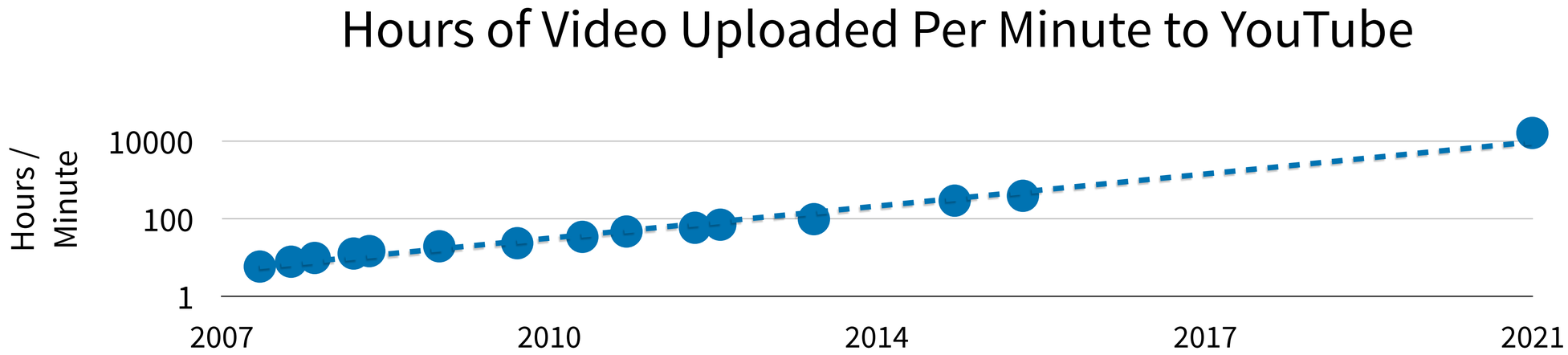
Perceptual Compression for Video Storage and Processing Systems

Amrita Mazumdar, Brandon Haynes, Magda Balazinska,
Luis Ceze, Alvin Cheung*, Mark Oskin

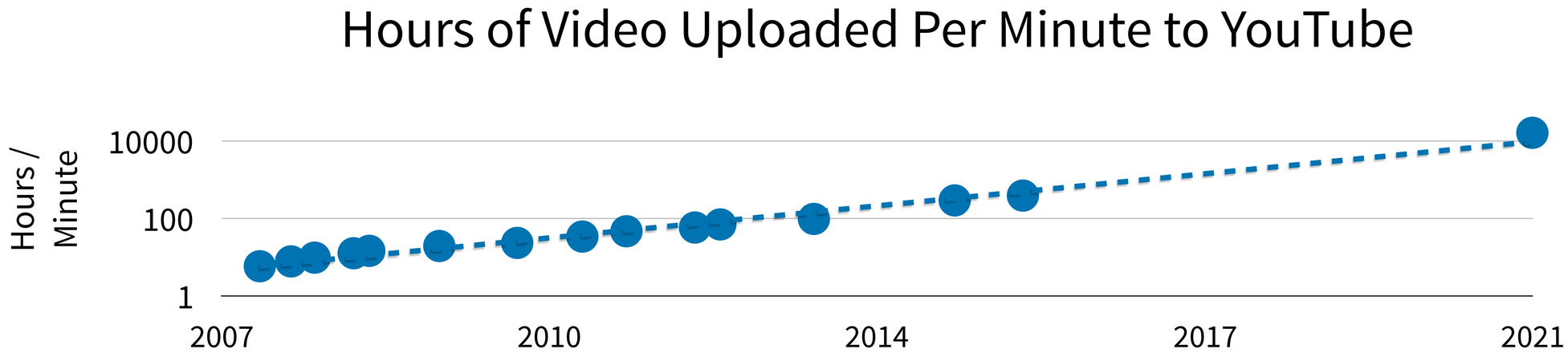
University of Washington
**University of California, Berkeley*

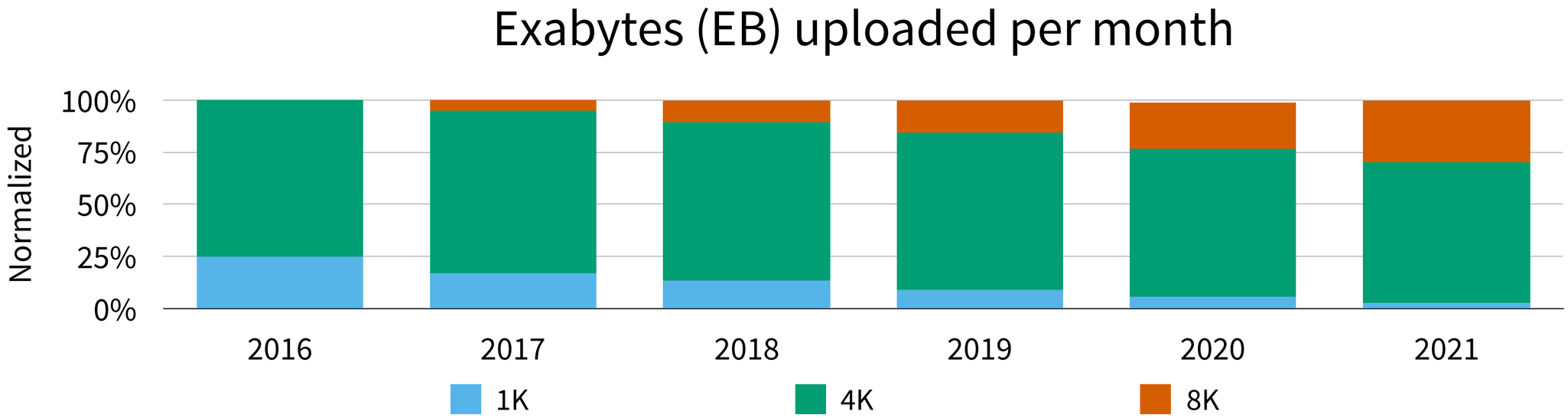
PAUL G. ALLEN SCHOOL

Video data is a fast-growing form of digital content



Video data is a fast-growing form of digital content





Video applications use compression to trade visual redundancy for file size

Baseline codec (HEVC) @ 20 Mbps
4 hours video playback

Video applications use compression to trade visual redundancy for file size

fine details (noise, high frequencies)

color perception

fast motion

Baseline codec (HEVC) @ 20 Mbps
4 hours video playback

This work: Integrating new perceptual cues with video distribution systems for reduced video sizes

new perceptual cues!

Baseline HEVC @ 20 Mbps
4 hours video playback

This work: 1 Mbps
6.5 hours video playback

Saliency is a powerful perceptual cue for reducing video size

4K 360° video
300 MB

AI-generated saliency map
only 15% of pixels are important

Saliency is a powerful perceptual cue for reducing video size

Saliency is one of many new perceptual metrics

How can video services take advantage of them?

Can we leverage perception in cloud-scale video services for better storage and performance?

AI-generated saliency map
only 15% of pixels are important

Leveraging perceptual cues at scale is difficult

Requires custom, outdated codecs

Recent work:

- outdated codec
- 1,500 lines of code
- already worse by newer codecs w/o saliency compression

Leveraging perceptual cues at scale is difficult

Requires custom, outdated codecs

No integration with storage manager

Video storage manager concerns:

- How big is a file?
- How can I compress it?
- How do I control quality?

Leveraging perceptual cues at scale is difficult

Requires custom, outdated codecs

No integration with storage manager

No interface for applications

Video app concerns:

- Can I process this like any other video?
- Do I have to re-package this video for standard players?

