或多个媒体流的过程与协议。

10.2 单播连接

单播流的连接管理应通过RFC 3261中定义的会话发起协议（SIP）来实现。SIP广泛应用于IP电话，也是EBU Tech 3326使用的连接管理协议。

根据7.7，所有接收器应支持单播（和组播）流，因此所有接收器应支持SIP。

10.2.1 SIP URI

在SIP应用环境下，音频设备是具有关联SIP URI的SIP用户代理。参照其他用户代理的SIP URI，SIP支持定位并连接到用户代理。支持连接的SIP URI可通过发现服务（见第9章）或其他方式（如：静态配置、专用目录服务等）获取。

在本标准中，SIP URI应使用*sip:* URI的表示方式。使用了*sips:* URI格式，则表明宜通过TLS进行安全连接管理。安全连接管理并不意味着流传输是安全的。流传输安全性属于连接管理进程的一部分。虽然RFC 3261强烈建议用户代理使用TLS，但由于本标准对于安全媒体流的互操作性（见6.4）未作规定，因此不要求支持TLS。

10.2.2 服务器模式和无服务器模式

SIP通常在SIP服务器协助和参与的情况下使用。不同类型的服务器执行不同的SIP网络任务。服务器可位于网络中终端站点可访问到的任意位置。借助服务器可以创建灵活且可扩展的连接管理系统。

无服务器模式在没有使用服务器的情况下，用户代理之间使用直接的点对点方式进行连接管理。无服务器模式适用于适度规模的场景。在此场景下，由于规模有限，服务器的作用较小，并且过度部署SIP服务器会增加系统复杂度。

为了执行点对点连接管理，呼叫方应通过一些方法来确定被呼叫方的网络联系信息（即主机名称或IP地址）。这些信息可通过发现服务（见第9章）、手动配置或更高层的协议来获取。在点对点连接管理过程中，所有SIP消息将直接发送至目标设备，而不是服务器。若设备遵循本标准，应响应此类请求。

对无服务器模式的支持并不表示服务器无需满足标准SIP运行环境的要求。需特别指出的是，设备仍需尝试发现SIP注册服务器并进行注册，且响应来自服务器的消息。

10.2.3 User-Agent 字段

SIP协议中的*User-Agent*包头字段用于传递终端站点的相关信息，管理员能借助此信息提高连接管理效率，并处理特定的应用问题。RFC 2616中14.43定义了*User-Agent*数据的格式。设备宜在*REGISTER*和*INVITE*消息中包含*User-Agent*包头字段。

注：有关SIP环境下*User-Agent*的安全问题，见RFC 3261中20.41。

10.2.4 格式协商

应使用RFC 3264中规定的提议/应答模型来协商网络连接的编码格式。本标准支持的编码格式，见7.2。

10.2.5 包时间协商

提议/应答模型未解决属性协商问题，例如包时间。“提议”支持多种包时间，这些包时间可以使用*ptime*和*maxptime*属性来指定，如8.2所述，“应答”应默认支持所接收到的*ptime*和*maxptime*对应的包时间。为了提高灵活性和可靠性，设备应尽可能满足RFC 5939中的能力协商条款。

10.3 组播连接

组播连接管理可不使用连接管理协议就能完成。在此简单的连接管理场景中，接收器不需要与发送器直接联系。

接收器通过发现服务（见第9章）或其他方式获取所需连接的会话描述。当接收器探测到需要连接的流时，能使用IGMP向网络发送可接收确认并开始接收此流。

附 录 A
(规范性附录)
媒体类别

A.1 概述

GB/T 25931—2010规定的精确时间协议旨在满足一系列同步应用。通过定义和利用类别，可以满足不同应用的不同需求。类别由应用服务的一系列运行参数和协议组件组成。

本附录所定义的媒体类别可满足媒体网络的需求。需特别说明的是，类别使缩短启动时间、提高时间精度，并兼容标准IP网络设备成为可能。

随着GB/T 25931—2010生态系统的发展，可预料到将会有更多的类别可用，其中一部分类别将适用于媒体网络。为应对这种情况，设备可扩展或更新本附录定义的媒体类别。

