

T/CAICI

中国通信企业协会团体标准

T/CAICI XXXX—XXXX

5G 慢直播建设规范

Construction specification for 5G slow live broadcast

（征求意见稿）

在提交反馈意见时，请将您知道的相关专利连同支持性文件一并附上。

XXXX – XX – XX 发布

XXXX – XX – XX 实施

中国通信企业协会 发布

目 次

前 言 III

引 言 IV

1 范围 1

2 规范性引用文件 1

3 术语和定义 1

4 5G 慢直播技术架构..... 2

 4.1 设备终端 2

 4.2 网络环境 2

 4.3 视联网..... 2

 4.4 视频直播..... 2

 4.5 音视频处理..... 3

 4.6 内容审核..... 3

 4.7 内容分发..... 3

5 5G 慢直播设备要求..... 3

 5.1 摄像头终端 3

 5.1.1 图像要求 3

 5.1.2 三防要求 3

 5.1.3 定焦场景 3

 5.1.4 动态场景 4

 5.1.5 无源场景 4

 5.2 非摄像头终端 4

6 网络环境与设备接入 4

 6.1 网络环境 4

 6.2 设备接入管理 5

 6.2.1 总体原则 5

 6.2.2 系统设计要求 5

 6.2.3 整体方案 5

7 云服务 9

 7.1 流媒体规范 9

 7.2 直播加速 9

 7.3 云储存 9

8 AI 能力技术要求..... 9

 8.1 人脸审核 9

 8.1.1 总体要求 9

 8.1.2 人脸检测 9

 8.1.3 人脸对齐 9

 8.1.4 人脸特征提取 9

- 8.1.5 人脸特征检索 9
- 8.2 智能审核分类 9
- 8.3 AI 审核策略..... 10
- 8.4 视频 DNA 技术 10
- 9 播控安全技术要求 10
 - 9.1 流媒体内容编码及文件格式要求 10
 - 9.1.1 流媒体属性编码及文件格式要求 10
 - 9.1.2 慢直播传输协议要求 11
 - 9.2 直播播出运行要求 11
 - 9.2.1 慢直播源站 11
 - 9.2.2 直播分发 11
 - 9.2.3 实时预览 11
 - 9.2.4 慢直播媒体数字版权保护 11
 - 9.2.5 慢直播日志 11
 - 9.2.6 系统性能要求 11
 - 9.3 内容审核运行要求 12
 - 9.3.1 审核规则库 12
 - 9.3.2 审核内容 12
 - 9.3.3 智能审核 12
 - 9.3.4 人工审核 12
 - 9.3.5 人工标注 12
 - 9.3.6 人工抽检 12
 - 9.3.7 审核结果管理 12
 - 9.3.8 审核沟通工具 12
 - 9.3.9 第三方智能审核接口 12
 - 9.3.10 审核流程管理 13
 - 9.4 送审主体管理(可选) 13
 - 9.5 审核组织管理(可选) 13
 - 9.6 审核培训管理(可选) 14
 - 9.7 审核监测分析 14
 - 9.8 数据接口 14
 - 9.9 系统性能 15
- 10 媒体宣传 15
 - 10.1 直播类型划分 15
 - 10.2 直播评级划分 15
 - 10.3 直播资源投入 15
 - 10.4 直播技术标准 15
 - 10.5 下线管理机制 15

前 言

本文件按照GB/T 1.1—2020《标准化工作导则 第1部分：标准化文件的结构和起草规则》的规定起草。

本文件由中国通信企业协会团体标准管理委员会提出并归口。

本文件主要起草单位：中国电信股份公司全渠道运营中心、天翼数字生活科技有限公司、央视国际网络有限公司、阿里巴巴（北京）软件服务有限公司。

本文件参加起草单位：深圳市腾讯科技有限公司、中兴通讯股份有限公司、思特威电子科技股份有限公司、深圳市佳思图科技有限公司、宇瞳光学科技股份有限公司、广西广播电视台都市频道、浙江日报全媒体视频影像部。

本文件主要起草人：张小刚、陈丹、张晨啸、邓德宝、高玥、曾守毅、吕小鹏、许伟、徐婷婷、王应飞、何辉辉、薛梦蛟、杨志斌、于小博、崔金、陈明明、庞霞、刘耀东、耿兆森、谭泱、刘澎浩、张占军、徐斌、徐文迪。

本文件为xxx发布。

引 言

本文件的发布机构提请注意，声明符合本文件时，可能涉及到……[条]……与……[内容] ……相关的专利的使用。

本文件的发布机构对于该专利的真实性、有效性和范围无任何立场。

该专利持有人已向本文件的发布机构承诺，他愿意同任何申请人在合理且无歧视的条款和条件下，就专利授权许可进行谈判。该专利持有人的声明已在本文件的发布机构备案。相关信息可以通过以下联系方式获得：

专利持有人姓名：……

地址：……

请注意除上述专利外，本文件的某些内容仍可能涉及专利。本文件的发布机构不承担识别专利的责任。

5G 慢直播建设规范

1 范围

本文件规定了 5G 慢直播系统建设规范和相关技术要求。

本文件适用于 5G 慢直播设备、网络、平台之间交互和通信，包括设备信息采集、网络配置、设备管理、流处理、AI 能力、内容播控、媒体分发等。

2 规范性引用文件

下列文件中的内容通过文中的规范性引用而构成本文件必不可少的条款。其中，注日期的引用文件，仅该日期对应的版本适用于本文件；不注日期的引用文件，其最新版本（包括所有的修改单）适用于本文件。

GB/T 38754-2020 IPTV 媒体交付系统技术要求 流媒体服务

YD/T 2807.5-2015 云资源管理技术要求 第 5 部分：存储系统技术要求

YD/T 3846.7-2021 互联网基础资源支撑系统信息交换接口规范 第 7 部分：内容分发网络（CDN）

3 术语和定义

下列术语和定义适用于本文件。

3.1

三维重建技术 3D reconstruction

通过深度数据获取、预处理、点云配准与融合、生成表面等过程，把真实场景刻画成符合计算机逻辑表达的数学模型。

3.2

定焦

定焦是指焦距是固定的，不能把眼前的景物拉近或推远，优点是比变焦成像质量好，光圈更容易做大。

3.3

帧率 (Frame rate)

