GPU-Based MultiPlatform Transcoding

Mahmut Ş. Sağiroğlu

Founded in 2007, Erlab Software is a privately-held company based in Istanbul, Turkey. Erlab Software is a technology company that develops innovative and value added solutions for the customers with emerging technologies.
Unique Position in the Market

Specialities:
- Internet TV,
- Content Management System (CMS),
- PlayReady DRM,
- Video Transcoding, Video Streaming,
- Microsoft Smooth Streaming,
- HTTP Live Streaming,
- Windows 8, Windows Phone 8,
- Kinect, Xbox, Microsoft Pixelsense,
- Enterprise Video Solutions,
- OTT Video Solutions,
- Smart TV

Market Position:
- Covered 100% of Internet TV / Video market in Turkey with his own solutions.
- Serving to 80% of the subscribers with Erlab’s Content Management Solutions
Encoding is the most time consuming process in video processing.

Video decoding / encoding / processing is inherently parallellizable.

New generation NVIDIA GPU’s have an embedded hardware for video encoding.
Motivation

- Video encoding is always on the agenda
  - %90 of the internet users watches video on any device
  - Lifelogging trend
  - Satellite Imagery
  - 2 trillion minutes (5 million years) of video content will cross the Internet each month in 2017

Challenges

Encoding and processing needs heavy mathematical operations
Erlab’s Running CPU Based Transcoding Solution

For Adaptive Streaming

- Source 1
  - Decode
  - Resize 1
  - Encode 1
  - Packetize
- Source 2
  - Decode
  - Resize 2
  - Encode 2
  - Packetize
- Source m
  - Decode
  - Resize m
  - Encode m
  - Packetize
- Transmit
CPU Based Solution

DECODING

PRE-PROCESSING
AND ENCODING

PRE-PROCESSING
AND ENCODING

POST-PROCESSING
AND DELIVERY

operations

F1 F2 F3 F4 F5 . . . Buffer is full . . .

F1 F2

F1 F2 . . .

F1 F2 . . .

F1 F2 . . .
# NVENC Features

<table>
<thead>
<tr>
<th>Feature</th>
<th>What it enables</th>
</tr>
</thead>
<tbody>
<tr>
<td>H.264 base, main, high profiles</td>
<td>Wide range of use-cases</td>
</tr>
<tr>
<td>Up to 8x HD encode (1080p @ 240 fps)</td>
<td>Faster than real-time encoding</td>
</tr>
<tr>
<td>Flexible ME, QP maps</td>
<td>Customizable quality, region of interest encoding</td>
</tr>
<tr>
<td>YUV 4:2:0 and planar 4:4:4 support</td>
<td>High quality encoding without chroma subsampling</td>
</tr>
<tr>
<td>Up to 4096 × 4096 in HW</td>
<td>High resolution encode</td>
</tr>
<tr>
<td>NVENC and CUDA parallelism</td>
<td>Simultaneous and parallel HW and CUDA encoding for increased performance</td>
</tr>
</tbody>
</table>
GPU Accelerated

Operations:

- Decoding
  - F1, F2, F3, F4, F5, F6, F7, F8, F9, ...
- Pre-processing and encoding
  - F1, F2, F3, F4, ...
  - F1, F2, F3, F4, ...
  - F1, F2, F3, F4, ...
- Post-processing and delivery
  - F1, F2, F3, ...

Time axis
## CPU vs GPU Transcoder (Criteria)

**System**  
2x Intel Xeon 2650v2  
2x NVIDIA GRID K1  
32 GB RAM

**Input**  
Resolution: HD 720p  
Length: 5.700 seconds

<table>
<thead>
<tr>
<th>Profile</th>
<th>Stream Type</th>
<th>Bitrate Kbps</th>
<th>Resolution</th>
</tr>
</thead>
<tbody>
<tr>
<td>Main</td>
<td>Smooth</td>
<td>3.400</td>
<td>720p</td>
</tr>
<tr>
<td>Main</td>
<td>Smooth</td>
<td>2.250</td>
<td>540p</td>
</tr>
<tr>
<td>Main</td>
<td>Smooth / HLS</td>
<td>1.500</td>
<td>540p</td>
</tr>
<tr>
<td>Main</td>
<td>Smooth / HLS</td>
<td>1.000</td>
<td>360p</td>
</tr>
<tr>
<td>Base</td>
<td>Smooth / HLS</td>
<td>650</td>
<td>360p</td>
</tr>
<tr>
<td>Base</td>
<td>Smooth / HLS</td>
<td>400</td>
<td>180p</td>
</tr>
</tbody>
</table>
CPU vs GPU Transcoder (Results)

- **Offline Encoder**: 2013 seconds
- **Concurrent Live Channels**: 380 seconds
- **Live Encoder**: 14 HD Channels 96 Encode

**5x**
For a single instruction, CPU requires 10 times more energy.

Cost Analysis

To get the same transcoding power…

Cost

$5x$

CPU

GPU

Cost

$
Final Products

Live Encoder
- Flexible Source
- Easy Management
- 20 HD + MultiPlatform Encoding
- Scalability
- Multi Channel-Multi Output

VOD Encoder
- Integration with other systems
- N+1 , N+M Redundancy
- Faster Encoding
- Streaming Anywhere
- High Extensibility
- Picture in Picture

Encoder SDK
- C++ .NET SOAP API
- Easy Development
- N+1 , N+M Redundancy
- Streaming Anywhere
- High Extensibility
- Picture in Picture

Green Technology
- DRM Encryption

Dr. Advertise Insertion
- Mosaic Screen

erlab
Summary

- With Erlab’s newest GPU based solution:
  - Encode 14 live HD channels concurrently
  - Execute 5 times faster
  - Decrease latency
  - OPEX: Save 67% of energy
  - CAPEX: Save 80% of Investment cost
Thanks...

Erlab’s growth awarded by Deloitte