A.2 媒体类别描述信息

在本标准中，媒体类别指定了媒体传输同步的属性及选项。该媒体类别符合GB/T 25931—2010协议规范要求，其基础架构和部分内容摘自GB/T 25931—2010中附录J。

媒体类别与GB/T 25931—2010中J.3的延时请求响应缺省PTP类别的不同之处有以下几个方面：

- 标识信息表明了媒体类别源于 AES；
- 当使用标准（GB/T 25931—2010）网络设备时，为了优化启动时间和精度，本附录减小了 *portDS.logSyncInterval* 和 *portDS.logMinDelayReqInterval* 属性值；
- 唯一支持的消息封装方式是 UDP/IPv4；
- 时钟的物理要求与 AES11—2009(R2014) 兼容；
- clockClass 标识值增加了 AES11—2009(R2014) 数字音频参考信号（DARS）精度级别，表 A.1 给出了媒体类别的 clockClass 值；
- 除请求响应机制外，还建议使用点到点的延时机制。

表A.1 媒体类别属性值

属性	取值
<i>defaultDS.domainNumber</i>	缺省初始化值应为 0，可取值范围为 0~255。
<i>portDS.logAnnounceInterval</i>	缺省初始化值应为 1，可取值范围为 0~4。
<i>portDS.logSyncInterval</i>	缺省初始化值应为-3，可取值范围为-4~1。
<i>portDS.logMinDelayReqInterval</i>	缺省初始化值应为 0，可取值范围为 -3 ~ 5 或 <i>portDS.logSyncInterval</i> 至 <i>portDS.logSyncInterval</i> +5，取两者之中更严格的值。
<i>portDS.logMinPdelayReqInterval</i>	缺省初始化值应为 0，可取值范围为 0~5。
<i>portDS.announceReceiptTimeout</i>	缺省初始化值应为 3，可取值范围为 2~10。
<i>defaultDS.priority1</i>	缺省初始化值应为 128。
<i>defaultDS.clockClass</i>	相较于 GB/T 25931—2010 中表 5，表 A.2 指定了更多值，以支持 AES11—2009(R2014) 时钟的物理特性，见 A.3.4。
<i>defaultDS.priority2</i>	缺省初始化值应为 128。
<i>defaultDS.slaveOnly</i>	如果该参数是可配置的，则缺省值应为 FALSE。
<i>transparentClockdefaultDS.primaryDomain</i>	缺省初始化值应为 0。
τ (见 GB/T 25931—2010 中 7.6.3.2)	缺省初始化值应为 1.0s。

对于表A.1中每个定义的范围，制造商可自行扩展。表A.2给出了表A.1中*defaultDS.clockClass*的属性值。

表A.2 媒体类别的 *clockClass* 值

<i>clockClass</i> (十进制)	说明	时标	是否能为 从时钟	是否是媒 体类别特 有
6	已同步到主参考时间源（如 GPS）的时钟	PTP	否	否
7	已指定为 <i>clockClass</i> 6 但已失去同步到主参考时间源的能力，且处于保持状态并符合保持规范的时钟	PTP	否	否
13	已同步到外部媒体时间源的时钟	ARB	否	否
14	已指定为 <i>clockClass</i> 13 但已失去同步到外部媒体时钟源的能力，且处于保持状态并符合保持规范的时钟	ARB	否	否
52	不符合保持规范的 <i>clockClass</i> 7 时钟的降级可选 A	PTP	否	否
58	不符合保持规范的 <i>clockClass</i> 14 时钟的降级可选 A	ARB	否	否
150	频率同步于精度为 $\pm 1\text{ppm}$ （例如，AES11—2009 (R2014) 的一级精度 DARS）的参考源、时间同步于主参考时间源的时钟	PTP	是	是
158	频率同步于精度为 $\pm 10\text{ppm}$ （例如，AES11—2009 (R2014) 的二级精度 DARS）的参考源、时间同步于主参考时间源的时钟	PTP	是	是
166	未指定容限的、先前已与主参考时间源同步的时钟	PTP	是	是
187	不符合保持规范的 <i>clockClass</i> 7 时钟的降级可选 B	PTP	是	否
193	不符合保持规范的 <i>clockClass</i> 14 时钟的降级可选 B	ARB	是	否
220	频率同步于精度为 $\pm 1\text{ppm}$ （例如，AES11—2009 (R2014) 的一级精度 DARS）的参考源、其时间先前未同步于主参考时间源的时钟	ARB	是	是
228	频率同步于精度为 $\pm 10\text{ppm}$ （例如，AES11—2009 (R2014) 的二级精度 DARS）的参考源、其时间先前未同步于主参考时间源的时钟	ARB	是	是
248	缺省值，如果没有其他的定义适用，则使用该 <i>clockClass</i> ；例如，未指定误差的时钟，其先前未与主参考时间源同步的时钟	ARB	是	否
255	仅作为从时钟的时钟	无	是	否