帧率 (Frame rate) 是以帧称为单位的位图图像连续出现在显示器上的频率（速率）。该术语同样适用于胶片和摄像机，计算机图形和动作捕捉系统。帧速率也可以称为帧频率，并以赫兹 (Hz) 表示。

3.4

码率

码率，又称比特率，是指单位时间内传送的比特 (bit) 数，单位为bps (bit per second)；

3.5

播控

播控就是：播出控制的意思，播出控制是电视媒体节目审核的最后把关。

3.6

CDN

全称是Content Delivery Network，即内容分发网络。其基本思路是尽可能避开互联网上有可能影响数据传输速度和稳定性的瓶颈和环节，使内容传输得更快、更稳定。通过在网络各处放置节点服务器所构成的在现有的互联网基础之上的一层智能虚拟网络，CDN系统能够实时地根据网络流量和各节点的连接、负载状况以及到用户的距离和响应时间等综合信息将用户的请求重新导向离用户最近的服务节点上。其目的是使用户可就近取得所需内容，解决 Internet网络拥挤的状况，提高用户访问网站的响应速度。

3.7

流媒体

流媒体（streaming media）是指将一连串的媒体数据压缩后，经过网上分段发送数据，在网上即时传输影音以供观赏的一种技术与过程，此技术使得数据包得以像流水一样发送；如果不使用此技术，就必须在使用前下载整个媒体文件。流式传输可传送现场影音或预存于服务器上的影片，当观看者在收看这些影音文件时，影音数据在送达观看者的计算机后立即由特定播放软件播放。

4 5G 慢直播技术架构

4.1 设备终端

慢直播设备终端，一般采用摄像头，摄像头（webcam）一般具有视频摄像/传播和静态图像捕捉等基本功能，它是借由镜头采集图像后，由摄像头内的感光组件电路及控制组件对图像进行处理并转换成电脑所能识别的数字信号，然后借由并行端口或 USB 连接输入到电脑后由软件再进行图像还原。

摄像头的工作原理大致为：景物通过镜头（LENS）生成的光学图像投射到图像传感器表面上，然后转为电信号，经过 A/D（模数转换）转换后变为数字图像信号，再送到数字信号处理芯片（DSP）中加工处理，再通过 USB 接口传输到电脑中处理，通过显示器就可以看到图像了。

4.2 网络环境

网络是用物理链路将各个孤立的工作站或主机相连在一起，组成的数据链路。通信网络是指将各个孤立的设备进行物理连接，实现人与人，人与计算机，计算机与计算机之间进行信息交换的链路，从而达到资源共享和通信的目的。

4.3 视联网

视联网是网络发展的重要里程碑，是互联网的更高级形态，是一个实时网络，能够实现互联网无法实现的全网高清视频实时传输，将众多互联网应用推向高清视频化，高清面对面。最终将实现世界无距离，实现全球范围内人与人的距离只是一个屏幕的距离。

视联网采用实时高清视频交换技术，可以在一个网络平台上将任何所需的服务，如高清视频会议、视频监控、智能化监控分析、应急指挥、数字广播电视、延时电视、网络教学、现场直播、VOD 点播、电视邮件、个性录制（PVR）、内网（自办）频道、智能化视频播控、信息发布等数十种视频、语音、图片、文字、通讯、数据等服务全部整合在一个系统平台，通过电视或电脑实现高清品质视频播放。

4.4 视频直播

视频直播是指利用互联网及流媒体技术进行直播，视频因融合了图像、文字、声音等丰富元素，声形并茂，效果极佳，逐渐成为互联网的主流表达方式。视频通过真实、生动的传播，营造出强烈的现场感，吸引眼球，达成印象深刻、记忆持久的传播效果；

直播业务一般采用组播的网络方式来实现。所谓组播就是利用一种协议将 IP 数据包从一个信息源传送到多个目的地，将信息的拷贝发送到一组地址，送达所有想要接收它的接收者。IP 组播是将 IP 数据包“尽最大努力”传输到一个构成组播群组的主机集合，群组的各个成员可以分布于各个独立的物理网络上。IP 组播群组中成员的关系是动态的，主机可以随时加入和退出群组，群组的成员关系决定了主机是否接收送给该群组的组播数据包，不是某群组的成员主机也能向该群组发送组播数据包。同单播或广播相比，组播效率非常高，因为任何给定的链路至多用一次，可以节省网络带宽和资源。

4.5 音视频处理

音视频处理是指对对音视频流协议进行处理及修改，从而实现预期的效果；常见的音视频处理包括编码、解码、转码服务；

编码(encode)

所谓视频编码方式就是指通过压缩技术，将原始视频格式的文件转换成另一种视频格式文件的方式。视频流传输中最为重要的编解码标准有国际电联的 H. 261、H. 263、H. 264。

解码(decode)

通过特定的解压缩技术，将某个视频格式的视频流转换成原始的视频格式。

转码(transcode)

使用视频转码技术将视频信号从一种格式转换成另一种格式。

4.6 内容审核

主要对包含色情、涉政、涉政敏感人物、图文违规、暴恐、违禁、广告等垃圾信息的文本、图片、音频、视频进行检测和识别，通过系统化的方式提供审核、打标、自定义配置等能力来保障企业内容安全。一方面为了净化网络，让人们在享受网络便利的同时，也是安全的。另一方面客户降低业务违规风险。一般公司产品包括图片审核、文本审核、音频审核、视频审核。

4.7 内容分发

内容分发网络(CDN)是构建在现有网络基础之上的智能虚拟网络，依靠部署在各地的边缘服务器，通过中心平台的负载均衡、内容分发、调度等功能模块，使用户就近获取所需内容，降低网络拥塞，提高用户访问响应速度和命中率。

