

# 面向数据中心的RISC-V视频转码卡应用实践

李天正

中国电信研究院

1

RISC-V视频转码卡研究背景

2

RISC-V视频转码卡解决方案介绍

3

RISC-V视频转码卡应用实践

# 视频流媒体和视频监控领域市场空间巨大

市场规模：千亿级别  
主战场：视频监控、视频流媒体

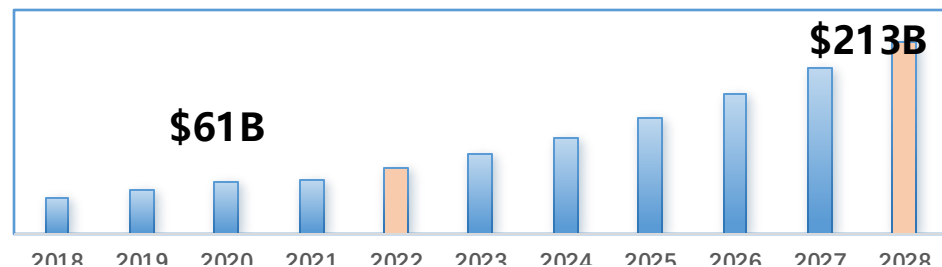
核心问题：视频传输带宽高、存储压力大、  
编码格式异构

2021-2025国内视频监控市场规模



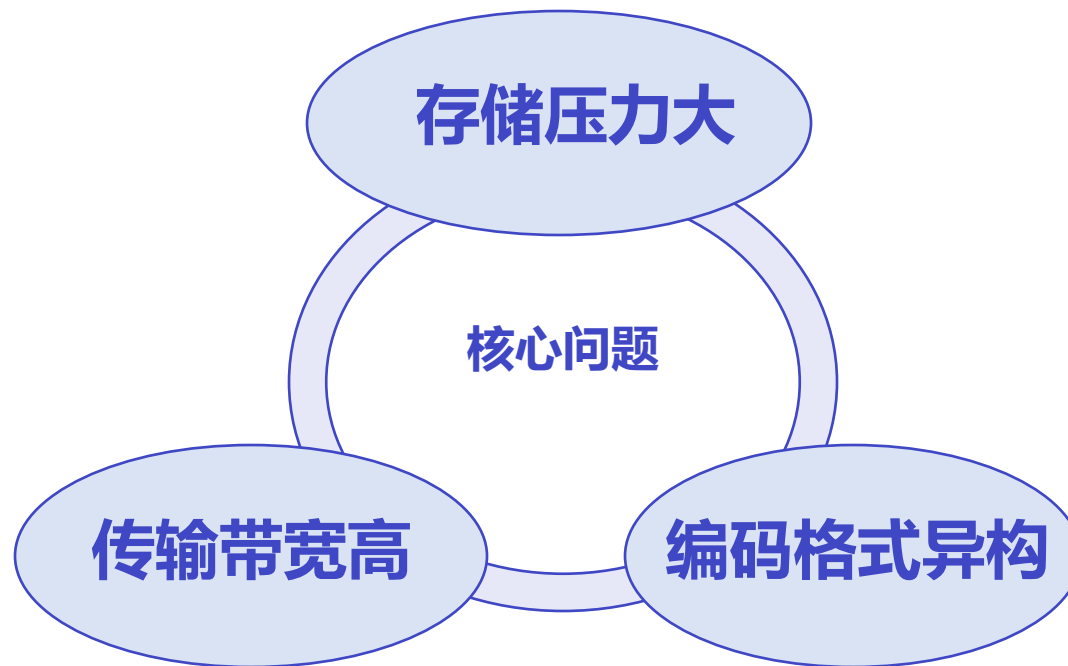
国内存量监控摄像头已超6亿路，年增过1亿路机构预测2024年市场规模超**4000亿元**，增速**12-16%**。