Leveraging perceptual cues at scale is difficult

Requires custom, outdated codecs

No integration with storage manager

No interface for applications

Goals

✅ Modern codecs

✅ API for storage

✅ Application portable

Vignette: a system for perceptual compression and storage

Vignette Compression

codec-free perceptual video compression framework

Vignette Storage

storage manager for perceptually-compressed videos

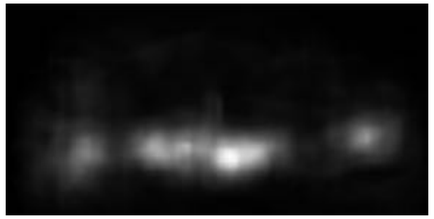
Reduces storage by up to 75-95% with little quality loss

Resulting videos use 50% less power on mobile phones

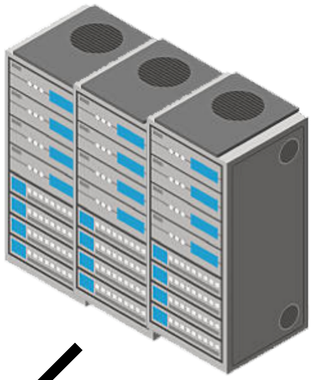
Vignette system architecture

1. Upload Video

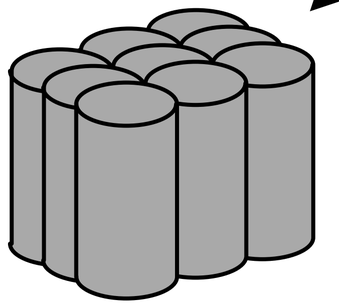
2. Produce ML-generated saliency map



3. Execute saliency-based compression



4. Write compressed video to storage manager



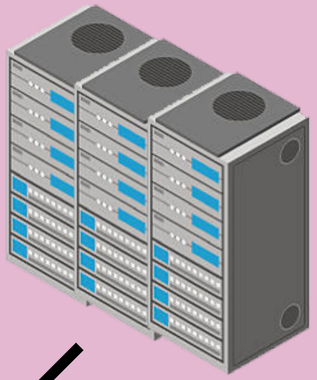
5. Stream compressed video to user

Vignette system architecture

1. Upload Video

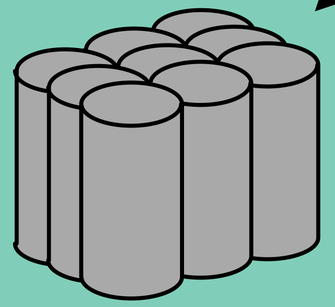
2. Produce ML-generated saliency map

3. Execute saliency-based compression



Vignette Compression

Vignette Storage



4. Write compressed video to storage manager

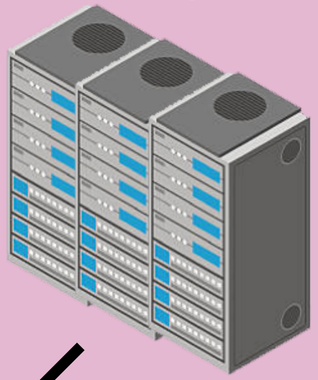
5. Stream compressed video to user

Vignette system architecture

1. Upload Video

2. Produce ML-generated saliency map

3. Execute saliency-based compression



Vignette Compression

Vignette Storage

4. Write compressed video to storage manager

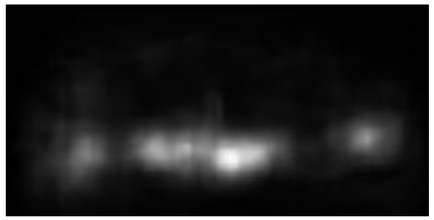
5. Stream compressed video to user

Vignette Compression

ingest video

Vignette Compression

generate perceptual map for the video



ingest video

Vignette Compression

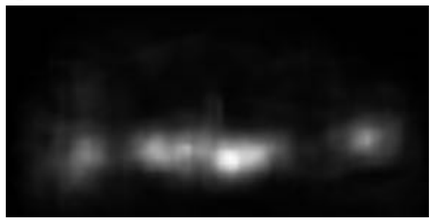
ingest video

generate perceptual map for the video

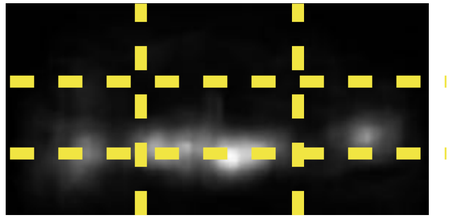
- **high-accuracy, compute-intensive neural network**
- **designed to easily switch to new models or perceptual metrics**

Vignette Compression

generate perceptual map for the video



ingest video



generate tiling configuration

Vignette Compression

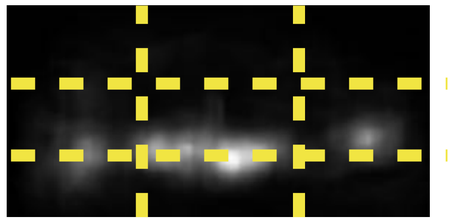
ingest video

Tiles:

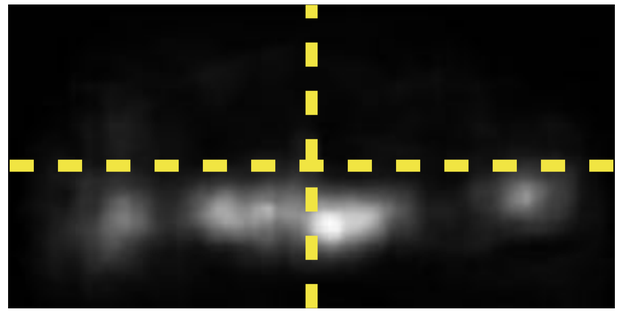
-codec-standard feature

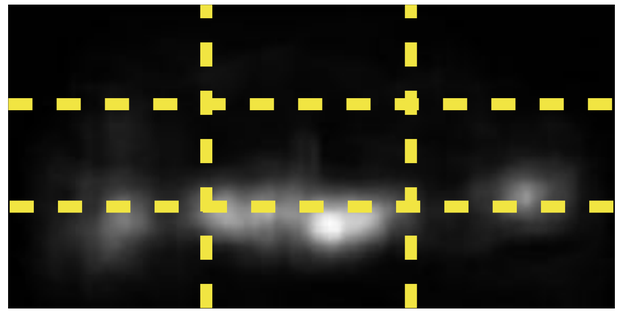
-divide video into smaller regions (some overhead)