5 5G 慢直播设备要求

5.1 摄像头终端

5.1.1 图像要求

摄像头终端图像要求包括：

- a) 有效像素应不低于 2880×1620 ；
- b) 支持 HDR 功能；
- c) 图像帧率应不低于 25fps；
- d) 图像效果突出锐度和对比度，不接受图像拖尾；
- e) 信噪比 SNR1 应不高于 0.45lux。

5.1.2 三防要求

三防要求包括：

- a) 应使用隔离式电源；
- b) 应支持 POE48V 供电和 DC12V 供电；
- c) 防水防尘等级应不低于 IP66；
- d) 防雷等级应不低于 $\pm 6kV$ 。

5.1.3 定焦场景

定焦场景要求包括：

- a) 像素应不低于 500 万像素；
- b) 应支持 4-5mm 定焦镜头；
- c) 光圈规格 F1.6 及以上；
- d) 靶面 1/2.7"；
- e) 水平视场角范围为 $80^\circ - 95^\circ$ ；

- f) 光学畸变范围为-35%-45%;
- g) 支持 POE 接口, 可选 TRRS 接口 (带串口)。

5.1.4 动态场景

动态场景要求包括:

- a) 像素应不低于 800 万像素;
- b) 帧率应不低于 40fps;
- c) 应支持 3.3mm-40mm 变焦;
- d) 光圈规格 F1.8 及以上
- e) 靶面 1/2.8” ;
- f) 水平视场角范围为 80° -8° ;
- g) 光学畸变 5%以内;
- h) 支持 POE 接口, 可选 TRRS 接口 (带串口)。

5.1.5 无源场景

5.1.5.1 定焦场景

定焦场景要求包括:

- a) 像素应不低于 500 万像素;
- b) 应支持 4-5mm 定焦镜头;
- c) 光圈规格 F1.6 及以上;
- d) 靶面 1/2.7” ;
- e) 水平视场角范围为 80° -95° ;
- f) 光学畸变范围为-35%-45%;
- g) 4G CAT.4 网络支持;
- h) 应支持 12V DC 接口, 可选 TRRS 接口 (带串口)。

5.1.5.2 动态场景

动态场景要求包括:

- a) 像素应不低于 800 万像素;
- b) 应支持 3.3mm-40mm 变焦;
- c) 光圈规格 F1.8 及以上
- d) 靶面 1/2.8” ;
- e) 水平视场角范围为 80° -8° ;
- f) 光学畸变 5%以内;
- g) 5G 网络支持;
- h) 支持 12V DC 接口, 可选 TRRS 接口 (带串口)。

5.2 非摄像头终端

太阳能储能套装要求包括:

- a) 太阳能电池板功率应不低于 100 瓦;
- b) 储能电池应不低于 12V/40AH;
- c) 应支持 12V DC 电源输出接口×2。

6 网络环境与设备接入

6.1 网络环境

应提供 POE 宽带接入, 有限网络带宽不低于 8Mbps。4G 或 5G 宽带覆盖区域, 无线带宽均值不应低

于 10Mbps。

6.2 设备接入管理

6.2.1 总体原则

依托运营商云网资源，基于运营商视联平台底座能力，构建慢直播服务设备接入平台。系统架构设计目标包括：

- a) 基于能力服务化形成高内聚、低耦合的业务能力中台，适应业务拓展；
- b) 支持千万级在线设备的就近接入和管控能力；
- c) 提供低带宽成本的 P2P 直连+转发混合的直播能力；
- d) 支持分省建设的云存储池集中接入管理和调度，支持跨云混合使用；
- e) 支持各种行业、各个场景下的慢直播业务技术需要；
- f) 基于分布式 NoSQL 提供高可用、高性能，支持千亿级数据的存储和快速检查。

6.2.2 系统设计要求

系统设计原则应符合可用性、可靠性、可扩展性等原则，具体要求如下：

- a) 可用性：用户对系统业务访问的可用性应达到 99.95%，平台应综合考虑系统的可靠性和稳定性进行规划、设计和建设，避免单点故障，满足系统整体的可用性要求；
- b) 可靠性：系统应支持高可靠性和稳定性，关键系统和数据应具备可靠性保护能力、容错能力、故障恢复能力、备份能力；
- c) 可扩展性：系统应具备良好可扩展能力，包括新增功能和模块的扩充，系统容量的扩展，功能的扩充和容量的扩展不应影响现有的系统架构和业务的开展和运营造成影响；
- d) 业务可扩展：根据业务的需求支持新的运营模式、计费模式等，能够根据用户需求增加新的增值应用；
- e) 开放性和兼容性：系统遵循相关的国内外技术标准和规范，系统可支持不同厂家设备的互操作性