所有 *clockClass* 保持规范是 96kHz 字时钟周期的 $\pm 5\%$ 。

A.3 媒体类别

A.3.1 标识

PTP 类别（见 GB/T 25931—2010 中 19.3.3）的标识值如下所示：

```
PTP Profile:
PTP profile for media applications.
Version 1.0
Profile identifier: 00-0B-5E-00-01-00
```

此类别由 AES 标准委员会指定。

A.3.2 PTP 属性值

各节点应实施本 PTP 类别中的所有要求，即指定缺省值或选项。无需用户配置就能应用这些缺省值或选项，即这些值或选项是制造商提供的。表 A.1 给出了媒体类别的属性值。

A.3.3 PTP 选项

设备可使用 GB/T 25931—2010 第 17 章中的任意选项。设备可使用 GB/T 25931—2010 中 16.1 定义的单播协商。除非特地被管理规程激活，这些选项都应是非活动状态。

节点管理应实现 GB/T 25931—2010 中规定的管理报文机制。

最佳主时钟算法应为 GB/T 25931—2010 中 9.3.2 规定的算法。

缺省路径延时测量机制应为GB/T 25931—2010规定的延时请求-响应机制，也宜实现对等延时机制。

注：每条链路只允许使用一种机制。边界时钟宜在使用不同路径延时机制的链路间使用。

A.3.4 时钟物理要求

时钟应满足AES11—2009 (R2014) 中5.2二级精度DARS的要求。也可使用一级精度时钟规范，这种情况下应通过A.2.2的*defaultDS_clockClass*将此信息告知其他网络链接者。

时钟应支持GB/T 25931—2010缺省类别。如果使用缺省类别，应满足与其关联的物理要求。

注：缺省类别要求有 $\pm 100\text{ppm}$ 的可调整范围，推荐具有 $\pm 250\text{ppm}$ 的可调整范围。

尽管单个时钟满足了AES11—2009 (R2014) 要求，但并不能保证时钟分发系统整体都满足AES11—2009 (R2014) 规范，因为网络性能也会影响分发的时钟质量。

附 录 B

（资料性附录）

IEEE 802.1AS 时钟域接口

B.1 概述

IEEE 802.1AS—2011为在增强型以太网网络中使用GB/T 25931—2010定义了一个架构，见IEEE 802.1BA—2011。此类型的网络通常称为音视频桥接（AVB）。如4.4所述，AVB网络可使用IEEE 802.1AS—2011同步类别替代第4章所述的类别（有关GB/T 25931—2010缺省的类别和媒体类别的详细信息见附录A）。很多情况下，可能需要对接两种类别，使其在异构网络中共存。

根据GB/T 25931—2010和IEEE 802.1AS—2011，PTP时钟被指定为边界时钟（BC）、普通时钟（OC）或透明时钟（TC），尽管802.1AS透明时钟也具备一些边界时钟的功能。设备可实现一或多个上述功能。PTP时钟拥有一或多个端口。OC至少拥有一个端口。TC和BC应有两个或以上端口。BC和OC上的端口工作状态可被设定为主或从。时钟的每一个端口与一个GB/T 25931—2010类别相关联。由于透明特性，TC能与多个时钟域连接并与相应的类别关联。