全球视频流媒体市场规模及预测



据预测，视频流媒体的市场价值 **2028年的超过 2130 亿美元**，复合年增长率约为 20%。

万路视频需PB级存储空间

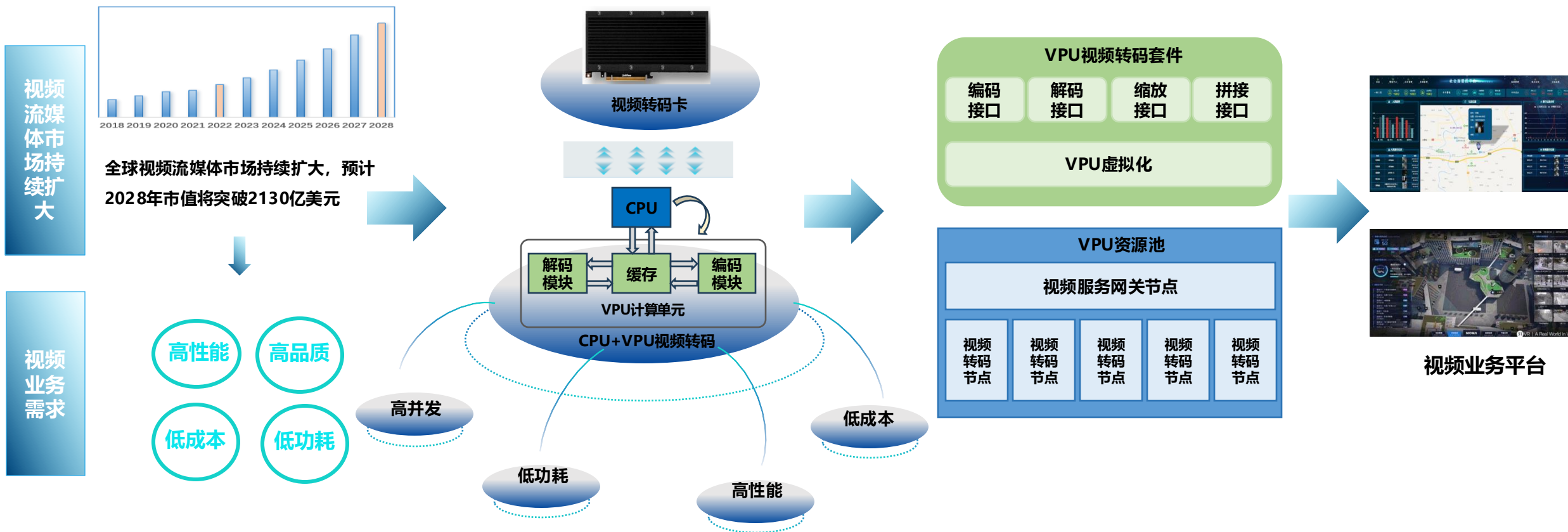


万路视频需GB级  
传输带宽

现网视频存在H264、  
H265等格式，且帧率、  
码率等均各不相同

# VPU可满足大规模视频业务的需求

VPU是一种全新的视频计算加速器，具有硬件化的编解码、图像拼接等功能，既可以维持视频的图像信息和画面质量，同时又能减少服务器CPU负荷、网络带宽的消耗，大幅降低数据中心的运营成本。



随着视频业务的大规模发展，对硬件AI能力和成本等方面也提出了更高的要求。

- 传统CPU转码效率低，GPU硬件成本高，不满足大规模视频业务的需求。
- ASIC转码效率高，但仅能用于单一场景，且硬件难以更新迭代。
- RISC-V VPU兼具高效的视频处理能力和高灵活性，还可用于边缘计算等场景。

方案	转码效率	灵活性	AI能力	功耗	硬件成本
CPU	低	高	低	中等	中等
GPU	中等	中等	高	高	高
ASIC	高	低	低	低	低
RISC-V VPU	高	高	中等	低	低

1

RISC-V视频转码卡研究背景

2

RISC-V视频转码卡解决方案介绍

3

RISC-V视频转码卡应用实践

TeleVPU集成了5颗国产高性能RISC-V芯片，支持40路1080p/25fps并行编码，单卡平均功耗60瓦，单路1080P视频转码功耗低至1.5瓦。



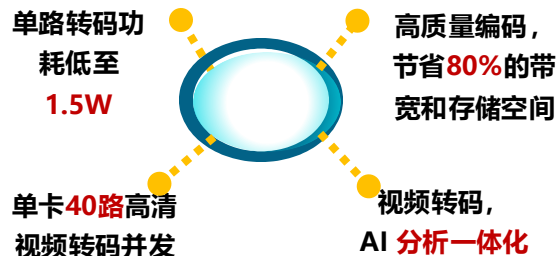
TeleVPU实物图

规格参数		
VPU	5 × 4核RISC-V 1.85GHz	
主机操作系统	CTyunOS、Ubuntu、CentOS、OpenEuler	
内存	5 × 8 GB	
单卡转码路数	分辨率	路数
	4K	10
	1080P	40
	720	80
解码视频格式	H.265、H.264、VP9/8/7/6、AVS/AVS+/AVS2.0、VC1、MPEG4	
编码视频格式	H.265、H.264、MJPEG	
图像处理能力	任意分辨率缩放、画面拼接、锐化、模糊、降噪、水印/图片/文字叠加	
AI算力	20Tops INT8	
功耗	单卡最大功耗75W，平均满负荷转码功耗60W	
工作温度	0 °C ~55 °C	
存储温度	-40 °C~ 75 °C	