6.2.3 整体方案

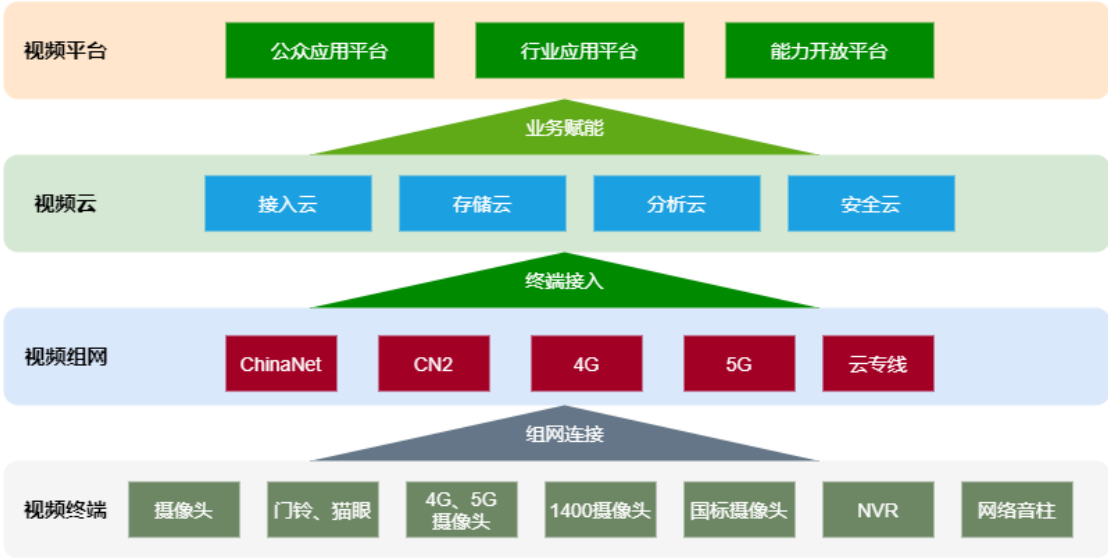
视联平台平台是基于“视频云+网+平台+终端”的云网边端安融合能力底座，实现了关键能力的标准化、组件化和云化，承载了面向公众用户的天翼看家服务、面向行业用户的标准化与定制化视频监控服务，提供了视频终端的标准化接入、云存储、监控、AI、大数据、运维等原子能力与组件，通过能力开放平台可支撑跨用户群融合场景和应用快速上线场景化服务，衍生视联平台宽带等服务形态，形成端到端全系列产品和解决方案。

视频终端：视频终端是视频监控业务的根本基础，视联平台基于标准化的嵌入式 SDK 技术形成了统一终端体系，目前主要包含的终端形态有普通摄像头、高清摄像头、AI 摄像头，NVR，网络音柱、猫眼、门铃等。

视频组网：视联平台平台基于运营商 ChinaNet、CN2、4G、5G 和云专线等基础网络能力，满足客户对公网、专网、有线网、无线网等混合组网的需求。

视频云：视联平台平台基于运营商数字化平台底座能力研发和部署接入云、存储云、分析云和安全云等视频相关核心云能力，作为视频业务的能力中台支撑业务发展。

视频平台：视联平台平台面向公众、行业和能力开发提供三大应用平台，有针对性地满足不同业务场景下的视频业务需要。



视联平台涉及到的端到端关键技术如下：

1) 多区域、多边缘分布式部署技术

传统的单一数据中心部署的方式在平台稳定性、平台支撑容量和扩容效率、平台服务质量监控、平台海量数据跨数据中心一致性同步、多数据中心调度与切换等方面存在问题，难以满足业务快速发展带来的系统容量和稳定性要求。

视联平台集约运营平台通过构建高效灵活的基于云原生的“多区域中心+多边缘”分布式技术架构，包含了系统的可扩展、弹性，及企业云的所有特性，保障了业务超量超常规扩展。针对上述关键问题，视联平台集约运营平台的对策如下：

(1) 构建“多区域中心+多边缘”分布式技术架构保证系统稳定性与并发性：针对关键服务双中心化部署，云化高可用集群架构，实现双服务负载均衡，故障自动切换，确保关键服务高可用，实现了用户无故障感知，实现了用户层的稳定性。构建“多区域中心+多边缘”分布式云网部署架构支撑并发。针对突发为超规模业务并发的情况，可伸缩，可扩展的业务架构是支撑的关键之一。针对摄像机和流媒体接入，视频类海量数据集中接入，存在中心节点资源大、网络带宽要求高、扩展性差等问题。视联平台集约运营平台创新使用边缘部署技术，将设备接入和流媒体转发业务下沉到省份部署，实现业务去中心化部署架构，通过动态调度技术，将设备就近接入，实现更好的用户体验，避免单点故障影响全网用户，大流量视频数据不出省，节省了大量骨干网流量，避免大量视频数据在骨干网影响。

(2) 基于云原生的资源编排、监控、调度技术与系统：提供容器自动化管理、治理和编排能力，整合了容器中的每层从虚拟层到容器层的系统日志、网络状态、多环境下的跨集群系统的日志及状态，实现全面的系统监控，达到复杂系统可控性目的。

(3) 混合式数据库和分布式部署技术：视联平台集约运营平台基于融合列式数据库和 NoSQL 混合技术实现的数据存储和检索能力，规避了传统关系型数据库在扩展性、可用性和性能上的问题，可支撑千亿级数据量，实现千万级设备量数据存储要求。

(4) 云资源池智能调度及快速切换技术及系统：视联平台集约运营平台基于多维度策略的动态调度技术的资源池调度管理系统，实现设备文件的就近接入和存储，减少跨省流量对骨干网的冲击，可实现资源池一键上线、下线，资源池故障时快速实现业务调度，减低用户的负体验。

2) 智能视频监控 AI 中台技术

视联平台集约运营平台 AI 中台是一套完整的智能模型全生命周期管理平台和服务编排体系，整合自研及第三方智能服务构建算法仓库、规范标准化接口打造服务开放共享生态、采用云边端协同架构提升服务处理性能，同时对服务研发角色、研发流程进行标准化、自动化管理，对业务应用提供个性化智能服务的迅速构建能力、实现第三方能力引入。其目标是将可复用的 AI 模型、服务沉淀共享，快速配置、组合、编排，快速输出满足用户个性化需求的智能应用。



3) 异构终端软件统一封装对接技术

针对多厂家、多种类设备接入难问题，视联平台集约运营平台创新地提出了视频终端软件统一 SDK 封装技术，实现了视频处理、终端与平台交换协议的统一封装，通过 SDK 封装，屏蔽了各大厂家技术实现差异，实现多商家和多产品快速接入，市场的快速推广。统一 SDK 区分设备端和客户端，设备端 SDK 需支持 Linux、LiteOS、Android 操作系统，并提供给摄像头、门铃等设备终端接入，客户端 SDK 需支持 Android、iOS、PC（win）等操作系统，并提供自营 APP 或第三方 APP 接入使用。通过客户端 SDK 对接平台，再通过平台对接设备 SDK，实现视频播放交互端到端的封装，屏蔽视频技术差异，实现视频播放优化。通过对 FFmpeg 视频编解码改造及优化，支持全格式视频码流，支持国标 GB/28181；支持 IPV6/IPV4 双栈网络，在 IPV6 不稳定情况下，采用 HappyEyeballs 算法进行快速回退至 IPV4。通过采用播放器应用层 DNS 缓存策略，减少 DNS 解析时间消耗，提升云回看视频播放点开速度。