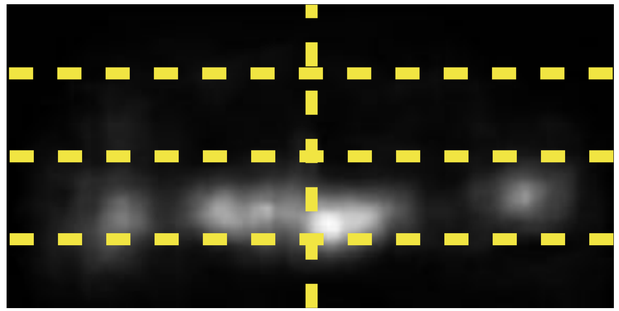
- drive quality per-tile



generate tiling configuration





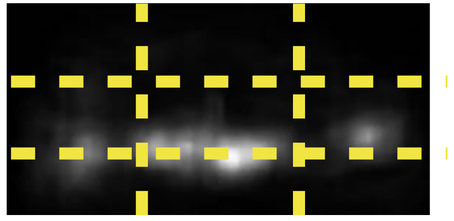


Vignette Compression

generate perceptual map for the video

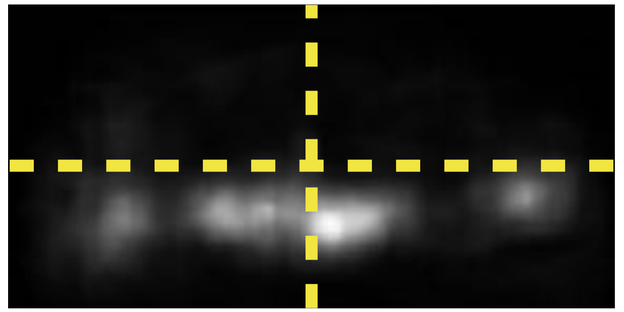
ingest video

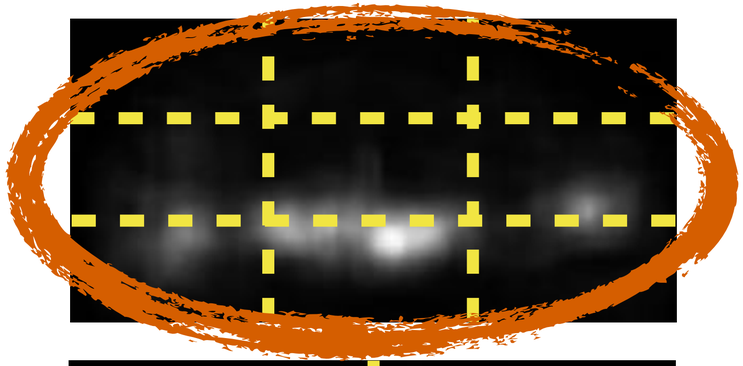
pick best quality, lowest overhead

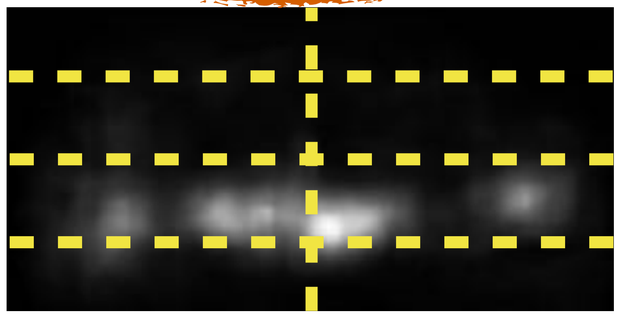


generate tiling configuration

exhaustive search process

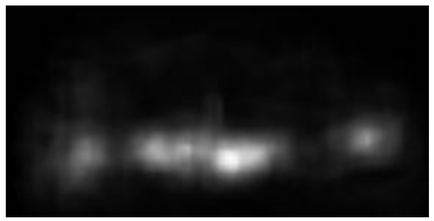




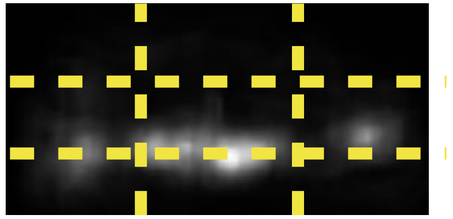


Vignette Compression

generate perceptual map for the video

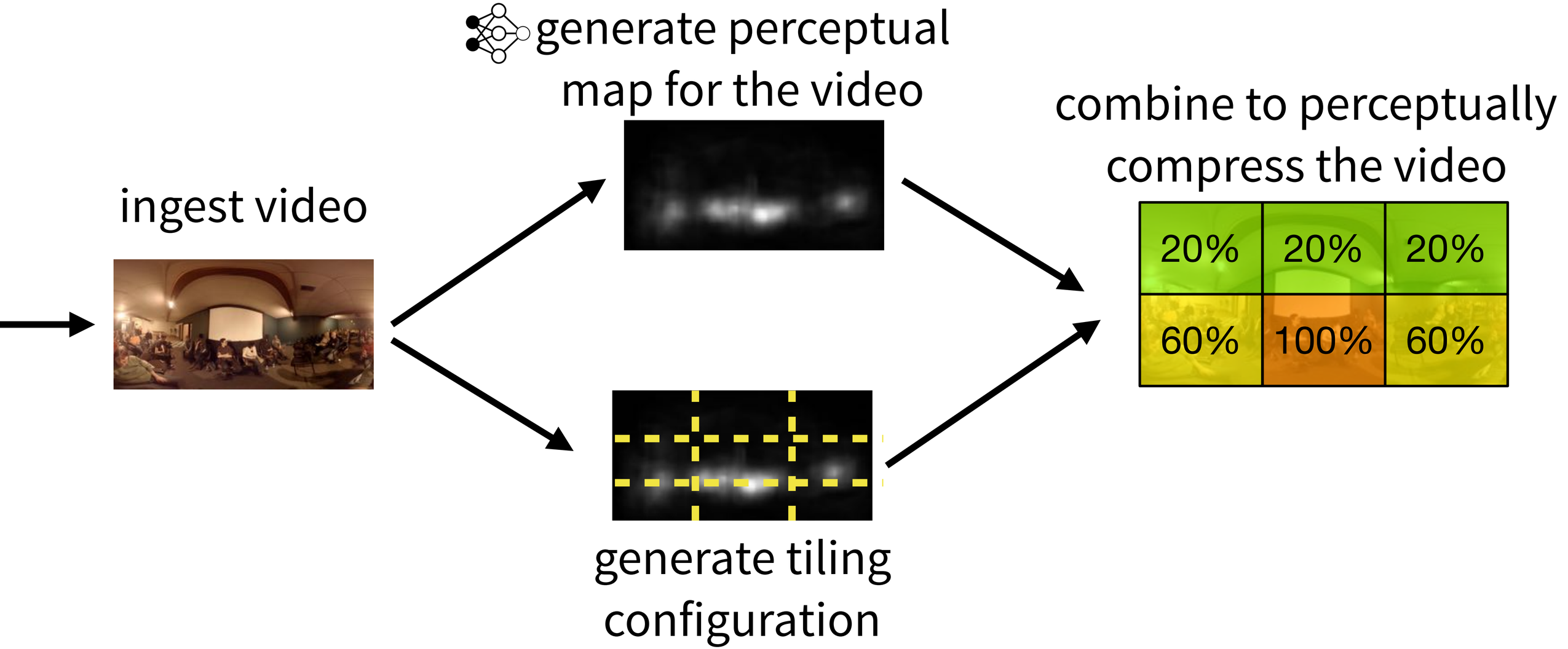


ingest video



generate tiling configuration

Vignette Compression

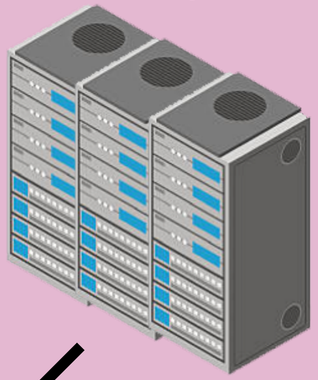


Vignette system architecture

1. Upload Video

2. Produce ML-generated saliency map

3. Execute saliency-based compression



Vignette Compression

Vignette Storage

4. Write compressed video to storage manager

5. Stream compressed video to user

Vignette system architecture

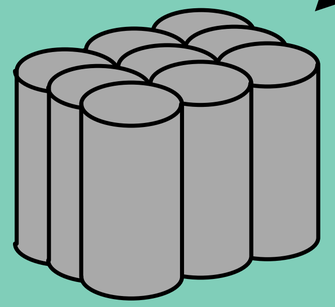
1. Upload Video

2. Produce ML-generated saliency map

3. Execute saliency-based compression

Vignette Compression

Vignette Storage



4. Write compressed video to storage manager

5. Stream compressed video to user

Vignette Storage uses compressed metadata to communicate tile configurations

up to 1.75x compressed video size

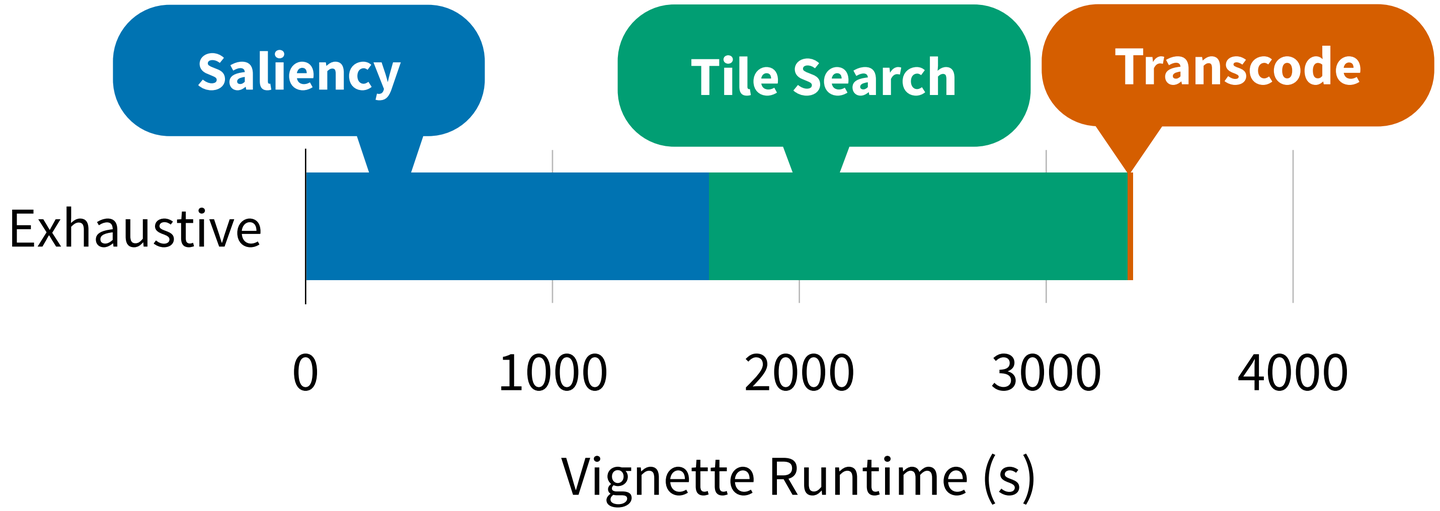
Vignette Storage uses compressed metadata to communicate tile configurations