TeleVPU集成了高并发、低功耗的实时视频转码能力，视频流经过转码可节省高达80%的带宽和存储空间，同时TeleVPU具备20 TOPS的AI算力支持对视频的智能化处理。

## 高并发、低功耗实时视频转码



### 转码并发

#### 单卡转码视频流路数@25fps

4K	H.264/H.265	10
1080P	H.264/H.265	40
720P	H.264/H.265	80
功耗	单卡最大功耗75W, 平均满负荷转码功耗60W	

## 视频处理与AI计算融合

- 单卡**20 TOPS AI**算力;
- 支持**实时转码**, 将不同编码格式的视频数据转换为其他格式;
- 支持**视频结构化并行计算**, 实现更加智能化的视频处理和分析功能。

### AI 计算

#### 单卡推理算力

INT8 20TOPS

**支持业务类型** 支持实时转码和视频结构化并行计算

## 支持多种图像处理能力

- 支持**缩放**, 最高支持**4096\*3072**分辨率的视频缩放;
- 支持锐化、模糊、降噪等**滤波能力**;
- 支持基于图片和文字叠加的**水印**。

### 图像处理

#### 硬件图像处理加速器

缩放	最高支持 <b>4096*3072</b> 分辨率
滤波	锐化、模糊 降噪
水印	图片、文字叠加

## 标准接口支撑灵活部署和扩展

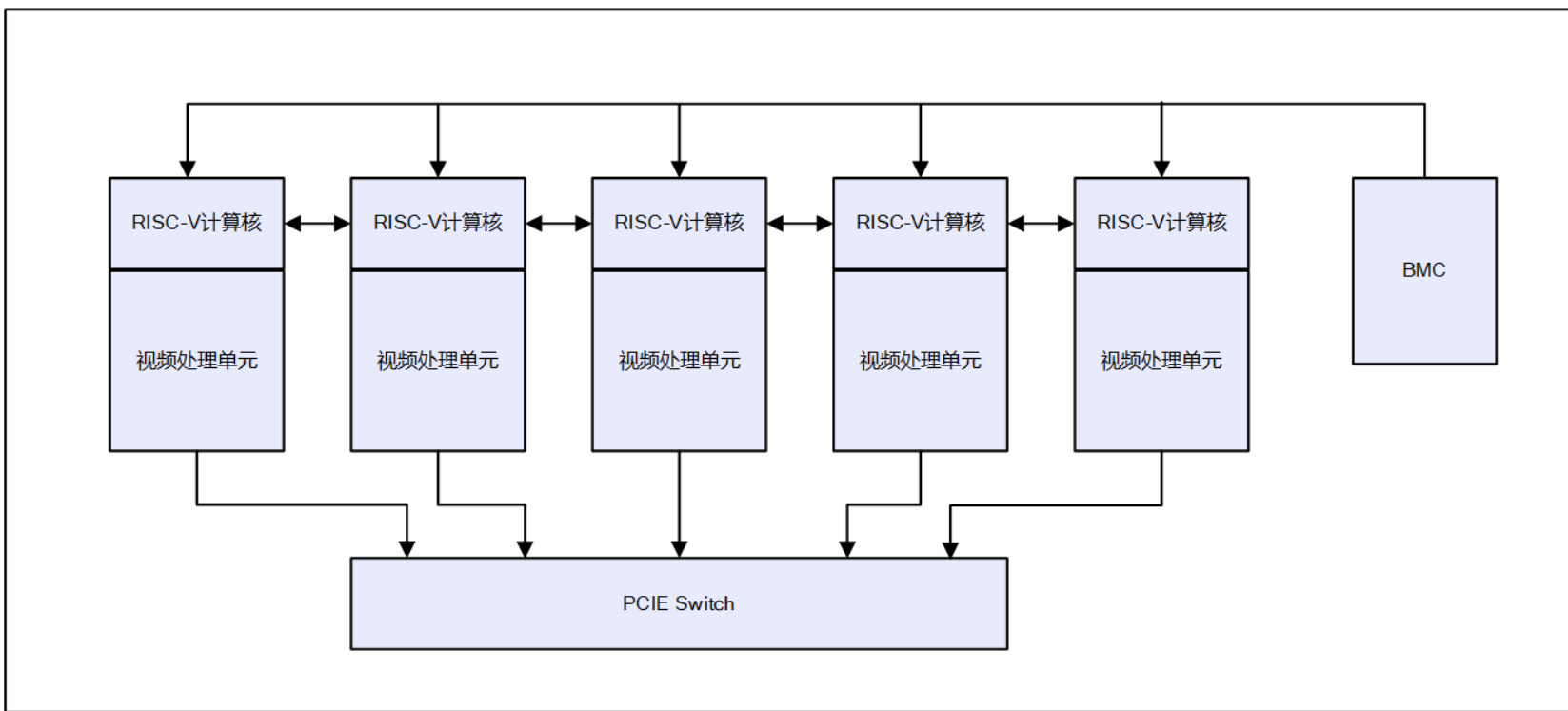
- 基于**FFmpeg API**快速集成进视频云系统, 主机侧服务软件提供完整的**生产环境使用和运维接口**;
- 标准4U服务器可集成**8个PCI-E**规格的VPU转码卡, 实现单台服务器**320路 1080P@25fps**实时转码。