承载方案

视联平台平台采用两级服务结构，即：通过集约平台承载用户的认证和业务流量，通过下沉的服务节点承载用户视频流和 AI 计算等大规模流量。同时，除公网接入和承载外，为了确保政企大客户的安全需求，视联平台产品支持通过专网接入和承载。

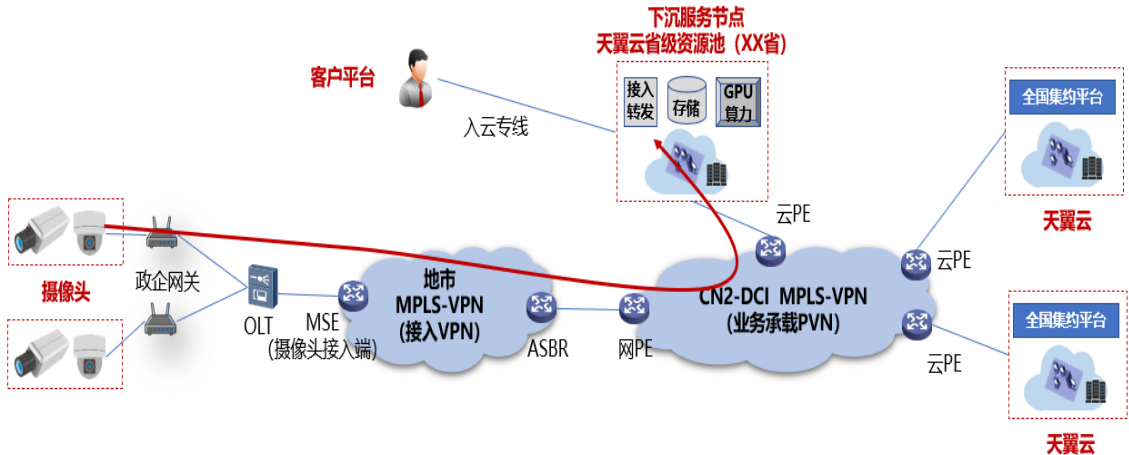
结合视联平台平台的服务特征，以“基于边缘服务在本地承载大规模业务流量，通过运营商新型城域网能力穿透中心和边缘云”为基本建设原则，基于运营商 CN2 骨干网络新建视联平台专用 VPN，承载视联平台集约业务平台和在各省下沉的服务节点；在各地市配置专用接入 VPN 网络，实现对视联平台业务终端流量的高速汇聚。

1) 视联平台专用网络

为确保视联平台业务面向各类客户的服务质量，运营商建立视联平台的专用 VPN。

视联平台专用 VPN 分为：业务承载 VPN、接入 VPN。其中业务承载 VPN 完成视联平台集约业务平台和各省下沉的本地服务节点互联，接入 VPN 负责用户终端和客户端的接入。

业务承载 VPN：基于骨干 CN2 配置，集约业务平台下挂在资源池所在省骨干 CN2 的 PE 设备下；各省视联平台下沉节点下挂于所在省 CN2 骨干 PE 设备下；接入 VPN：各本地网基于 PON 组网专线接入视联平台专用 VPN。



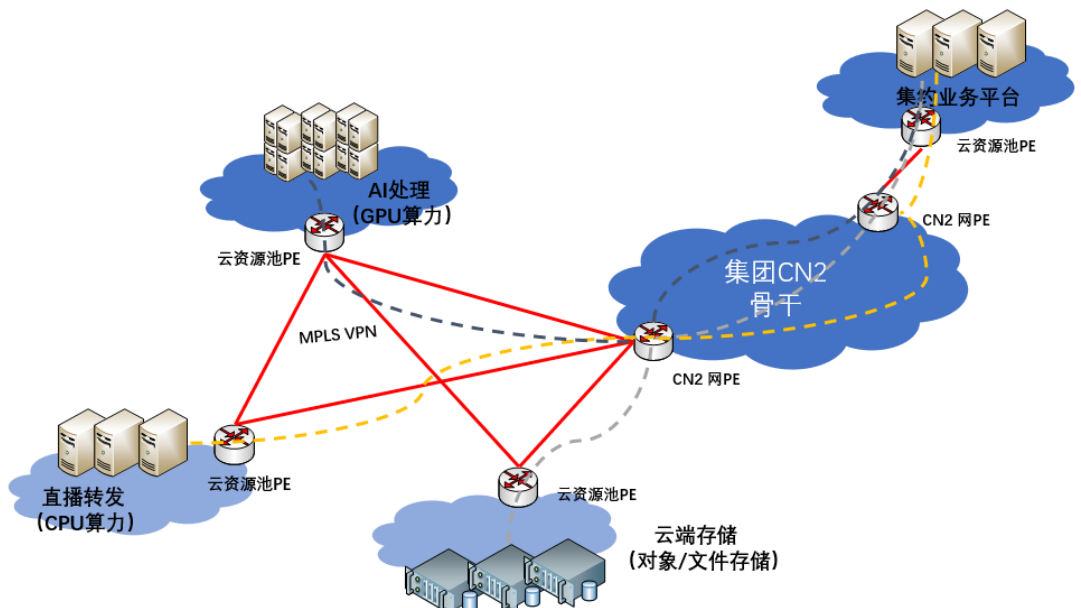
2) 全国中心和各省下沉服务的部署

视联平台的全国中心节点采用多 AZ 高可用部署,负责用户和终端校验、负责业务订购关系的管理、负责对接管理和调度各本地服务节点。

本地服务节点由接入转发点+存储点+算力点构成,基于边缘云承载。本地服务节点包括:

- (1) 接入/转发点: 每个省至少配置一个接入/转发点。接入/转发点由通用算力构成,负责频流接入、转码、面向用户或开发方转发,逐步在地市下沉。
- (2) 存储点: 每个省至少配置一个存储点,由视频云存设备构成,用于存储云回看文件,随着规模增长逐步在地市下沉。
- (3) 算力点: 每个分节点按业务需求配置算力点。算力点由 GPU 计算资源构成,用于处理实时视频流、云存内容。

对于已建的本地服务节点,其接入/转发(通用 CPU 算力)、存储、算力通过所在云资源池 PE 与在省 CN2 骨干 PE 互联接入视联平台 VPN,实现与集约平台安全、高速互联。省内客户通过视联平台专用 VPN 享受低时延、高可用的服务保障。



通过云网一体的组网架构,以“中心云-边缘云”承载数字生活业务,按业务流量调度不同层级云服务,通过新型城域网穿透各级云。认证和管理流量在中心平台承载,视频流、云存储和 AI 计算流量在边缘节点承载,充分发挥运营商云网融合的能力优势。

7 云服务

7.1 流媒体规范

慢直播应满足 GB/T 38754-2020《IPTV 媒体交付系统技术要求 流媒体服务》相关技术要求；

7.2 直播加速

慢直播相关的 CDN 基础设施应满足 YD/T 3846.7-2021《互联网基础资源支撑系统信息交换接口规范 第7部分：内容分发网络（CDN）》相关技术要求。

7.3 云储存

慢直播涉及的云存储部分应满足 YD/T 2807.5-2015《云资源管理技术要求 第5部分：存储系统技术要求》相关技术要求。

8 AI 能力技术要求

8.1 人脸审核

8.1.1 总体要求

人脸审核是基于人的脸部特征信息进行身份识别的一种生物识别技术，对图片和视频源中面部特征进行提取分析。

人脸检测和人脸识别应涵盖人脸检测与分析、五官定位、人脸对比与搜索、跨年龄人脸识别等功能。人脸识别流程应包括：人脸检测、人脸对齐、特征提取、特征检索四部分。

8.1.2 人脸检测

人脸识别的第一个环节是找到图像里面的人脸，提取出最可能是人脸的边界框，人脸检测算法应支持 30×30 像素以上人脸能被识别。

8.1.3 人脸对齐

对识别出的侧脸、扭头、倾斜等变形人脸进行校准，以得到更靠近正脸的图像。其中人脸对齐算法应支持以下条件内的对齐处理，

- f) 俯仰角小于 30 度；
- g) 偏航角（左右侧旋转）小于度；
- h) 横滚角小于 180 度。

8.1.4 人脸特征提取

得到经过人脸对齐处理的图像后，应使用一个特征向量来表示一张人脸图像，支持在人脸特征检索环节进行相似度计算。

8.1.5 人脸特征检索

人脸识别的最后一个环节是人脸特征检索。从已经在人脸库里的特征里找到最相似的一组特征，并根据数据库的人脸 ID 信息得到此人的相关信息返回。

8.2 智能审核分类

针对慢直播业务和场景应包括涉黄、涉恐、涉暴、涉政、违禁、广告等敏感内容提供全方位的智能审核能力。

- a) 涉黄 AI 识别模型应具备分级模型，分为涉黄、性感等一级标签及二级标签，涉黄 AI 模型具备分类学习和识别模型；
- b) 涉政暴恐 AI 识别模型应具备针对特殊符号、暴恐组织、暴力事件进行分类识别的能力，并提

- 供相似度数据；
- c) 不良场景 AI 识别模型应具备针对吸毒、赌博、不文明手势等场景进行分类识别的能力，并提供相似度数据；
 - d) 广告 AI 识别模型应具备针对网址、二维码、水印等场景进行分类识别的能力，并提供相似度数据；
 - e) 视频文字 AI 识别模型应具备针对视频中文字的敏感内容、涉政内容、色情内容等行分类识别的能力，并提供相似度数据；
 - f) 语音 AI 识别模块应具备违规语音检测及低俗语音检测的分类识别能力。

8.3 AI 审核策略

通过 AI 图像智能识别、OCR 文字识别、ASR 语音识别等技术，并结合实时更新维护的特征库，针对慢直播业务和场景中重要人物和敏感人物、涉黄、涉恐、涉暴、涉政、违禁、广告等敏感内容提供全方位的智能审核能力。

针对多种慢直播业务和场景中审核尺度以及不同时期的审核要求，可调整变更审核策略，结合业务需要，应提供灵活的自定义配置功能。

8.4 视频 DNA 技术

基于视频基因检索和动态比对技术，精确匹配已审视频和待审视频的重重复度查找，便于查看差异点，并快速定位至非相同片段。

基于深度神经网络技术提取媒资库中所有视频的视频基因并建立基因检索库，从而实现多模态检索功能，应包括以图搜图、以图搜视频以及以视频搜视频功能，同时应支持在视频中标注待检索图像或视频出现的时间点。

9 播控安全技术要求

9.1 流媒体内容编码及文件格式要求

9.1.1 流媒体属性编码及文件格式要求

- 慢直播编码格式宜支持以下视频属性编码标准。
- a) ISO/IEC14496-10/H.264 MPEG-4 第 10 部分 AVS1/AVC/H.264 MainProfile level3 或 HighProfile level3；
 - b) ISO/IEC23008-2:2013 AVS2/HEVC/ITU-T H.265 MainProfile level3 以上。
- 编码参数应具备高清、超高清编码要求规范，具体见表 1。

表 1 慢直播编码参数要求

参数	高清	超高清
码流类型	复合流	复合流
分辨率	1920×1080	3840×2160
帧率	25fps	50fps
编码方式	H264	H265
I 帧间隔	50	50
SVC	关闭	关闭
扫描模式	逐行	逐行

