

SRT协议特点及在直播中的应用

屈 晖

重庆广播电视集团(总台) 重庆市 渝北区 401147

摘要: 在远距离传输领域, 光纤传输, 卫星传输, 一直是传统信号传输主流方式。随着互联网的不断发展, 网络传输灵活、成本低、应用领域广泛的优势逐渐凸显。毋庸置疑, 现今存量最大的直播协议是RTMP, 但随着新技术的不断发展与使用场景的不断拓展, 继续使用RTMP会令人感到有些力不从心。可靠、低延时互联网传输协议—SRT协议应运而生。

关键词: 信号传输; 网络传输; 低延时SRT协议

前言: SRT是由Haivision和Wowza共同创建的SRT联盟所发起的互联网传输协议, 是一种较为常用的开源、免费应用规范, 具有较强的灵活性, 同时能够在不同制造商生产的产品之间工作。SRT作为一种开源低延迟视频传输协议, 目前受到十分广泛的欢迎。借助于这一传输技术, 可以实现普通互联网环境下的多地、多维传播, 安全可靠, 视频内容高清。

1 SRT 协议原理分析

SRT协议可以针对互联网环境下不可预测的场景提供安全、可靠的数据传输, 目前广泛应用在流媒体传输领域。理论上SRT可以传输任意类型的数据, 但是在进行传输的过程中, 实现了对实时音频、视频的优化, 所以目前仍然主要应用于跨公共互联网点对点传输当中。在实际工作中, 搭配使用不同厂商的SRT设备也能够实现高可靠、低延时的音视频传输, 这对于用户来说非常方

便和灵活。

1.1 SRT协议和UDP协议

SRT是由UDT(UDP-based Data Transfer)协议改进而来, SRT协议保留了UDT协议大多数的核心概念和机制, 同时引进了一些改进和增强功能, 其中主要包括针对实时音视频的流量控制、增强的拥塞控制、控制数据的修改、加密机制的改进。

图1展示了裸露的UDP协议在有损网络上的传输效果, 测试中利用软件Netem模拟了具有丢包、抖动、延迟的有损网络, 在源端将音视频信号进行信源编码, 编码后输出的TS流具有可变比特率(VBR)和固定的帧间隔, 但是在跨越网络后输出的码流特性已经完全改变, 固定的帧间隔也因为网络的抖动发生了改变, 实际上解码器很难从这样的码流中恢复出正常的音视频信号。

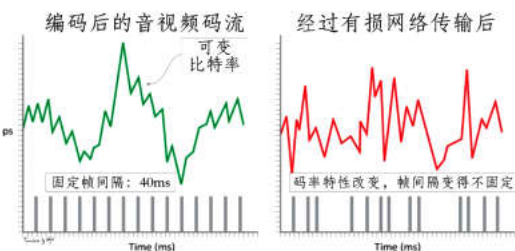


图1 UDP 协议在有损网络下的传输性能

由图2可以看到, SRT协议很好地克服了有损网络环境中的丢包和抖动, 在输出端很好地还原了输入的码流, 其借助于有效的差错控制以及精确的反馈信号、时间戳等, 根据延时量, 对缓冲区等进行定义。可以实现有损网络下的流媒体传输。在这当中, 差错控制主要采用的是自动重传请求机制(ARQ)。

2 传输链路的安全冗余量 (Secure-Margin)

SRT传输链路的可靠性配置, 需要结合实际情况, 进

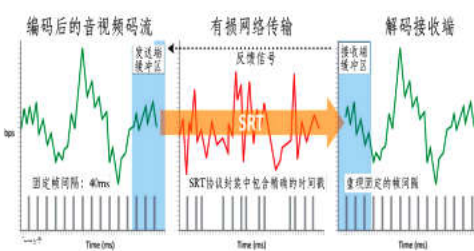


图2 SRT 协议在有损网络下的传输性能

行设计, 最大限度保证发送缓冲区的使用量与延时量相比更低, 切需要保证接收缓冲区的使用量需要大于零。这句话的原理将在下文详述, 它实际上对这一链路的两个临界崩溃点进行了描述, 而SRT链路安全冗余量表示了链路离临界崩溃点还有多少余量。