### 接口与环境

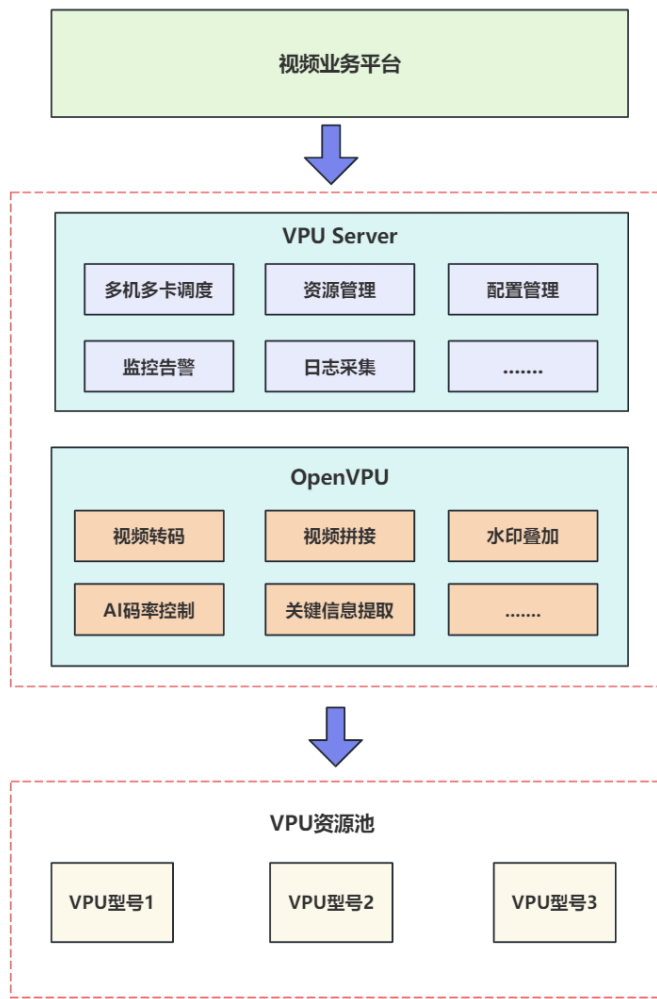
接口	FFmpeg、Codec/Filter 主机侧接口	流管理	创建、删除、维护转码流
		设备状态	实时获取设备工作状态、硬件状态
		流状态	转码流的状态，带宽、帧率、负载
		维护	转码卡远程重启、升级、日志提取
主机操作系统		CTyunOS、Centos、Ubuntu、openEuler	
FFmpeg		FFmpeg 2.8.6、3.2.7	



- 5颗国产高性能RISC-V芯片并联，通过PCIE接口与主机通信。
- 单独的板卡级BMC管理，以融入现有的服务器集群运维体系。
- 芯片间独立工作，任务和视频流可在芯片间实时切换。