音频编码	AAC	AAC
采样率	48kHz	48kHz
音频码率	64kbps	256kbps

音频编码应支持 MPEG-4 AAC、HEAAC 中至少一种编码标准。

9.1.2 慢直播传输协议要求

慢直播系统应至少支持 RTMP、RTSP、HTTP、GB28181 流格式中的一种。

- a) 流接收模式：应支持通过 GB28181、RTSP、RTMP、SRT 协议将节目流数据传输至系统；
- b) 流节目应支持但不限于：音频流支持 AAC、AC3 等格式；直播流支持 RTMP、HLS 等协议。
- c) 互联网多终端播放至少支持 HLS、RTMP 等多种格式一种。

9.2 直播播出运行要求

9.2.1 慢直播源站

系统应具备慢直播汇聚、分发、版权保护能力，支持对慢直播的流进行策略配置，按照相关行业互联网直播规范要求，对网络视听节目内容进行网络直播。

9.2.2 直播分发

慢直播内容分发应遍布全球的加速节点，可实现全网分发，满足客户最后一公里的低时延的视频监控和录像回看传输。

9.2.3 实时预览

系统应支持超添加设备完毕后且设备在线的情况下，对设备进行远程视频监控实时预览。并支持开启轮巡监控，按照系统策略依次开展预览监控。

9.2.4 慢直播媒体数字版权保护

系统应具备网络视听节目内容数字版权保护功能，实现对节目的加密、解密、验证、授权，具体要求包括，

- a) 应提供音视频加密功能，针对音视频内容进行版权保护；
- b) 应提供音视频验证功能，提供终端用户观看验证授权观看内容；
- c) 应提供音视频解密功能，支持对视频按照 DRM 标准进行解密内容。

9.2.5 慢直播日志

系统应支持将直播切换结果通过系统交互层接口反馈给节目送审主体，并应长期保存节目内容直播切换报告，包括但不限于节目标识信息、切换过程记录，切换日志，保存时间应不少于 180 天。

9.2.6 系统性能要求

系统性能要求如下，

- a) 单台服务器部署条件下，应支持对分辨率不大于 1920×1080 的直播流进行大于 200 路并发的接入处理；
- b) 单台服务器部署条件下，对音视频节目分发性能应大于 200 路；
- c) 系统性能稳定，应支持 7×24 小时不间断运行；
- d) 系统运行过程中所有服务器 CPU 的平均使用率应低于 80%；
- e) 系统运行过程中总存储使用率应低于 90%；
- f) 单台服务器部署条件下，应支持至少 1000 个用户同时访问。

9.3 内容审核运行要求

9.3.1 审核规则库

系统应具备违规内容审核规则库，支持对审核规则的策略配置，按照相关行业审核规范要求，对网络视听节目内容进行审核。

9.3.2 审核内容

系统支持的审核内容应包括但不限于点播节目、直播节目、用户上传节目、评论信息、弹幕信息、账号信息、短视频、广告素材等。

9.3.3 智能审核

应具备网络视听节目内容智能审核功能，实现对节目的辅助审核，具体要求包括：

- a) 应提供图像智能审核功能，鉴别图像是否包含违规内容；
- b) 应提供文字智能审核功能，分析文本语义是否包含违规内容；
- c) 应提供视频智能审核功能，支持对视频按照关键帧或定时抽帧的方式进行图像提取、特征识别、分析和比对，判断是否包含违规内容；
- d) 应提供音频智能审核功能，支持对普通音频、方言类音频和多语种音频，进行语音识别、分析和比对，判断是否包含违规内容；
- e) 支持审核任务分配，对智能审核后的节目和审核结果，应自动建立人工审核任务并分配给系统审核小组或者指定审核小组进行人工审核；
- f) 支持根据人工审核和人工标注的结果，对智能审核模型进行优化训练和迭代更新，保障违规内容的及时、有效检出；
- g) 支持审核结果作为结构化标签，对接到内容智能分析系统中，以配合后续进一步的节目内容编辑、剪辑应用。

9.3.4 人工审核

系统分配的人工审核任务，应提供智能审核后的节目原文件、智能审核过程中对节目进行识别、抽帧等处理的中间结果以及智能审核标注的疑似违规内容信息，人工审核操作事项包括但不限于：通过、修改、删除等。

9.3.5 人工标注

对审核员发现的智能审核未标注的违规节目内容，系统应提供审核数据人工标注敏感标签的功能。

9.3.6 人工抽检

应支持按系统预设的比例抽取审核通过和未通过的节目内容审核结果数据进行人工抽检复核。

9.3.7 审核结果管理

应支持将节目内容审核结果和审核进度状态通过系统交互层接口反馈给节目送审主体，并应长期保存节目内容审核报告，包括但不限于节目标识信息、审核过程记录，违规原因及结果，保存时间应不少于180天。

9.3.8 审核沟通工具

应在系统内提供与审核流程相关联的即时通讯、人员群组、留言等多种沟通工具，用于审核过程协同和审核结果沟通。

9.3.9 第三方智能审核接口

宜提供第三方智能审核接口功能，能够通过API、RPC、WEBSERVICE等服务调用接入第三方的智能审核算法模型，用于对系统中送审节目的智能审核，并可配置不同类型的送审节目调用不同的审核模型接口进行审核。

9.3.10 审核流程管理

9.3.10.1 预审

系统具有预审库或待审库，存储节目送审主体提交的待审节目内容样本、资质信息等材料，预审支持对送审内容合规性、格式正确性、责任信息完备性等方面进行审核。

9.3.10.2 三审一检

系统对待审节目宜包括初审、复审、终审和抽检四个功能，所有审核工作可分配给不同的审核员进行审核，任意审核环节不通过则退回。

9.3.10.3 审核管理

应建立并实施一套与系统审核流程配套的规范管理制度，包括但不限于审核过程信息的存储制度、审核人员的考核激励制度、审核信息的公示制度以及用于监督管理的通报制度等。

9.3.10.4 申诉机制

如送审主体对审核结果有疑问，可提出申诉理由，再交由第三方审核判定是否属于审核员失误。

9.4 送审主体管理(可选)