2.1 发送端缓冲区冗余量

发送缓冲区的作用是用来保存有可能需要重传的数据包, 即那些还没有收到肯定应答(ACK)的数据包, 如果发

送端收到了关于某个数据包的肯定应答,该数据包将从发送端缓冲区踢出。如果一直没接收到回复的ACK信

号,该数据也不能永远保存在发送端缓冲区中,SRT协议规定了保存的最长时间为延时量。



图3 发送端和接收端缓冲区(较好状态)

图4 发送端和接收端缓冲区(较差状态)

参考图3,并结合SRT的工作原理,可以将延时量视为一个同时在发送端与接收端滑动的窗口(从左到右)。在上图当中,1号到6号数据包,均已收到肯定应答(ACK),已从发送端缓冲区踢出,7号数据包仍然在缓冲区保存,等到接收方对于相关询问的回复,这一情况,属于一种较为良好的传输状态。发送端缓冲区的使用比例很小,大部分数据都已经被及时接收并且收到肯定应答。

参考图4,接收方并未接收或者获取到3号数据包,并向发送端回复了否定应答(NAK),因为诸多原因使得这一过程出现延误,导致3号数据包已经处在窗口最左侧,随着窗口的滑动下一步它将被从发送端缓冲区踢出。在此种情况下,倘若3号数据包在重新传输过程中,仍然出现丢失或则是延误的情况,就会直接导致整个数据包出现丢包问题,进一步影响解码端的图形。此外,还需要注意的是图中发送端缓冲区已经被填满,这说明传输当中处于丢包临界点。根据分析可以得出结论:发送端缓冲区被占用的比例越少,链路越安全,若发送端缓冲区被占满,链路则很有可能发生丢包。根据SRT的工作机制可知:数据保存在发送端缓冲区的最长时间为延时量(Latency),相反保存在发送端缓冲区的最短时间为链路的单次往返时延(RTT)。如果把发送端缓冲区被填满定为临界崩溃点,则发送端缓冲区冗余量SendBuffer-Margin的定义为:一段时间内缓冲区的最大空余量除以发送端缓冲区的最大值,则公式如下:

$$\text{发送端缓冲区冗余量} = (\text{延时量ms} - \text{发送端缓冲区最大占用量ms} / \text{延时量ms}) * 100\%$$

2.2 接收缓冲区冗余量

接收缓冲区其主要作用是对收到的数据包进行排序,根据SRT协议在SRT包头所记录的时间戳,结合实际需求,进行分配。排序具有两个方面的目的,第一是为了能够解码,第二是能够及时的找出达到数据包。并向

发送端返回否定应答(NAK),等待重传。在理想状态下,接收缓冲区应该保证其使用量略低于延迟量。倘若在这一过程中,占用量变为零,则表示在进行解码后,容易出现图像问题。

结合SRT的工作机制可以看出,接收缓冲区当中的数据存放时间与延时量相比,应当略低。如果将使用量为零视为其临界崩溃点,则缓冲区冗余量则可以定义为:某段时间内接收缓冲区最小占用量与最大值的比,具体公式如下:

$$\text{接收缓冲区冗余量} = (\text{接收端缓冲区最小占用量ms} / \text{延时量ms}) * 100\%$$

2.3 链路的安全冗余量

根据SRT的工作机制,发送端和接收端的缓冲区状态是相互影响的,并且状态较差一端的缓冲区冗余量实际上决定了链路的安全冗余量,那么定义链路的安全冗余量(Secure-Margin)等于发送端缓冲区冗余量(SendBuffer-Margin)和接收端缓冲区冗余量(ReceiveBuffer-Margin)两者的最小值。