TeleVPU硬件架构



整体软件架构

## VPU Server

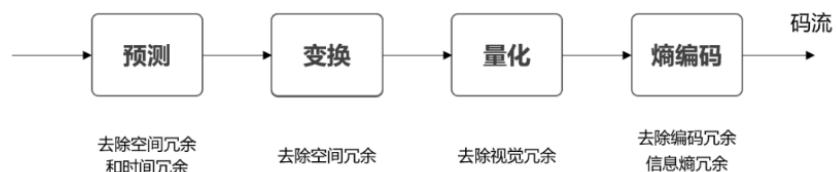
基于k8s平台实现VPU资源池化，负责任务调度、资源管理、监控告警等功能。

## OpenVPU

屏蔽不同VPU型号软件的差异，实现标准化的VPU软件接口。

- **视频处理**：在主机上由后台服务负责多卡的任务调度、视频流分配等任务。在VPU上由主程序负责调度管理具体的计算任务。
- **AI码率控制**：基于AI能力在转码过程中动态调节码率。
- **关键信息提取**：通过算法识别提取视频中的关键帧或片段。

## 视频压缩原理—去除冗余信息



原始视频采集数据中存在大量冗余信息，视频压缩编码即去除冗余数据、保留图像有效信息

- 时间冗余：视频前、后帧存在大量相似图像部分，如监控画面背景，对于重复出现的画面和物体，等间隔的重复产生大量时间冗余数据
- 空间冗余：同一帧画面里，存在大量相似图像块，如天空、墙壁等，各相似图像块的采集数据产生大量空间冗余数据
- 视觉冗余：视频图像原始数据中，存在人眼无法分辨的数据，如色彩的精细色度差异、亮度的辨别阈值等，产生了视觉冗余数据
- 信息熵冗余：表达某一信息所需要的比特数总比理论上表示该信息所需要的最少比特数要大，之间的差距就为信息熵冗余



IPC摄像机：重采集、轻压缩



分辨率：1920×1080  
帧率：25帧/秒  
编码格式：H.264

4Mbps  
带宽



VPU：深度压缩去冗余



分辨率：1920×1080  
帧率：25帧/秒  
编码格式：H.264

0.8Mbps  
带宽

1

RISC-V视频转码卡研究背景

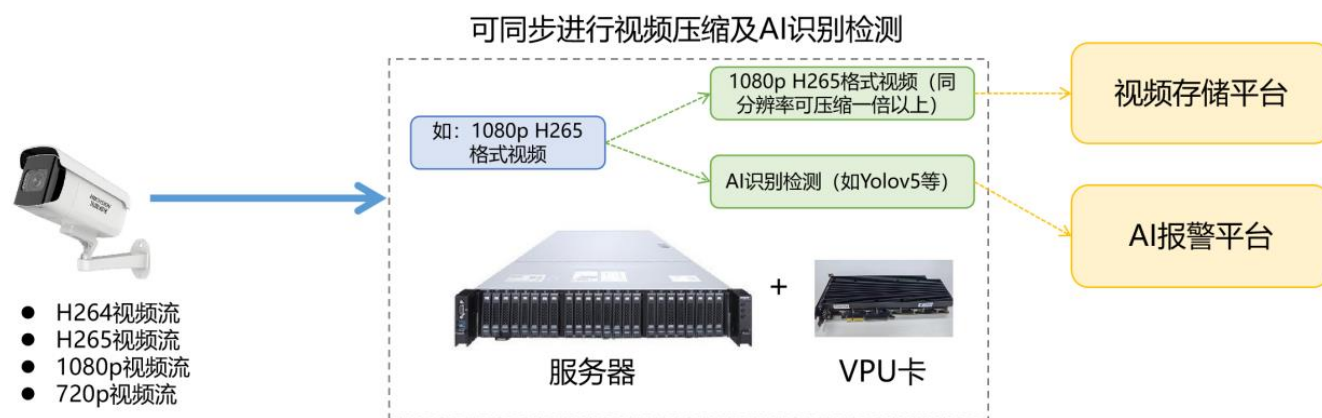
2

RISC-V视频转码卡解决方案介绍

3

RISC-V视频转码卡应用实践

在一视频监控业务中，存在数十万路规模视频流，且视频数据需存储三个月以上，占用了大量的存储资源。针对上述问题，在其数据中心现有服务器安装TeleVPU进行测试，测试视频转码压缩、多路处理等能力。