B.2 边界时钟接口

BC的所有端口通常使用相同的类别，但不同的端口也可关联不同的类别，后者在使用不同类别的GB/T 25931—2010网络的各部分之间创建了接口。适用于媒体网络的类别包括：本标准采用GB/T 25931—2010中附录J定义的缺省类别、附录A中定义的媒体类别以及IEEE 802.1AS—2011定义的AVB网络类别。

当BC被用于不同网段间同步时，所有网段应采用最佳主时钟算法（BMCA），并且应使用同一版本的GB/T 25931—2010协议和相同的域标识符。这样做利于将运行缺省类别的网段与运行媒体类别的网段进行连接。由于IEEE 802.1AS—2011使用以太网专用BMCA，因此将GB/T 25931—2010网段与IEEE 802.1AS—2011网段进行连接时，此方法不适用。

以此方式使时钟互连时，不会有时钟源信令的问题。通过BC接口可简单地获得更大的PTP域。BC接口两边采用相同的同步标识（见第8章）。

B.3 普通时钟接口

多端口OC能在不同端口运行不同类别。多端口OC能用于同步多个网段，但限制条件如下：OC最多只能使一个端口处于从模式；时钟分配时，应确保OC不作为多个网段的从时钟；所有网络使用同一全局同步源。

因此，多端口OC能被多个PTP网络同时用作最高级时钟，与PTP网络类型、版本和类别无关。此功能受限于对最高级时钟的配置要求，事实上目前未定义此类互连的时钟源信令。

B.4 可溯源参考

RFC 7273支持的任意时钟源能指定为“可溯源”。这意味着时钟是同步到TAI全球参考时钟。理论上，任何可溯源时钟可用来替代其他可溯源时钟。实践中，这种方式的意义无从考证。PTP网络的同步精度是可控的，而两个可溯源参考时钟之间的差异是无法指定的，且一般未知。一个特例是，当两个参

考时钟均为GPS时，由于GPS精度是100ns左右，与PTP的精度相当。使用GPS最高级时钟的多个PTP网络能像同源时钟一样工作。最高级时钟广播*Announce*消息或管理消息均包含*clockClass*信息，其同步性能信息比RFC 7273更丰富。

B.5 AVB 网络用作边界时钟

由于IEEE 802.1AS—2011时钟域中的高精度时钟分发，因此将整个IEEE 802.1AS—2011域作为分布式边界时钟用于任何其他PTP域成为可能。虽然此用法不必将IEEE 802.1AS—2011域同步到任何穿过它的PTP域，但展示了另一种AVB集成方法，该方法提高了互操作性。

附 录 C

（资料性附录）

网络 QoS 配置建议

C.1 概述

支持高质量媒体流的网络应提供这些服务所需的QoS。本标准考虑到QoS的需求，在6.3中指定了如何在网络中标记网络流量。虽然网络特性不在本标准范围之内，有关如何根据IETF区分服务规范在IP网络中实施QoS，本附录提供了通用的配置指南。

C.2 区分服务网络配置

C.2.1 区分服务

区分服务是对网络流量进行分级并以不同方式处理的框架。区分服务被认为是一个粗颗粒度的QoS体系结构，因为它是根据类别对流量进行分级，而不是对流量个体进行操作。

逐跳行为（PHB）是执行区分服务的基础。在每台路由器或交换机中，会对流量进行分级并根据此分级重新传输。当出现网络拥塞时，拥有更高优先级的流量会被立即重传，而低优先级的流量会在网络设备中缓冲等待，也可能被全部丢弃。相对于未分级的网络流量（如流量个体），在分级的网络流量上使用PHB可获得简单、可扩展的QoS解决方案。

以太网设备为每个端口分配了多个出口队列，通常支持PHB。不同级别的流量会单独排队。当出口端口可以传输数据包时，由选择算法选择一个队列并从其中选出一个数据包。网络设备中会使用多个选择算法。按理说，最简单的选择算法是严格优先级算法。根据此算法，会从拥有最高优先级的非空队列中选择最早的数据包进行传输。因此高优先级的流量会拥有低的延时，而较低优先级的流量可能会长时间排队，甚至导致数据丢失，此现象称为饥饿状态。通过加权轮询调度算法和保障最小带宽算法等更为复杂的选择算法可以缓解饥饿状态，但是拥有最高优先级的流量延时会变高。