根据SRT的机制和延时量的含义,实际上延时量和RTT一起定义了数据包能够被重传的次数,这个次数等于延时量除以RTT,我们可以依此进一步定义重传次数冗余量的概念:

$$\text{重传次数冗余量} = (\text{延时量ms} * \text{安全冗余量} \% / \text{链路往返时间RTTms})$$

3 配置 SRT 流的策略

3.1 测量网络链路的基础参数

在配置SRT流之前我们必须测量自身链路的可用带宽、丢包率、单次往返时延RTT,并注意观察这些指标是否在变动以及变动的范围。

3.2 设定延时量

在SRT协议中,延时量是一项十分重要的参数,其数值固定,可以设置的范围为80ms-8000ms。数值越大,则

SRT链路当中，对于差错的控制能力也就越强。不过，随着总延时不断增进，结合前期测试选择，可以在链路当中获得具有较强可靠性的延时机制，保证平衡。

3.3 设定带宽开销比例

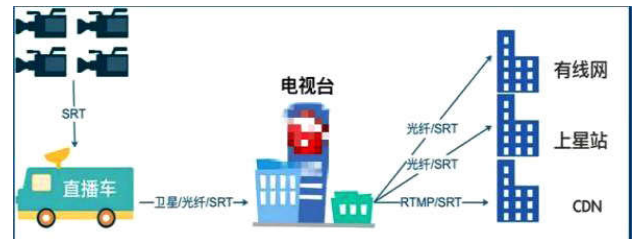
带宽开销(Bandwidth Overhead)是一个百分比参数，主要根据音视频码流比特率进行计算，其初始默认值为25%，一般情况下，不建议超过50%。其主要原因是如果设置过大，会导致用户分配额外宽带引起丢包。带宽开销比例与传输反向控制数据拥塞度密切相关，故而也影响链路差错控制能力。在时间当中，SRT协议中每丢失一个数据包，接收端就会消耗大约400bps的可用带宽。出现链路崩溃的情况下，缓冲区数据会逐渐渗透到解码器当中，基于数据需求对其进行填补，等待恢复。缓冲区数据消耗则需要由额外划定带宽开销弥补。处理完相关

在链路崩溃发生后，首先缓冲区数据会暂时填补解码器对数据的持续需求，紧接着链路恢复后，额外划定的带宽开销会弥补之前对缓冲区数据的消耗，在处理完这次突发事件后，整个SRT链路会恢复正常。

4 SRT 在电视直播中的应用

4.1 SRT在泛广电领域内的应用较早，在很多技术区域，已经在使用SRT技术。首先从拍摄机位的信号来说，

传输到直播车或演播中心，都可以使用SRT。另外制作好的节目传输到电视台，以前都是使用卫星或者光纤之类比较昂贵的传输方式，现在也可以通过公共互联网使用SRT技术来实现。在电视台播出之后传给各个分发商，这些分发商包括传统的有线网、上星站、无线覆盖或者直接对接CDN。对于CDN或者直播平台，我们之前是使用RTMP，但现在也有一些流媒体服务器的解决方案使用SRT作为上传推流的方式。



4.2 随着5G网络飞速发展，5G网络的覆盖率，稳定性越来越好。以5G网络为支撑，SRT传输协议配合相应硬件设备构建一个直播传输，回传链路。由于SRT的开源特性，它在工作中使用起来是非常方便的，和其他的单位或公司对接也相当便捷。因为他不会区分品牌、软件、硬件编码器等等。极大的提升传输效率和兼容、便捷性。



结束语

最近几年，随着信息技术的发展，在广电领域当中，SRT协议的应用愈发广泛。无论是远程制作、现场直播还是国际数据传输或者上行推流，都发挥重要作用。实际的SRT链路部署当中，需要进行大量测试，测试过程中，需要系统观察安全冗余量和缓冲区状态图，实现对断链路安全性的有效判断。另外还需要学会调整延时量Latency，保证安全冗余量的同时满足不同的直播场景对延迟的需求，不同直播场景有不同的设置策略。

参考文献：

[1] Haivision.SRT Protocol Technical Overview[M/OL].(2018-10) [2022-02].www3.haivision.com/srt-protocol-technical-overview.

[2] Haivision.SRT Open Source White Paper[M/OL].(2019-1) [2022-02].www3.haivision.com/srt-open-source-wp.

[3] SRT Alliance. SRT Deployment Guide,v1.1,Issue 01[M/OL].(2018-10) [2022-02].www3.haivision.com/srt-alliance-guide.