Making Systems Faster: Distributed Video Processing

COS 418/518: Distributed Systems
Lecture 20

Wyatt Lloyd

[Grey slides from Qi Huang’s SOSP 2017 Talk]
Distributed Video Processing Outline

• Motivation for video processing
  • (How streaming video works)

• Legacy design

• SVE design

• Why SVE is faster than legacy
SVE: Distributed Video Processing at Facebook Scale

Qi Huang

Petchean Ang, Peter Knowles, Tomasz Nykiel, Iaroslav Tverdokhlib, Amit Yajurvedi, Paul Dapolito IV, Xifan Yan, Maxim Bykov, Chuen Liang, Mohit Talwar, Abhishek Mathur, Sachin Kulkarni, Matthew Burke, Wyatt Lloyd

Facebook, University of Southern California, Cornell, Princeton
Video is growing across Facebook

- **FB:** 500M users watch **100M hours** video daily (Mar. 16)
- Instagram: **250M** daily active users for stories (Jun. 17)
- All: **many tens of millions** of daily uploads, **3X** NYE spike
Processing is diverse and demanding

Pt. 1
Legacy System Scaling Challenges

Pt. 2
SVE Impact of Design

Input video

Processing

Re-encoding

Thumbnail

Video Classification

02
Legacy: upload video file to web server
Legacy: preserve original for reliability
Legacy: process after upload completes

Client → Web Server → Original Storage → Processing

She is having so much fun with #MSQRD
Legacy: encode w/ varying bitrates

Client → Web Server → Original Storage → Processing

- **720P**: 4Mbps
- **1080P**: 16Mbps
- **480P**: 1.5Mbps
Legacy: store encodings before sharing

Client

Web Server

Original Storage

Processing

Final Storage

- 1080P 16Mbps
- 720P 4Mbps
- 480P 1.5Mbps

- 1695x738
- 890x791
- 891x742
- 471x757
- 107x573
- 131x757

She is having so much fun with #MSQRD
Sharing with adaptive streaming

Client

Web Server
720p
480p

FBCDN

Final Storage
Focus: pre-sharing pipeline

All steps from when a user starts an upload until a video is ready to be shared
How Long Does This Take? (Latency)

Client → Web Server → Original Storage → Processing → Final Storage
How Long Does This Take? (Latency)

1 MB Video \approx 1 \text{ secs}  
8 Mbps link

16 MB Video \approx 16 \text{ secs}  
1 Mbps link

**SVE paper stats:**

- **Video Size**
  - \leq 1\text{MB} \quad 10\% \text{ of uploads over 10 seconds}
  - 3-10\text{MB} \quad 50\% \text{ of uploads over 10 seconds}
  - 300\text{MB} \quad 50\% \text{ of uploads over 9 minutes}
  - 1\text{GB}
How Long Does This Take? (Latency)

Web Server
How Long Does This Take? (Latency)

(pipelined with uploading)

SVE paper stats:

<table>
<thead>
<tr>
<th>Stat</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>median</td>
<td>200 ms</td>
</tr>
<tr>
<td>90%</td>
<td>650 ms</td>
</tr>
<tr>
<td>99%</td>
<td>900 ms</td>
</tr>
</tbody>
</table>
How Long Does This Take? (Latency)

SVE paper stats:

10% of all video take ≥ 1.3 s

Proportional to video size:

Most videos over 100 MB take over 6 seconds
How Long Does This Take? (Latency)

SVE paper stats:

<table>
<thead>
<tr>
<th>Video Size</th>
<th>Percentage Taking Over</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-3MB</td>
<td>20% take over 10 seconds</td>
</tr>
<tr>
<td>100-300MB</td>
<td>50% take over 1 minute</td>
</tr>
<tr>
<td>&gt;1GB</td>
<td>23% take over 10 minutes</td>
</tr>
</tbody>
</table>
Serial pipeline leads to slow processing
Let’s Make This Faster!

Discuss with your neighbors!
Speedy: harness parallelism

Users can share videos quickly

- Overlap fault tolerance and processing
- Overlap upload and processing
- Parallel processing
Architectural changes for parallelism
Architectural changes for parallelism

- Client
- Web Server
- Preprocessor
  - Scheduler
  - Worker
  - Worker
  - Worker
- Final Storage
- Original Storage
Overlap fault tolerance and processing

Client → Web Server → Write-through Cache → Scheduler → Worker → Final Storage

Original Storage
Overlap upload and processing

Client → Web Server → Preprocessor → Worker → Final Storage

Split into segments

Scheduler

Worker

Worker

Worker

Original Storage
Overlap upload and processing

Client → Web Server → Preprocessor → Worker → Final Storage

...upload in progress

Scheduler
Parallel processing w/ many workers

Client → Web Server → Preprocessor → 720P Encode → Final Storage

...upload in progress

Client

Web Server

Preprocessor

Original Storage

Scheduler

720P Encode

480P Encode

Thumbnail

Final Storage

20
Parallel processing w/ many workers

- Client
- Web Server
- Preprocessor
- Original Storage
- 720P Encode
- 480P Encode
- Thumbnail
- Final Storage

...upload in progress
Parallel processing w/ many workers

Client → Web Server → Preprocessor → Original Storage → Final Storage

...upload in progress

Scheduler

720P Encode

480P Encode

Thumbnail
Parallel processing w/ many workers
Three sources of parallelism

Overlap fault tolerance and processing
Overlap upload and processing
Parallel processing
Let’s Make This Faster!

SVE is enabling these options while still supporting lap uploading and encoding. One major challenge we over-tensity on pre-sharing latency: 1) upload less data, and 2) over-us with two options for decreasing the effect of upload la-
available to the client, which we cannot improve. This leaves
tively. This demonstrates that upload time is a significant part

take 1 minute, 3 minutes, 9 minutes, and 28 minutes, respec-
300 MB–1 GB, and

50%. For the large size classes of 30–100 MB, 100–300 MB,
more than 10 seconds. For the 3–10 MB size class, the per-
est size class (ure
video is a significant part of the pre-sharing latency. Fig-

The time required for a client to upload all segments of a
day period in June 2017 unless otherwise specified.
over MES (§
quantifies the latency improvement that the SVE provides
why these choices lead to lower latency. This section also
latency for the MES and SVE designs to provide intuition for

and processing (§
in parallel (§
ping uploading and processing (§
section describes how SVE provides low latency by overlap-

store

A

B

C

D

E

Pre-Sharing Latency

Legacy

Store

Upload

Process

SVE

Store

Upload

Process

A

B

C

D

E
Results: 2.3x ~ 9.3x speedup

Relative speedup vs. Video size buckets:

- < 3M: 2.3x
- 3M ~ 10M: 3.0x
- 10M ~ 100M: 3.7x
- 100M ~ 1G: 6.1x
- > 1G: 9.3x
Results: 2.3x ~ 9.3x speedup

Overlap upload & processing

Relative speedup vs. Video size buckets:
- < 3M: 2.3x
- 3M ~ 10M: 3x
- 10M ~ 100M: 3.7x
- 100M ~ 1G: 6.1x
- >1G: 9.3x
Results: 2.3x ~ 9.3x speedup

Parallel Processing

Video size buckets:
- <3M: 2.3x
- 3M ~ 10M: 3x
- 10M ~ 100M: 3.7x
- 100M ~ 1G: 6.1x
- >1G: 9.3x

Relative speedup
Summary

- Motivation for video processing
  - (How streaming video works)

- Legacy design – Serial processing was slow

- SVE design – Three sources of parallelism make SVE faster
  - Overlap upload and processing
  - Overlap fault tolerance and processing
  - Parallel processing