在支持区分服务的网络中，是通过每个IP包头中的一个6位区分服务代码点（DSCP）值对流量级别进行识别的。这些值由生成流量的终端站点进行分配，见6.3。

本标准对网络中的媒体流提供三种流量级别。6.3为每种级别指定的DSCP建议值被用于按区分服务相关的请求注释（RFC）文件配置的网络。

注：并不是所有网络都是按照区分服务规范配置的。由于这些RFC文件允许较大的配置范围，因此允许用不同的方式在网络中配置和部署区分服务。

C.2.2 时钟

时钟流量是由低数据率（小于100kbps）、低频度（低于100数据包每秒）的UDP数据包（每包约100字节）构成。虽然时钟流量对数据包丢失的敏感度不高，但对延时非常敏感，尤其是延时变化，网络工程师称之为延时抖动。对于时钟流量而言，大的延时抖动会使网络间时钟传递的精度降低。

由于在某些网络中，网络管理和路由控制流量的优先级要高于其他应用流量，建议为GB/T 25931—2010时钟流量分配最高优先级。在大部分区分服务网络中，通过为时钟流量指定快速转发（EF）来实现。快速转发通常会进入严格优先级队列，此队列会先于其他队列传输。

注：相较于之前的网络应用（如文件传输、电子邮件传输、Web浏览），VoIP（Voice Over IP）是对性能要求更高

的网络应用，因此EF也通常用于VoIP流量中。由于VoIP通常使用20ms的包时间，对于时间的严格程度与本标准定义的媒体流最大差2个量级。但VoIP流量的小的数据包（100字节）与低数据率（10kbps每次呼叫）的特性使其产生的干扰很小。

配置了快速转发区分服务并愿意信任终端站点分级的网络会按原样接受流量。安全性考量更多的网络会基于地址和端口分配对时钟流量进行识别和分级，如表C.1所示。

表 C.1 时钟流量标识

流量	目的地址	协议	目的端口
时间敏感事件消息	224. 0. 1. 129	UDP	319

C.2.3 媒体流

媒体流的特点是高频度（每流每秒高达8000个数据包）和高数据率（每个音频通道1Mbps以上）。数据包大小取决于流中承载的音频通道数，范围能够从小于100字节（针对单通道流）到填满1500字节的以太网MTU。媒体数据包应以可靠的方式及时传输。媒体数据中的任意数据包丢失将表现为音频丢失。对于最高性能的应用，数据包的网络延时若超过250μs，可能因太晚到达目的地而无法使用，这将表现为音频丢失。音频丢失对可听度的影响取决于应用。越专业的应用对数据包丢失的容忍度越低。

6.3中建议发送器使用DSCP值34来标记AF41媒体流量。本章也允许根据不同流的QoS需求使用其他的DSCP值。

建议将媒体流作为高优先级的保证转发级。有关保证转发PHB的定义，参考RFC 2597。对媒体使用多个保证转发级，可使某些应用受益。保证转发的实现可能要求在网络建设中分配足够带宽，从而使流量如预期一样传输。如果用低于快速转发的第二优先级队列来实现保证转发，则可以实现最佳性能。通常更倾向于使用加权轮询调度选择算法来实现保证转发。

配置了AF41区分服务并愿意信任终端站点分级的网络会按原样接受流量。安全性考量更多的网络会基于地址和端口分配对媒体流量进行识别和分级，如表C.2所示。

表 C.2 媒体流量标识

流量	目的地	协议	目的端口
缺省的 RTP 媒体	单播或组播	UDP	5004
缺省的 RTCP 管理	单播或组播	UDP	5005
6.4 中提到的非缺省的 RTP 和 RTCP 流量	单播或组播	UDP	用户端口 (1024~65535)

注：有关某一具体设备对媒体流量的其他或备用端口分配的详细信息，参考产品文档或咨询制造商。

C.2.4 尽力转发

尽力转发（BE）流量是与本标准相关但未归类为时钟或媒体流量的流量，包括与发现服务和连接管理相关的消息流量。有关发现服务和连接管理的详细信息，见第9章和第10章。

终端站点通常使用DSCP值0来标记BE流量。因为所有网络都会尽力转发流量，未被归类的流量也被认为是BE流量，通常不需要其他特殊配置来容纳BE流量。

附 录 D

(资料性附录)