9.4.1 送审主体注册

对外提供第三方节目审核服务的系统应具有送审主体注册功能，送审主体可通过在线填写相关信息完成注册认证，注册审核通过后可接入并获取系统提供的审核服务。

9.4.2 送审节目提交

应支持送审主体通过系统接口、FTP 传输和专线接入等方式提交送审节目，应支持对传输文件的完整性校验功能，宜支持断点续传功能。

9.4.3 审核结果查看

应向节目送审主体提供审核结果查看功能，能够查看审核结果和审核数据内容。具体要求如下：

- a) 审核通过的数据内容应被做标记，并支持播放查看被打标记的数据内容；
- b) 审核不通过的数据应支持返回节目送审主体，整改后重新上传系统审核。

9.4.4 审核流程编排

送审主体可根据自身送审节目内容审核的个性化需求，自定义设置多级审核流程，审核数据可在系统中自动根据编排的审核流程进行查看。

9.5 审核组织管理(可选)

9.5.1 组织管理

应具备审核组织的创建和分级管理功能。高等级的审核组织可管理自身及下级所属组织的审核数据，对跨组织的审核流程结果须其中等级较高的组织审核通过后方可生效。

9.5.2 审核员管理

应具备节目内容审核员管理功能，包括上岗前的岗位责任书签署、培训考试，工作过程中的制度规范、继续教育、职位晋升、绩效考核以及离岗后的脱密期管理等。

9.5.3 审核员等级管理

应具备审核人员等级管理功能，包括但不限于初级审核员、中级审核员、高级审核员、审核专家，并可根据审核员的培训考核结果进行等级晋升管理。

9.5.4 结算流程

应具备审核人员的收益计算、结算等功能。

9.5.5 审核员评估

应具备对每个审核员的审核质量进行评估、评测的功能，并提供评估结果的定期汇总和反馈功能。

9.6 审核培训管理(可选)

9.6.1 证书管理

应具备网络视听节目内容审核证书管理功能，由管理员在系统中建立审核员相关的各种资质证书，审核员进驻系统或等级评定时可在系统中通过关联证书进行资格审查。

9.6.2 课程管理

应具备培训课程信息的建立、增加、修改、删除、作废、过滤、排序、发布等功能。

9.6.3 课件管理

应具备对培训中使用的培训课件进行上传、修改、删除、配置课程关联等管理功能。

9.6.4 考试管理

应具备题目分析、题库管理、试卷管理、考试安排、考试监控、考试阅卷等功能。

9.6.5 在线学习

应具备审核人员注册、报名、课程学习等功能。

9.7 审核监测分析

9.7.1 审核行为监测

应具备审核人员日常审核行为的记录、监测功能，并提供审核员审核指标分析、效能分析等分析模型。

9.7.2 大屏监控

应采用可视化的方式实现对系统运行情况的有效监控，具备系统故障告警和审核流程监控预警等功能。

9.7.3 质量监控

应具备对审核数据的质量监控功能，可对漏审或错审数据进行监测、记录和告警。

9.7.4 运营统计

应具备对入驻系统的用户数量、审核节目总量、已通过审核总量、未通过审核总量等运营数据进行分析显示和报表输出功能。

9.8 数据接口

对送审节目和相关数据应支持通过以下接口调用的方式送审：

- a) HTTP 接收模式：送审主体通过 HTTP/HTTPS/RPC 接口将节目数据传输至系统；
- b) 消息推送模式：送审主体通过消息队列接口将节目数据流推送至系统；
- c) 数据库读取模式：送审主体提供节目数据库只读权限，系统通过数据库接口读取节目数据；
- d) 送审接口的节目和素材应支持但不限于：
 - 音频文件支持 MP3、WAV 等格式；
 - 图片文件支持 JPG、PNG、BMP、GIF 等格式；

- 视频文件支持 FLV、MKV、TS、MP4、WMV、AVI、RMVB、3GP 等格式；
- 文本文件支持 TXT、CSV 等格式；
- 直播流支持 RTMP、HLS 等协议。

9.9 系统性能

系统性能要求如下：

- a) 单台服务器部署条件下，对分辨率不大于 1024×768 的图片智能审核处理性能应大于 15 张每秒；
- b) 单台服务器部署条件下，对音频节目智能审核处理性能应大于 6 倍实时播放速率；
- c) 单台服务器部署条件下，对分辨率不大于高清 (1920×1080) 的视频节目智能审核处理性能应大于 3 倍实时播放速率；
- d) 系统性能稳定，应支持 7×24 小时不间断运行；
- e) 系统运行过程中所有服务器 CPU 的平均使用率应低于 80%；
- f) 系统运行过程中总存储使用率应低于 90%；
- g) 单台服务器部署条件下，至少应支持 100 个用户同时访问。

10 媒体宣传

10.1 直播类型划分

平台应针对慢直播类型进行体系划分，包括但不限于泛娱乐、游戏、体育、电商、其它等。

10.2 直播评级划分

平台应针对慢直播内容进行评级标准划分，包括但不限于 S+级，S 级，A 级，B 级等相应评级标准。

10.3 直播资源投入

应建立完善的慢直播资源扶持体系，运营平台应根据直播内容评级，给予形式多样的媒体传播资源矩阵。

10.4 直播技术标准

慢直播视频接入应满足如下最低标准：

- a) 编码：H264；
- b) 高宽比：横屏 16:9，竖屏：9:16；
- c) 分辨率：720P, 1080P；码率：2m-1m；
- d) 帧率：20-60fps；
- e) 关键帧间隔：2s；
- f) 音频：aac-1c, 64-128kb/s。

10.5 下线管理机制

互联网平台应制定相应人工审核或 AI 审核机制，对违法违规直播内容进行及时下线处置。