TeleVPU在视频监控业务中应用示例



TeleVPU部署试验



原始视频 (1080P 25fps H.264) , 在保持分辨率、帧率和编码格式不变的情况下进行压缩



原始视频 (1080P 25fps, 大小为151MB)



压缩后视频 (1080P 25fps, 大小为31.4MB, 降低80%)

总时长约22小时的视频片段（1080P 25fps H.265）大小总计39G，全部压缩为720P，使用单卡5芯并行压缩

```
./lvaccl_transcoding -2 1 source/5/3-32041300001317001291-1697261350-1697261652.mp4 ou
50-1697261652.mp4 -s 1280:720 -r 10 -b 600000 -g 50 -bf 0 -profile main -preset defaul
source/5/3-32041300001317001291-1697257136-1697257438.mp4
5/3-32041300001317001291-1697257136-1697257438.mp4
./lvaccl_transcoding -2 1 source/5/3-32041300001317001291-1697257136-1697257438.mp4 ou
36-1697257438.mp4 -s 1280:720 -r 10 -b 600000 -g 50 -bf 0 -profile main -preset defaul
source/5/3-32041300001317001291-1697260446-1697260748.mp4
5/3-32041300001317001291-1697260446-1697260748.mp4
./lvaccl_transcoding -2 1 source/5/3-32041300001317001291-1697260446-1697260748.mp4 ou
46-1697260748.mp4 -s 1280:720 -r 10 -b 600000 -g 50 -bf 0 -profile main -preset defaul
source/5/3-32041300001317001291-1697258342-1697258642.mp4
5/3-32041300001317001291-1697258342-1697258642.mp4
./lvaccl_transcoding -2 1 source/5/3-32041300001317001291-1697258342-1697258642.mp4 ou
42-1697258642.mp4 -s 1280:720 -r 10 -b 600000 -g 50 -bf 0 -profile main -preset defaul

real    14m24.280s
user    10m55.561s
sys     4m52.113s
```

263个视频片段总时长约22小时

```
[root@localhost vpuoneday]# du -h source/
3.4G    source/12
2.4G    source/1
3.2G    source/2
3.6G    source/3
3.7G    source/7
3.3G    source/8
3.2G    source/11
3.5G    source/4
3.2G    source/9
3.5G    source/6
3.9G    source/10
2.7G    source/5
39G     source/
```



```
[root@localhost vpuoneday]# du -h out_h265/
497M    out_h265/12
381M    out_h265/1
490M    out_h265/2
452M    out_h265/3
594M    out_h265/7
524M    out_h265/8
432M    out_h265/11
453M    out_h265/4
449M    out_h265/9
548M    out_h265/6
499M    out_h265/10
427M    out_h265/5
5.7G    out_h265/
```

压缩后总大小为5.7G



TeleVPU提供了监测工具用来监测视频转码卡的运行状态，此外也提供了HTTP接口，外部服务可以通过访问特定的URL获取JSON格式的状态信息。

module id(s)	card id	if no.	model name	ip address	port	mac address	in error	tasks count	decoding load	encoding load	scaling load	temp. (C)	power (W)	firmware version
2	177	3	UA3170	192.168.100.4	80	48:DA:35:60:17:FB	no	8	26.44%	99.90%	25.19%	59.37	6.00	V0.2.0.8
3	177	2	UA3170	192.168.100.2	80	48:DA:35:60:18:15	no	8	26.44%	99.90%	25.19%	60.25	6.00	V0.2.0.8
4	177	4	UA3170	192.168.100.3	80	48:DA:35:60:18:09	no	8	26.44%	99.90%	25.19%	56.73	6.00	V0.2.0.8
5	177	1	UA3170	192.168.100.6	80	48:DA:35:60:18:11	no	8	26.44%	99.90%	25.19%	55.41	6.00	V0.2.0.8
6	177	5	UA3170	192.168.100.5	80	48:DA:35:60:18:13	no	8	26.44%	99.90%	25.19%	53.21	6.00	V0.2.0.8