AVB 网络传输

D.1 概述

本标准定义如何使用IP传输来承载媒体数据。以太网是支持IP传输的通用网络。AVB可被认为是增强型以太网。IEEE有一套标准对AVB进行描述。其中与网络传输相关的标准有：

- a) IEEE 802.1AB——音视频桥接（AVB）系统；
- b) IEEE 802.1Q 的第 34 章——时间敏感流的转发和排队；
- c) IEEE 802.1Q 的第 35 章——流注册协议（SRP）。

AVB不但能改善媒体传输的性能，还能优化时钟分发。有关后者的详细信息，见附录B。

D.2 AVB网络传输

D.2.1 概述

AVB网络提供两种基本方法以传输具有互操作性的媒体流。

方法一：以AVB时间敏感流的方式传输具有互操作性的媒体流。该方法利用了AVB的增强特性，但存在局限，即媒体流只能在AVB网络域中传输，不能发送到非AVB网络，因此不兼容非AVB设备。

方法二：以“其他流量”的方式传输具有互操作性的媒体流。该方法具有较大的灵活性，可在AVB和非AVB网络间传输及连接非AVB设备。但此传输模式未使用AVB的QoS和注册服务，分配给非时间敏感流的网络带宽比例通常限制为25%。

D.2.2 以AVB时间敏感流的方式传输具有互操作性的媒体流

D.2.2.1 概述

符合本标准的AVB设备可选择使用AVB预留带宽和QoS传输媒体流，此方法具备以下优势：

- 预留带宽确保媒体流在网络中可靠传输；
- 与其他使用IP传输的AVB设备具有潜在兼容性；
- 优异的带宽和延时性能。

使用此方法的设备直接连接到支持AVB标准和协议的以太网交换机。在AVB网络中，有两类媒体流。A类优先级最高，在兼容IEEE 802.1BA的网络中拥有2ms的延时保证。B类优先级较低，但仍高于网络中的其他非流式数据，在兼容IEEE 802.1BA的网络中拥有50ms的延时保证。这两类流都可用于具有互操作性的媒体数据传输。

设备在将具有可互操作性的媒体数据作为AVB媒体流传输之前，先预留带宽。使用流预留协议实现预留带宽的方法，见IEEE 802.1Q的第35章。在预留过程中，发送器会指定流类别（A或B），以及在度量时间间隔内该类可以产生的最大流量值。A类度量时间间隔为125 μ s，B类为250 μ s。

由于AVB网络会根据以太网的目的地址来识别流，因此AVB流通常会使用组播目的地址。具有互操作性的媒体流传输宜使用7.7所述的组播IP地址。根据RFC 1112，由于多个IP组播地址会映射到同一以太网组播地址，因此为AVB时间敏感性流选择组播地址时，要关注7.7中所述的注意事项和考虑其他情况。

D.2.2.2 发送器行为

发送器（AVB术语中称为Talker）使用SRP在“*Talker Advertise*”消息向AVB网络提供如下信息：

<i>StreamID</i>	64位全球唯一标识符，由两部分组成：发送器的48位以太网MAC地址和发送器生成的16位唯一标识符
<i>Destination address</i>	用于流传输的以太网组播目的地址
<i>VLAN identifier</i>	所有AVB媒体数据包包含一个IEEE 802.1Q VLAN标记。此标记包含一个12位VLAN标识符。AVB流的缺省VLAN标识符为2
<i>MaxFrameSize</i>	媒体流数据包的最大值，见第7章 <i>MaxFrameSize</i> 包含IP包头，不考虑以太网开销 注： <i>MaxFrameSize</i> 即以太网帧最大有效载荷长度
<i>MaxIntervalFrames</i>	发送器在一个度量时间间隔内可传输的最大帧数 如7.2所述，由于允许的包时间大于或等于AVB度量时间间隔，此处始终为1
<i>Data Frame Priority</i>	3代表A类，2代表B类
<i>Rank</i>	1代表正常流量，0代表紧急流量