up to 1.75x compressed video size

this work: video + low-overhead metadata

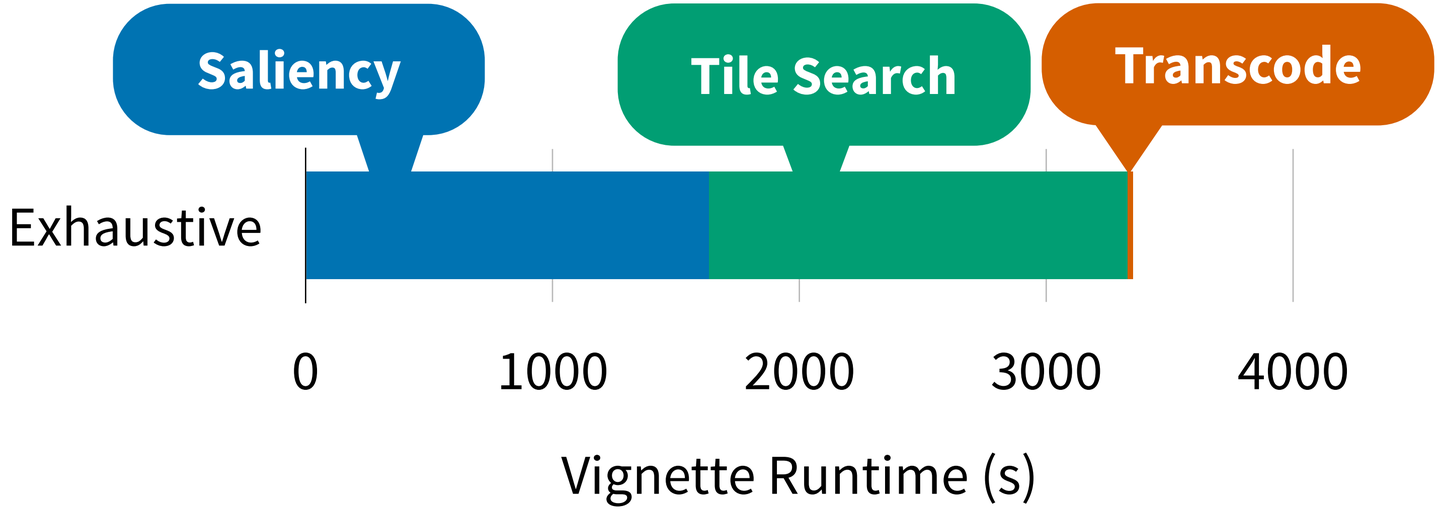
metadata: tile config & qualities (8-100 B)

Vignette Storage provides a faster heuristic search to find tile configurations



Vignette Storage provides a faster heuristic search to find tile configurations

motion vectors in video

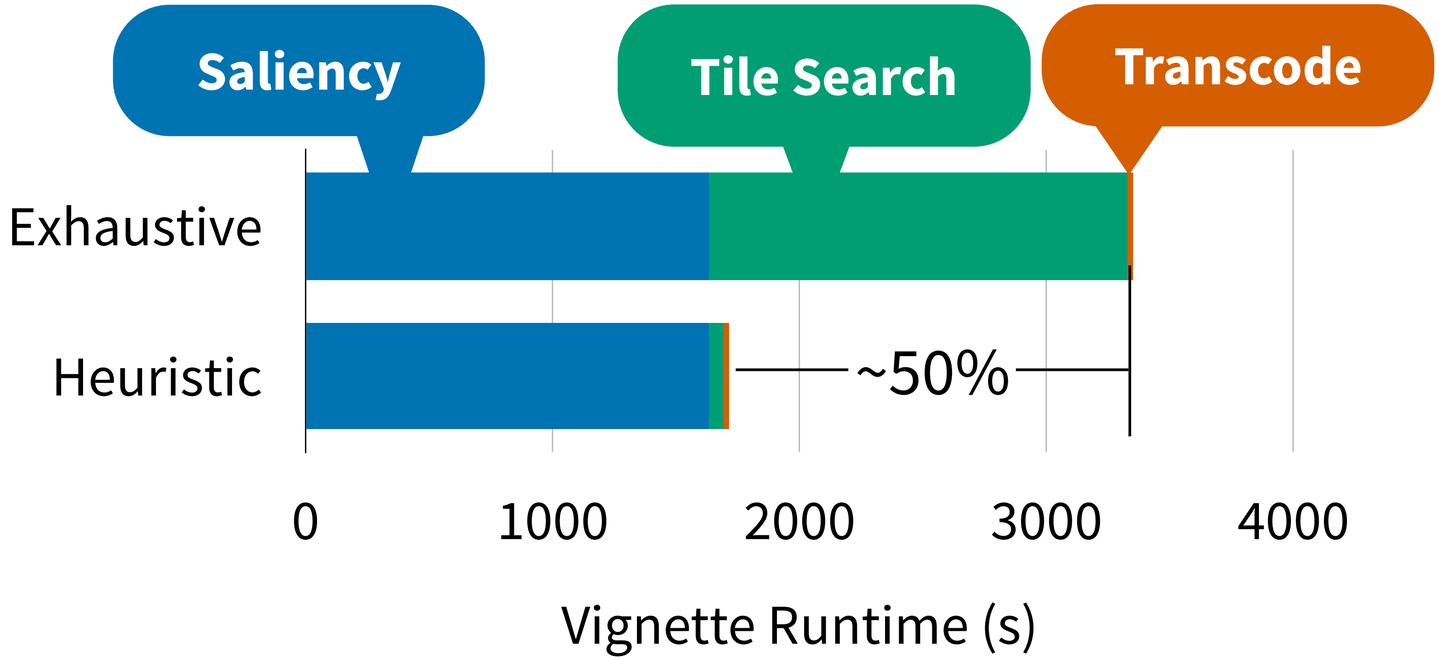


New algorithm using motion vectors to guide tile size

Key Insight: splitting motion vectors across tiles results in larger videos

Vignette Storage provides a faster heuristic search to find tile configurations

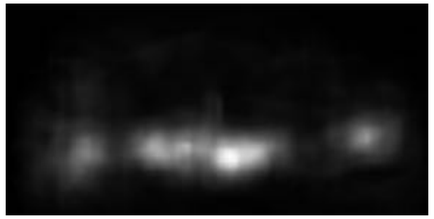
motion vectors in video



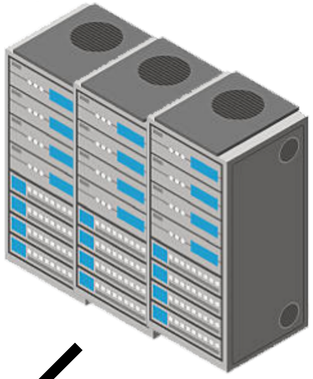
Vignette system architecture

1. input video

2. ML-generated saliency map

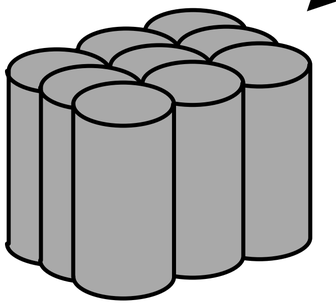


3. saliency-based compression



more details in the paper!

4. storage manager



5. streaming to user, (optional) eye-tracker feedback

Evaluation & Results

Baseline HEVC @ 20 Mbps
4 hours video playback

Vignette @ 1 Mbps
6.5 hours video playback

Evaluation & Results

Compression Performance and Quality

How small are Vignette videos?

How much quality loss is there?

Cloud Performance

Is Vignette practical to use in the cloud?

How does Vignette affect TCO?

Evaluation & Results

Compression Performance and Quality

How small are Vignette videos?