TeleVPU模块信息页面

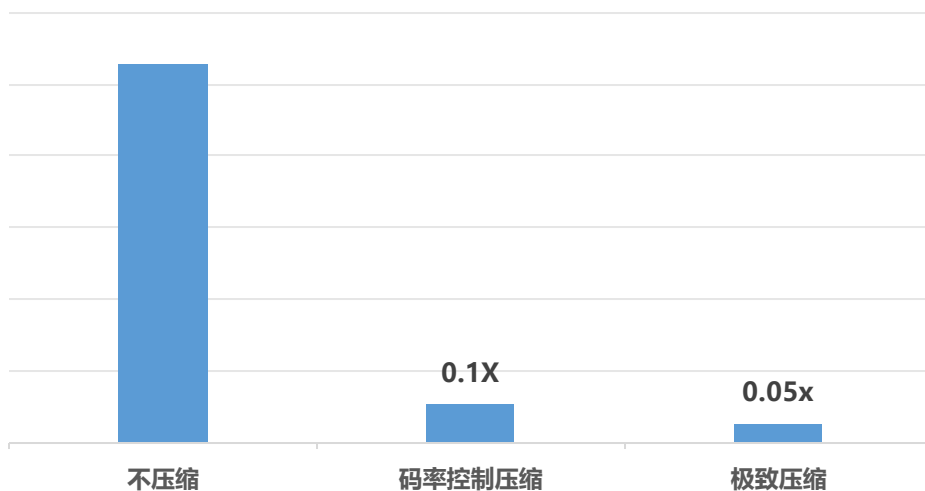
```
"code": 0, // 执行成功返回0，其他值表示失败
"reason": "ok.", // 失败的原因
"modules_infos": [{ // 所有模块信息
  "id": 0, // 模块ID
  "ip": "192.168.100.64", // 模块IP
  "mac": "00:04:4B:8C:B7:49", // 模块MAC
  "port": 80, // 模块http监听端口
  "card_no": 1, // 模块所在的卡的ID
  "br_port_no": 3, // 所在的网卡在网桥中的序号。
  "on_bridge": 1, // 是否处在PCIE桥中
  "temperature": 43, // 模块温度
  "power": 5.508000, // 模块功耗
  "voltage": 12.072000, // 模块电压
  "current": 0.264000, // 模块电流
  "available_decoding_ability": 920678400, // 模块空闲可用的解码能力
  "available_encoding_ability": 479692800, // 模块空闲可用的编码能力
  "available_scaling_ability": 1921766400, // 模块空闲可用的缩放能力
}
```

JSON格式模块信息

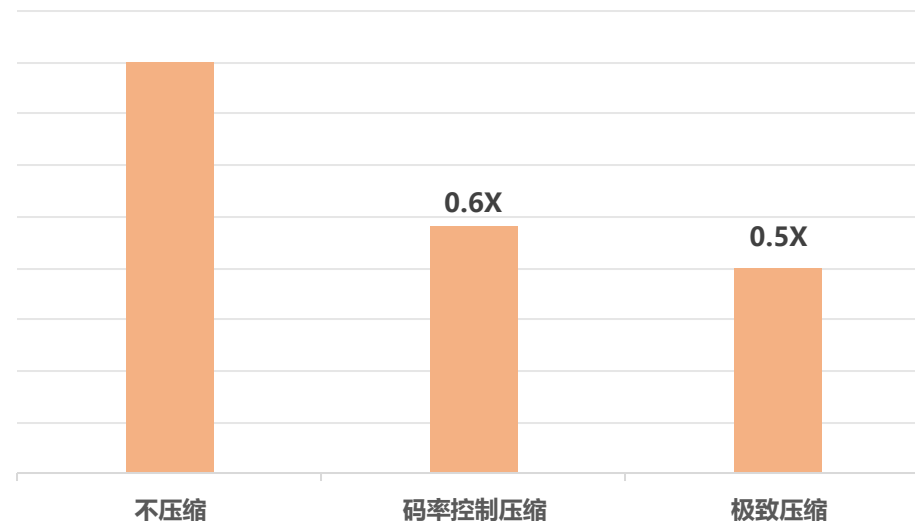
按照1万路1080P H.264视频，三个月视频存储周期计算：

- 在保证视频分辨率不变的情况下，采用变码率压缩方案，最大可节省90%存储空间及40%成本。
- 在码率控制基础上采用极致压缩方案，最大可节省95%存储空间及50%成本。

## 视频存储容量



## 方案成本



## 主机定制

定制RISC-V服务器集成16卡  
VPU，提高单机最大转码路数

## 构建生态

将OpenVPU贡献开源社区，  
共同构建VPU软硬件生态。

**谢 谢!**