SRP分配的固定带宽数为： $MaxFrameSize \times MaxIntervalFrames$ （一个度量时间间隔内）。注意，由于具有互操作性的流的包时间大于等于AVB度量时间间隔，因此实际使用的带宽可能明显少于预留值。在设计AVB网络时要考虑这一点。

收到SRP的“*Listener Ready*”或“*Listener Ready Failed*”的答复后，发送器才能传输流数据包。发送器以第7章所规定的IP数据包格式传输流数据包。AVB网络使用组播过滤方式强化带宽分配，从而只支持对流数据使用组播目的地址（见IEEE 802.1Q中35.2.2.8.3）。在AVB网络中按照此方式传输的流数据包使用IP组播目的地址。IP组播目的地址将映射到以太网组播目的地址，根据RFC 1112的6.4节，地址范围为01:00:5e:00:00:00至01:00:5e:7f:ff:ff。

IP数据包经组帧后在以太网中传输。该以太网帧包含IEEE 802.1Q VLAN标签，该标签带有VLAN标识符和优先级代码点，且与相应SRP的*VLAN Identifier*和*Data Frame Priority*属性匹配。发送器对类内的流数据包的传输定时与可信因子(credit-based)整形器的定时一致，见IEEE 802.1Q中8.6.82。

D.2.2.3 接收器行为

接收器（对应的AVB术语为Listener）使用SRP以*Listener Declaration*的形式向网络提供的*StreamID*。如果流可用，接收器将接收“*Talker Advertise*”，然后发送MVRP成员请求，请求加入*Talker Advertise*中指定的VLAN。同时，流数据包开始到达，接收器宜使用与非AVB网络中相同的方法处理这些流数据包。

D.2.3 以“其他流量”的方式传输具有互操作性的媒体流

D.2.3.1 概述

尽管AVB网络为每一条链路上的媒体流量预留的带宽最高可达75%，但这个数值是可配置的，这样其他流量才有可能获得更大比例的带宽。AVB QoS算法（可信因子整形器）考虑了其他网络流量的传输需求，使其在每个度量时间间隔内能够在网络中传输。对于仅承载A类流的AVB网络，度量时间间隔为125

微秒；对于承载B类流或A类和B类流都承载的网络，度量时间间隔为250微秒。虽然其他流量的优先级低于AVB媒体流量，但其他流量可根据自身情况优化。这样就为可互操作性媒体流提供了一个机会，使其作为其他流量在AVB网络中传输。此方法具有以下优势：

- 无需实现 SRP；
- 无需实现可信因子整形器；
- 可以在 AVB 网络边界传输媒体数据，包括在独立的 AVB 网络域之间桥接；
- 无 AVB 网络设备时也可工作；
- 不会因度量时间间隔与包时间失配而造成带宽浪费；
- 对发送器和接收器的行为无特殊要求。

D. 2. 3. 2 发送器行为

对发送器的行为无特殊要求。在一些AVB网络中，发送器可能希望使用IEEE 802.1Q VLAN标签来传送QoS等级。AVB网络通常适应于以太网协议，可不检查IP包头中的DSCP字段。

VLAN标识符的值取决于网络中VLAN的配置。尽管并不是所有网络设备都支持0作为VLAN标识符，但本标准仍使用0可表示缺省的VLAN成员。有关VLAN标记中优先级字段的建议值，见表D. 1。

表 D. 1 根据 802. 1Q 将媒体流量标记为“其他流量”在 AVB 网络中传输的 PCP 建议值

流量	PCP
时钟	6
媒体	5
尽力转发	0

注：根据IEEE 802.1Q中6.9.4的规定，以PCP值2和3来标记的AVB媒体流量，在AVB域边界处将被重新映射为0。

D. 2. 3. 3 接收器行为

对接收器的行为无特别要求。接收器能接收带有IEEE 802.1Q标签的媒体数据包即可。

附录 E (资料性附录) 发现系统

E.1 概述

本标准不强制要求使用发现服务，但是第9章规定可通过SIP URI识别潜在的连接。SIP URI能通过发现系统或其他方式获得。

E.2 Bonjour

Bonjour是一套零配置网络技术，由Apple公司开发并作为公开文档发布。在Bonjour中，相关发现技术包括RFC 6762中规定的组播DNS（mDNS）和RFC 6763中规定的DNS服务发现（DNS-SD）。