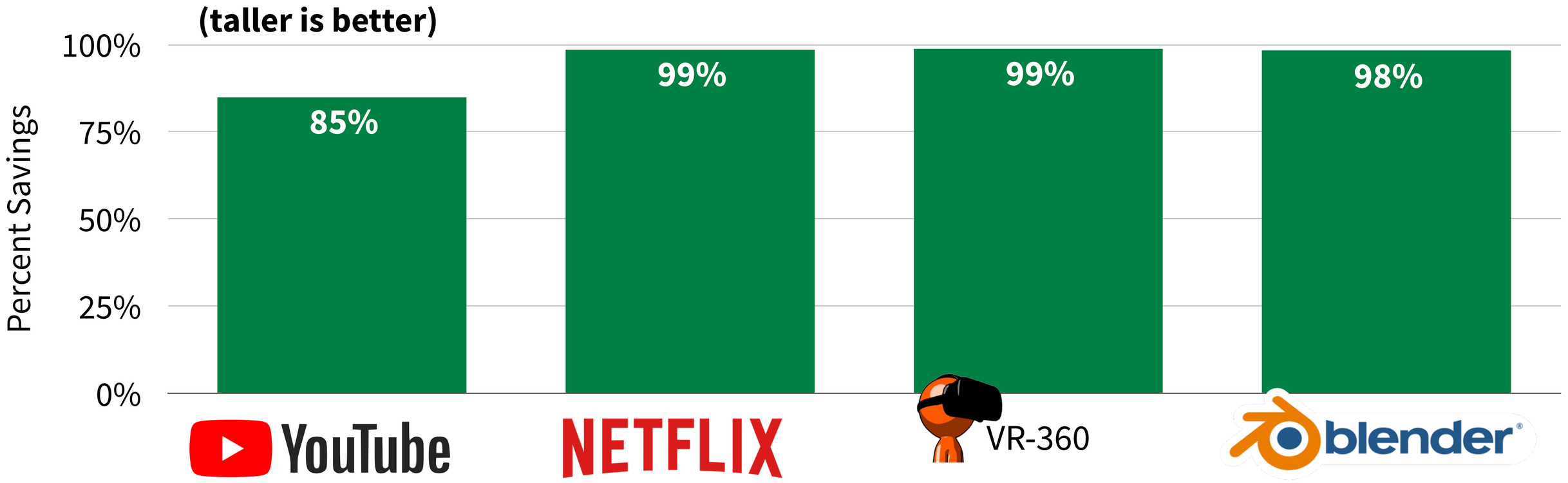
How much quality loss is there?

Cloud Performance

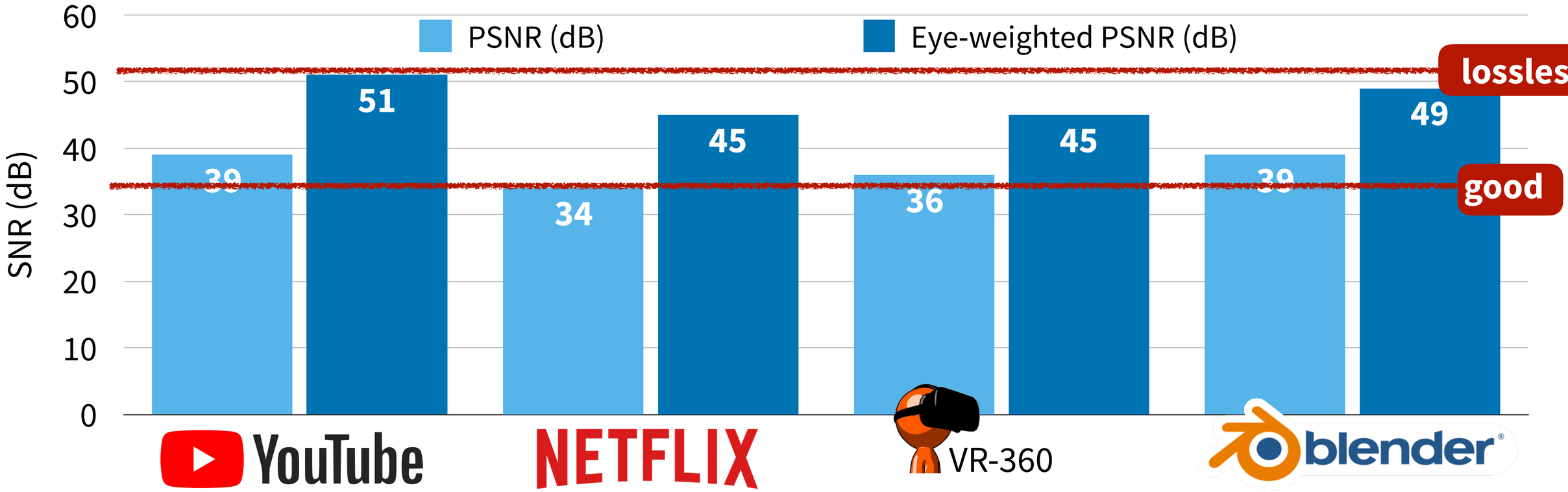
Is Vignette practical to use in the cloud?

How does Vignette affect TCO?

