How GPU Helps to Power Next Generation of ArcVideo Video Products and Service

Jin Huang
jhuang@arcvideo.com
CTO
ArcVideo Inc.
Agenda

• ArcVideo introduction
• How GPU could help for media processing
• Real Cases Studies
• New Business Opportunities
What is ArcVideo

• ArcVideo is our Enterprise Video Solutions and Services, including
  – Codec engines, containers and streaming protocols, developed by own over 10+ years
  – End device players, which are validated by 100M+ OEM devices shipping
  – Cloud services

• It’s a suite including most video related workload, including Broadcasting Level Video Transcoding, Video Processing, Video Analytic, Video Streaming, Video Playback, Video Big Data, and Cloud implementation

• ArcVideo Solutions focus on Content Providers, Broadcasting companies and TV Stations, and are expending to Education, Telecom, Security as well.
ArcVideo™ Live
High density and broadcasting quality real-time stream transcoding/encoding system

ArcVideo™ Core
High performance and professional quality file transcoding/encoding system

ArcVideo™ Cloud
Transcoding and Delivering Service for both Public and Private Cloud

ArcVideo™ Supervisor
All IP based video contents monitoring and intelligent error detection for TV channels and OTT multi-screen

ArcVideo™ ICR
Intelligent Content Recognition based on IVA, Face Recognition, Object Recognition, etc technologies

ArcVideo™ Ingest
Intelligent and high performance Ingest system, turns content into easily used assets

ArcVideo™ Player
Playback, User Interaction and User Data Collection SDK for content providers, running on all kinds of end devices
GPU Advantages for Video Industry

- Hardware accelerated decoding and encoding
  - Fast Decoding capability with good error resilience
  - NVENC for multiple sessions of encoding, with various quality level and latency mode choices
- CUDA accelerated video post processing
  - Adaptive Deinterlacing/Frame Rate Up-conversion
  - Various video enhancement algorithms
- Deep Learning acceleration for intelligent video analytic workload
  - Face Recognition
  - Object Recognition

- Less Servers and Space, save up to 80%
- Fast delivery time, up to 85% faster
- Less Power consumption, save up to 60%

= One ArcVideo GPU Server
ArcVideo Customized parts:

• Transcoding pipeline optimization
  – Seamless buffer sharing between HW Decoding, Video Processing/Analyzing and HW Encoding
  – Handling various streaming content dynamic change
  – Reduce unnecessary overhead moving uncompressed buffer

• Modified NVENC with customization of QP
  – Better rate control over NVENC

• CUDA accelerated video processing and video analytic
  – Scaling/Video composition/CC/Subtitle

• CUDA accelerated video Codec
  – MPEG2 422/444, 10bit, HDR
  – Apple ProRes / Avid DNxHD
  – Perceptual Based Coding

• Deep Learning acceleration
  – Face Recognition
  – Object Recognition, like Car and Cloth

• User interaction for live event broadcasting
  – Real time AR rendering
NVCUVID

**Benefit**

- Major Codec Support like MPEG2/H.264/HEVC
- High Performance
- Industry level of bit-stream error tolerant

**User scenarios**

- Real time Multi-Screen Video Transcoding and Delivering
  - 1 HD input, multiple resolution output
  - With various video post processing
  - 20+ simultaneous sessions per 1U server
- Video Wall support 100 SD or 32 HD channels per GPU board easily
- Low latency APP/Game streaming to STB
NVENC

Benefit

• Impressive Performance, with various Quality Levels for different user scenarios
  – High Quality/High Performance/Low Latency

User scenarios

• Real time Multi-Screen Video Transcoding and Delivering
  – 4x 4K HEVC and 8x 4K H.264 real time transcoding sessions for one single M60

• Offline Fast Transcoding
  – 32x HD/4~8x 4K faster than real time (1x M60)
CUDA Accelerated Video Processing

**Benefit**

- Easy to customize
- High Performance
- Highly efficient pipeline to access and interop with HW Decoding and HW Encoding

**User scenarios**

- Fast Video processing and analyzing
  - Scaling/Composition/ Anti-Shaking/Denoise/Deblock/Adaptive Color enhance/Sim4K (quality enhancement)
  - Video quality and fault detection
  - Special Video Codec support, like MPEG2, etc.
  - Video analyzing
What ArcVideo could benefit

• Impressive video performance comparing to traditional CPU or DSP/FPGA multimedia solutions:
  – Good Hardware accelerated Decoding/Encoding performance, and tons of CUDA cores
• Highly customizable via CUDA programming
  – Easy to customize CUDA accelerated video post processing and video analytic algorithms
  – Flexible CUDA programming to easily fit customer request in very short time
• Balanced GPU and CUDA core configuration
  – Both Tesla and GRID provide various combination of GPU and CUDA core to fit different user scenarios
  – Mature server vendors ecosystem to find reliable GPU servers, depends on task burden, pick multiple GPU board and achieve highest density
CNTV – Broadcasting full IP based monitoring solution

- Monitoring hundreds of HD and SD video inputs
- Need low bandwidth
- Monitoring remotely from any mobile devices

- Full IP based solution, and easily remote control and manage anywhere
- Support most popular streaming protocols and video formats
- Support up to 48 HD or 144 SD video simultaneously in 1U server
- Easy to expand to support 4K input and output
ChinaCache CDN—Arcloud Game Live Broadcasting

- Arcloud live video platform bring reliable game broadcasting service to millions audiences
- Help Game companies to stream thousands of games you love, with high quality and low latency

- Multi-Screen transcoding support up to 4K
- Support most popular streaming protocols and video formats
- Low latency transcoding
New Business Opportunities
-Perceptual Video Coding

• Perceptual video coding attracts lots of interests recently
• HVS (human visual system) characteristics help to exploit perceptual redundancy and improve video compression efficiency, but involves more computing
  – Luminance, Contrast sensitivity, fovea, etc
  – SSIM/M-SSIM/CW-SSIM/VIF/VQM
  – Perceptual Noise from Spatial and Temporal Perspectives
  – Region of Interest
New Business Opportunities
- 4K/HFR/HDR/10bit/VR

- Pursuing higher image quality leads to higher resolution, higher frame rate, dynamic range (bit depth), chroma sampling scheme (4:2:2, 4:4:4)
- CPU performance increase only 30-50% per year, hard to catch Live 4K/HDR/HFR/VR streaming request
- Next generation Codec HEVC/AVS2 are 3-10x more complex than current MPEG2/H.264
- Video content analyzing and monetizing with Deep Learning acceleration
  - Face Recognition, Car, Cloth
New Business Opportunities
- Private/Public Cloud Video Service

- GPU Virtualization helps service providers to move video business to private/public cloud to expand their business quickly and with affordable cost
- Ensure high demanding streaming business performance, like 4K, VR live broadcasting

- Single purpose and dedicated hardwares moving towards general purpose and multi-tenant servers, server Virtualization especially GPU Virtualization will play more important rule for multimedia service
  - Video Editing and Transcoding share same physical or virtual server, work at day for editing, and night time for transcoding
  - Video analytic services with GPU accelerated Video Decoding and Recognition capability could save lots of costs comparing to CPU instance
Global Leader in Multimedia Solution
Feedback

• Email: jhuang@arcvideo.com

• Please complete the Presenter Evaluation sent to you by email or through the GTC Mobile App. Your feedback is important!