如果这些技术用于发现服务，RTP会话宜由其SIP URI进行播发，见IETF *draft-lee-sip-dns-sd-uri*。

E.3 SAP

当使用SAP或类似机制向潜在接收器分发SDP描述信息时，可在组播流上实现简单的连接管理，见10.2。

如果使用SAP，则宜使用RFC 2974中的SAP第2版本。SAP消息的有效载荷中宜承载第8章定义的SDP描述信息。组播会话宜公告使用RFC 2974第3章中指定的目的地址；全网范围的组播会话宜公告使用224.2.127.254目的地址，管理域的组播会话均需公告使用该域的最高地址。

E.4 Axia发现协议

Axia发现协议运行于系统级的专属组播通道。处于这些组播的传输范围内且订阅了组播群组的每台设备都能收到所有其他设备的公告。

所有设备定期生成短的在线公告，并以更长的时间间隔发布描述性通告。描述性通告包含设备属性以及可传输的流清单和流属性。描述数据更新后，立即发布新的描述性通告。公告数据允许每位参与者创建网络中的其他参与者清单和可用的流清单，以协助连接管理。有关连接管理的详细信息，见第10章。

E.5 Wheatstone Wheatnet IP发现协议

Wheatnet IP发现协议运行于系统级的专属IP组播通道（“公告通道”）。处于这些组播的分发范围内且加入了组播组的设备都能收到其他设备的公告。

Wheatnet IP协议会从系统的节点中选出一个路由管理节点（*Route Master*）。新设备接入Wheatnet IP系统后，其将在公告通道发出进入公告（*Howdy*）。路由管理节点确认此消息并发起点名，系统中的每台设备依次用组播的方式报告各自的源和目的地。路由管理节点将为每台设备分配一段组播地址，用于

组播设备输出的流。一旦新设备收集了所有源和目的地信息，就会在系统中激活。路由管理节点会定期轮询每台设备，以确认所有设备仍在线。

任何设备都能作为路由管理节点使用，且每台设备都包含成为路由管理节点所需的所有相关信息。因此，若路由管理节点丢失，则路由管理节点将无缝迁移到另一个节点。

参 考 文 献

- [1] IEEE 802.1BA, Audio Video Bridging (AVB) Systems, Institute of Electrical and Electronics Engineers (IEEE), US.
- [2] RFC 894 – A Standard for the Transmission of IP Datagrams over Ethernet Networks, Internet Engineering Task Force
- [3] RFC 2597, Assured Forwarding PHB Group, Internet Engineering Task Force
- [4] RFC 3170, IP Multicast Applications: Challenges and Solutions, Internet Engineering Task Force 0
- [5] RFC 6762, Multicast DNS, Internet Engineering Task Force RFC 6763, DNS-Based Service Discovery, Internet Engineering Task Force draft-lee-sip-dns-sd-uri, SIP URI Service Discovery using DNS-SD, Internet Engineering Task Force.
<http://datatracker.ietf.org/doc/draft-lee-sip-dns-sd-uri/>
- [6] AES67—2015: “2015-09-21 printing AES standard for audio applications of networks - High-performance streaming audio-over-IP interoperability”, Audio Engineering Society, New York, NY., US.
- [7] Geoffrey M. Garner, Michel Ouellette and Michael Johas Teener (2012-09-27). “Using an IEEE 802.1AS Network as a Distributed IEEE 1588 Boundary, Ordinary, or Transparent Clock”. 2010 International IEEE Symposium on Precision Clock Synchronization for Measurement Control and Communication (ISPCS) (IEEE).
-

中 华 人 民 共 和 国
广 播 电 影 电 视 行 业 标 准
**高性能流化音频在 IP 网络上的
互操作性规范**
GY/T 304—2016

*

国家新闻出版广电总局广播电视规划院出版发行

责任编辑：王佳梅

查询网址：www.abp2003.cn

北京复兴门外大街二号

联系电话：(010) 86093424 86092923

邮政编码：100866

版权专有 不得翻印