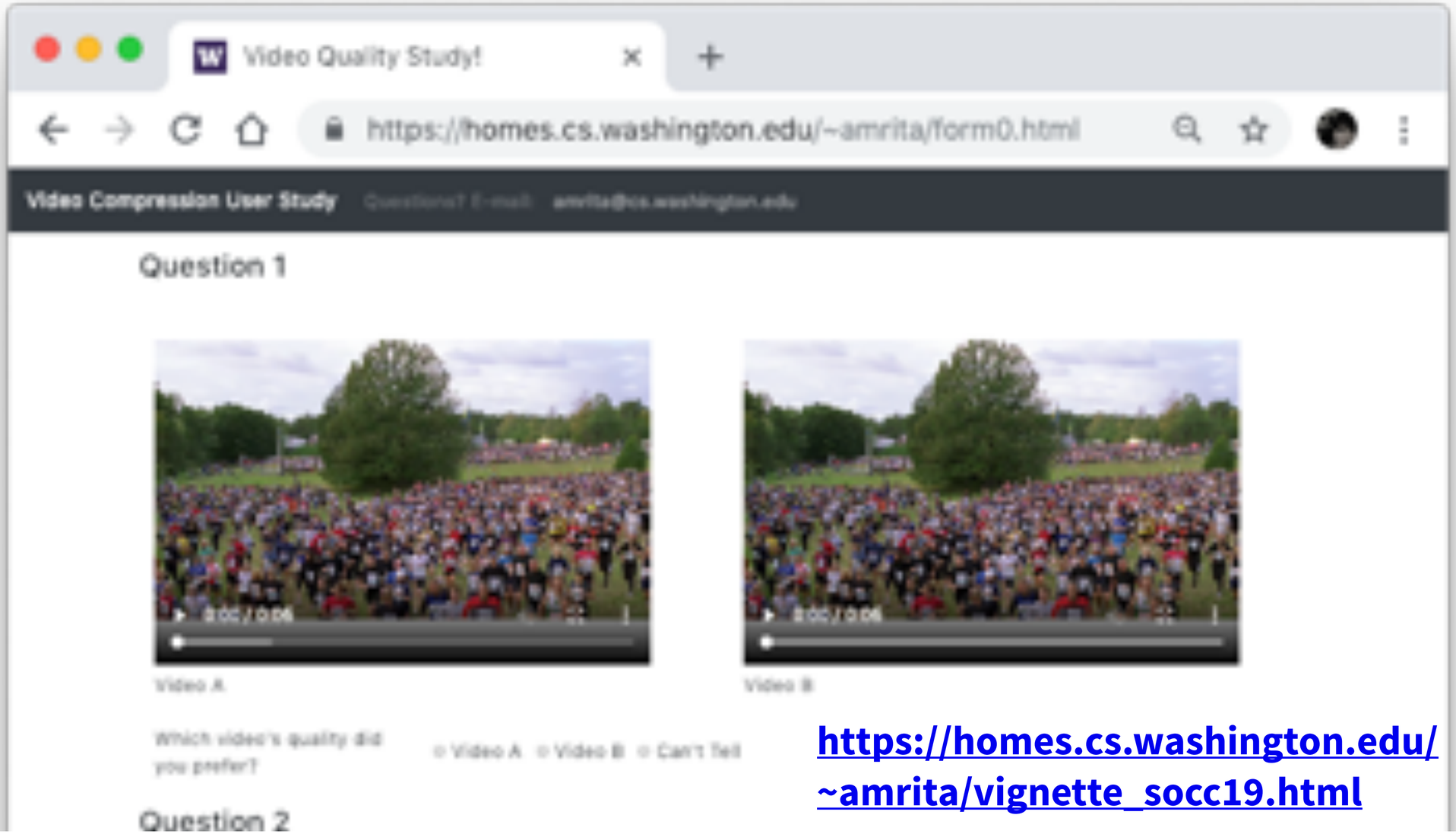
Vignette Storage Savings



Vignette Quality

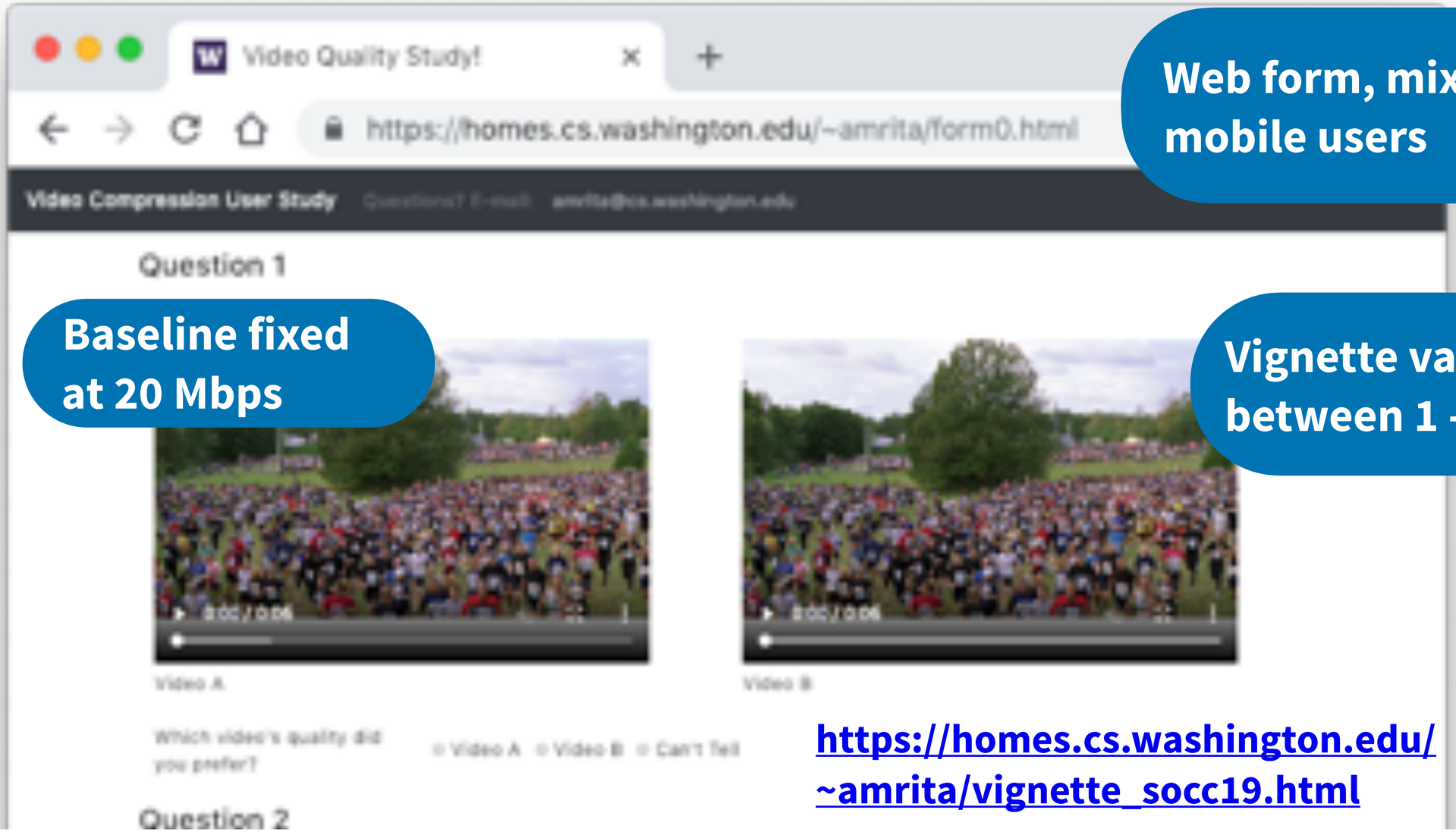


Vignette Perceptual Quality User Study



https://homes.cs.washington.edu/~amrita/vignette_socc19.html

Vignette Perceptual Quality User Study



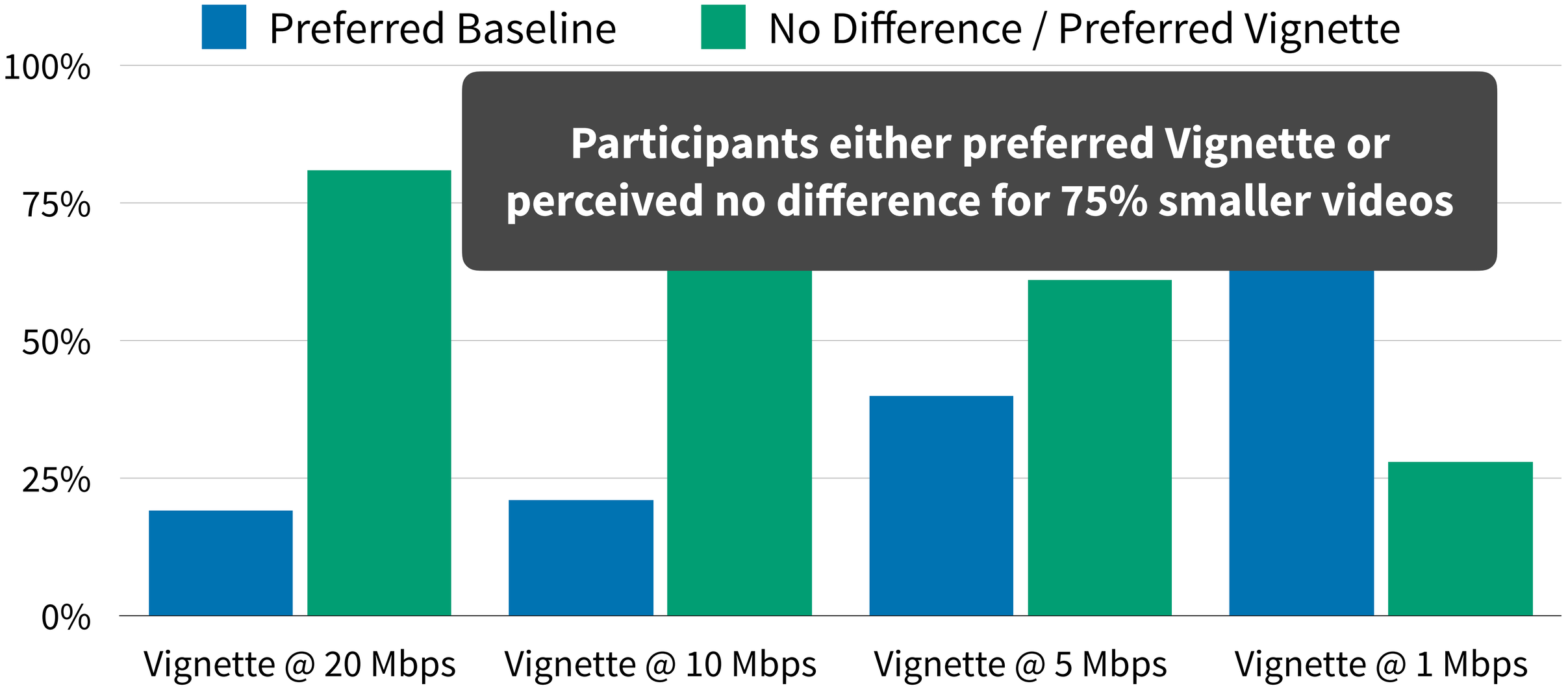
Web form, mixed desktop/ mobile users

Baseline fixed at 20 Mbps

Vignette varied between 1 - 20 Mbps

https://homes.cs.washington.edu/~amrita/vignette_socc19.html

Vignette Perceptual Quality User Study



Evaluation & Results

Compression Performance and Quality

Vignette videos are 85% smaller while maintaining quality.

For 75% smaller videos, users did not notice quality impact.

Cloud Performance

Is Vignette practical to use in the cloud?

How does Vignette affect TCO?

Evaluation & Results

Compression Performance and Quality

Vignette videos are 85% smaller while maintaining quality.

For 75% smaller videos, users did not notice quality impact.

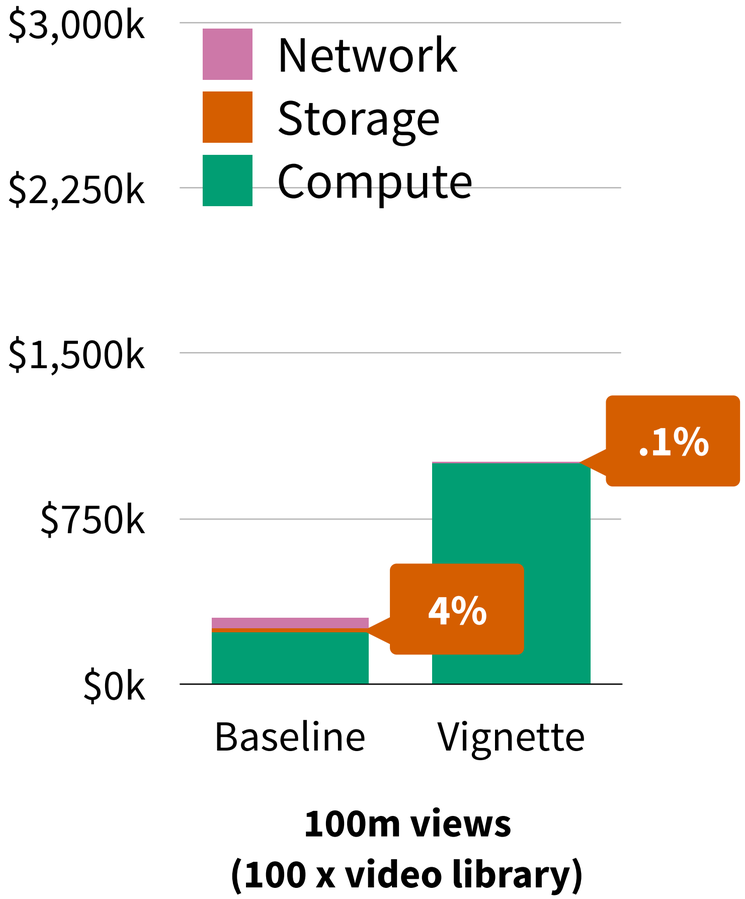
Cloud Performance

Is Vignette practical to use in the cloud?

How does Vignette affect TCO?

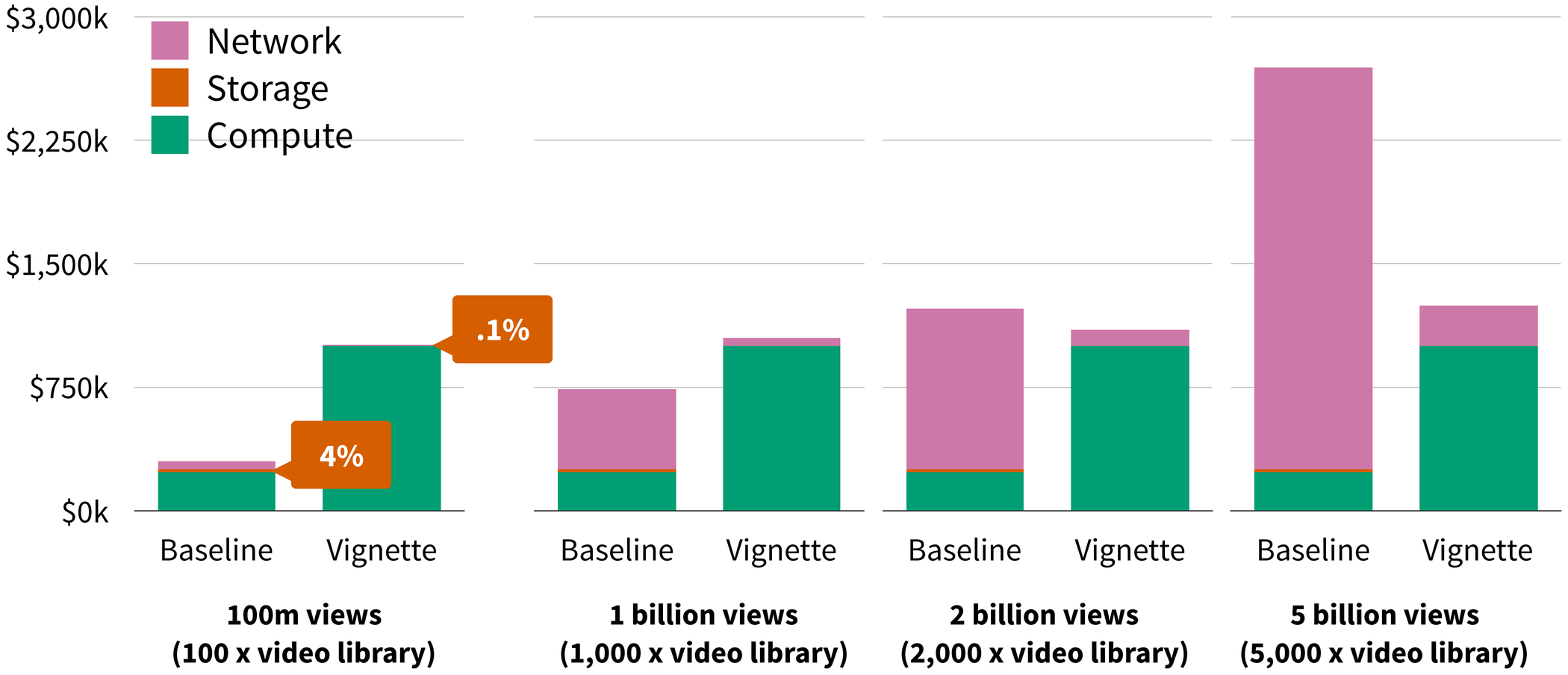
TCO analysis on AWS datacenter

storing and streaming a 1-million video library



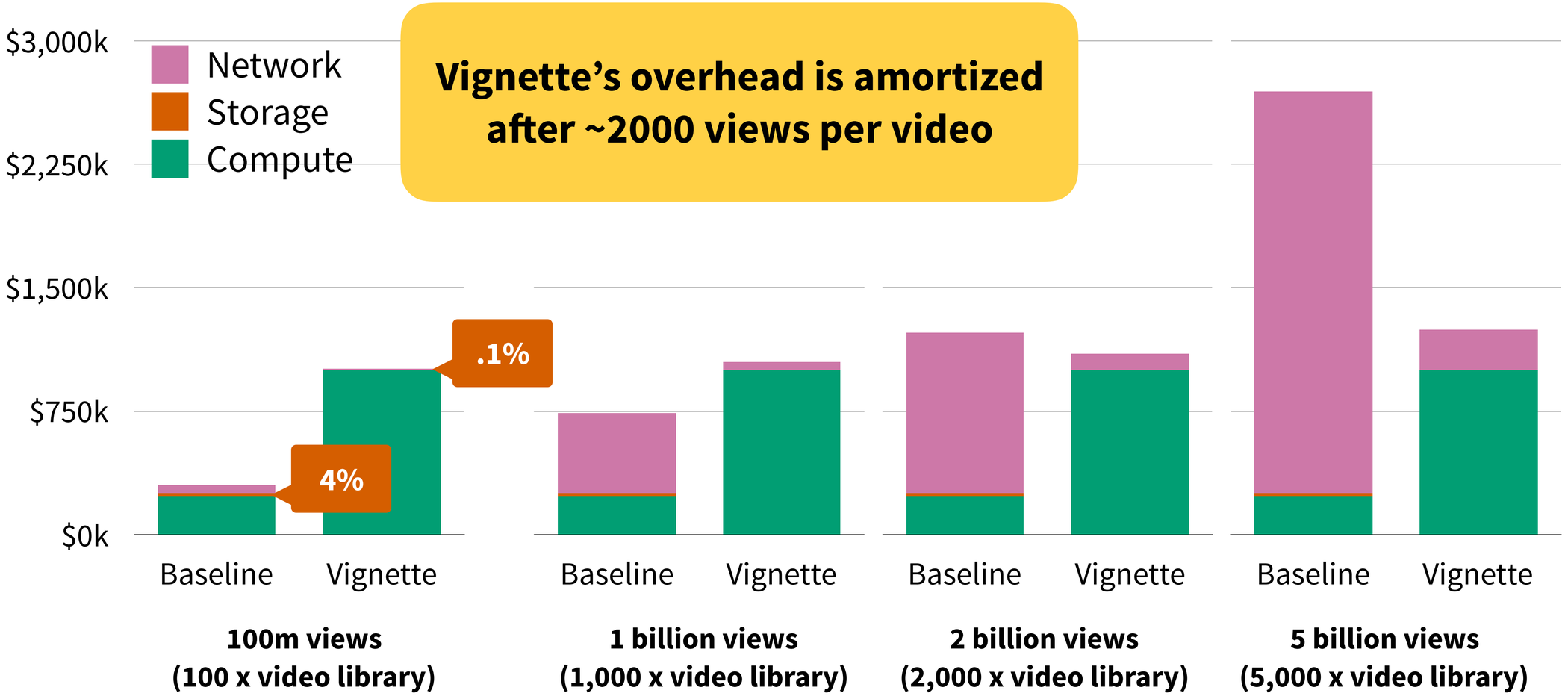
TCO analysis on AWS datacenter

storing and streaming a 1-million video library

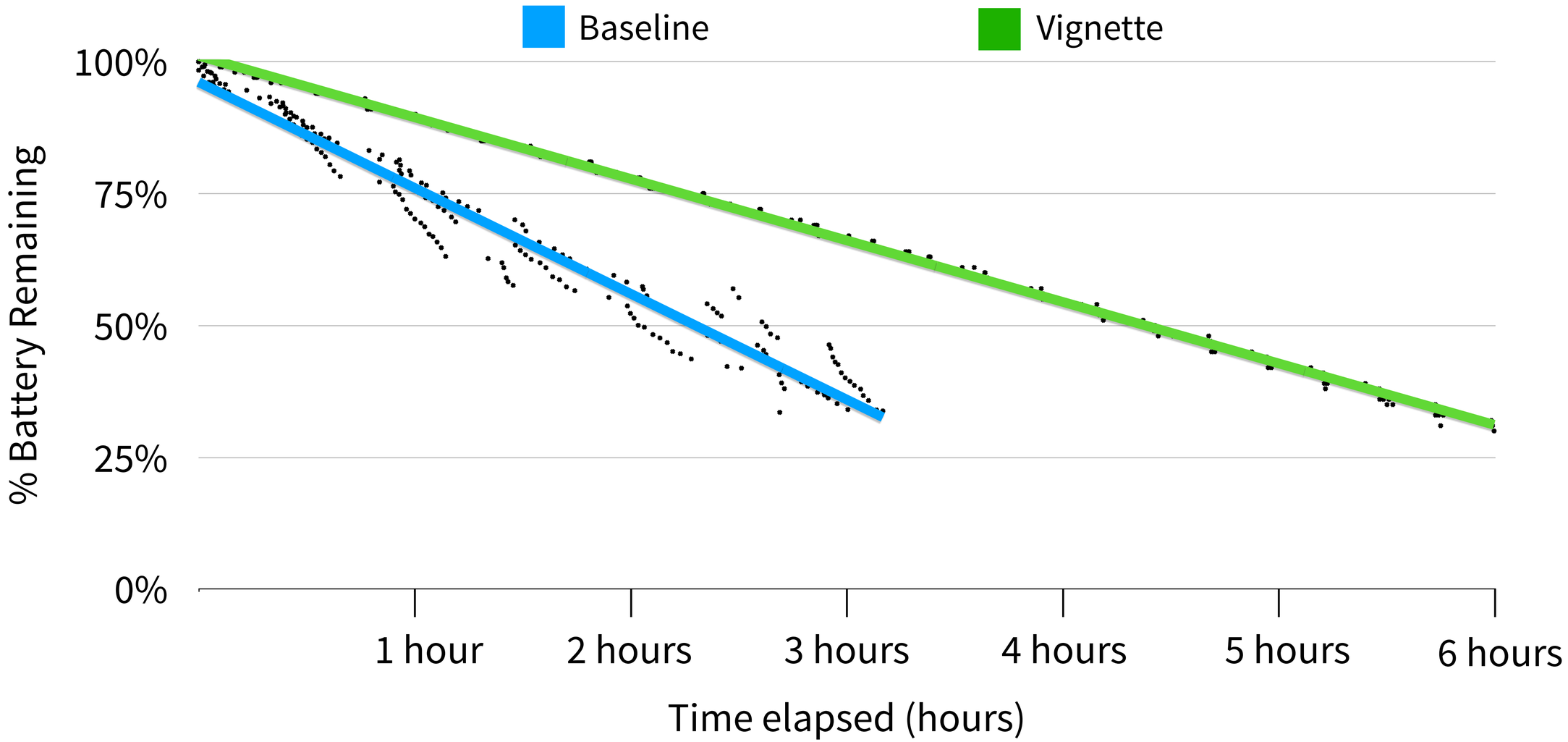


TCO analysis on AWS datacenter

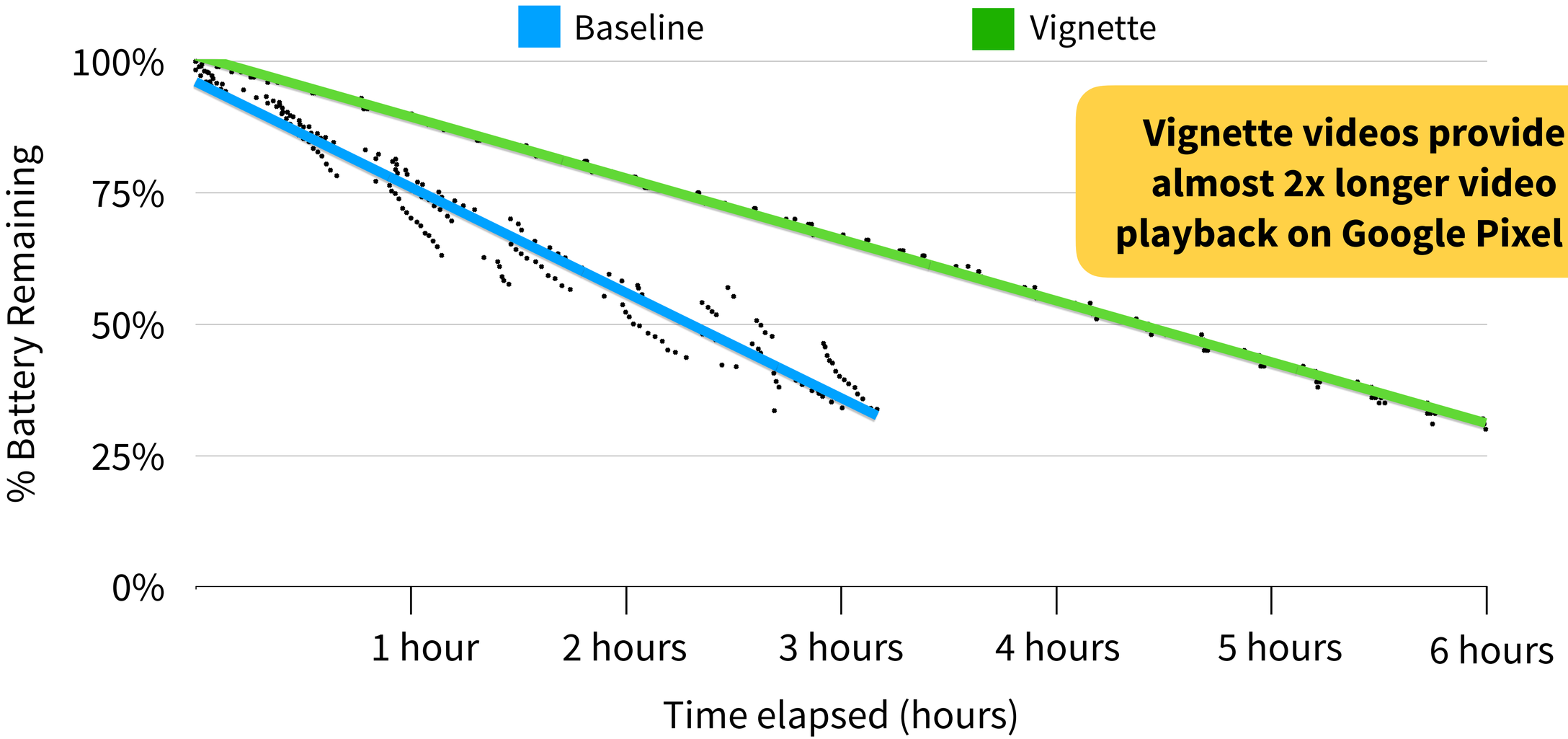
storing and streaming a 1-million video library



Results: Mobile Power Analysis



Results: Mobile Power Analysis



Vignette videos provide almost 2x longer video playback on Google Pixel 2

Evaluation & Results

Compression Performance and Quality

Vignette videos are 85% smaller while maintaining quality.
For 75% smaller videos, users did not notice quality impact.

Cloud Performance

Vignette amortizes large transcode costs in ~2000 streams / video.
Vignette videos also reduce power during mobile playback.

Vignette: a system for perceptual compression and storage

Vignette Compression

codec-free perceptual video compression

Vignette Storage

storage manager for perceptually-compressed videos

Reduces storage by up to 75-95% with little quality loss

Resulting videos use 50% less power on mobile phones

Thank you!

https://homes.cs.washington.edu/~amrita/vignette_socc19.html

PAUL G. ALLEN SCHOOL