SERF:

Enabling Optimization of Socially Sourced Images using Psychovisual Enhancements

Aaron Koehl
Dept. of Physics, Computer Science, and Engineering
Christopher Newport University

Haining Wang
Dept. of Electrical and Computer Engineering
University of Delaware

MMSys'16 – Klagenfurt, Austria – May 12, 2016

Overview

- Introduction
- Social Media
- Lossy Compression
- Psychovisual Enhancements
- Experiment
- Results

What is Psychophysics?

> "Psychophysics seeks to quantitatively investigate the relationship between psychological events and physical events, or more specifically, between sensations and the stimuli that produce them."

—Britannica

TRADITIONAL WEB CLIENTS

Mobile Devices

Search Engine
Crawlers, Scrapers,
and Bots

NON-TRADITIONAL WEB CLIENTS

Client Side

Web Application Ecosystem

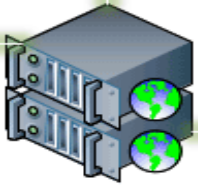
TRADITIONAL WEB CLIENTS

Mobile Devices

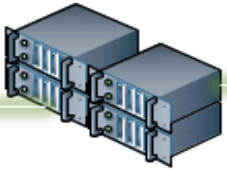
Search Engine
Crawlers, Scrapers,
and Bots

NON-TRADITIONAL WEB CLIENTS

Image Servers



Web Server / Proxies



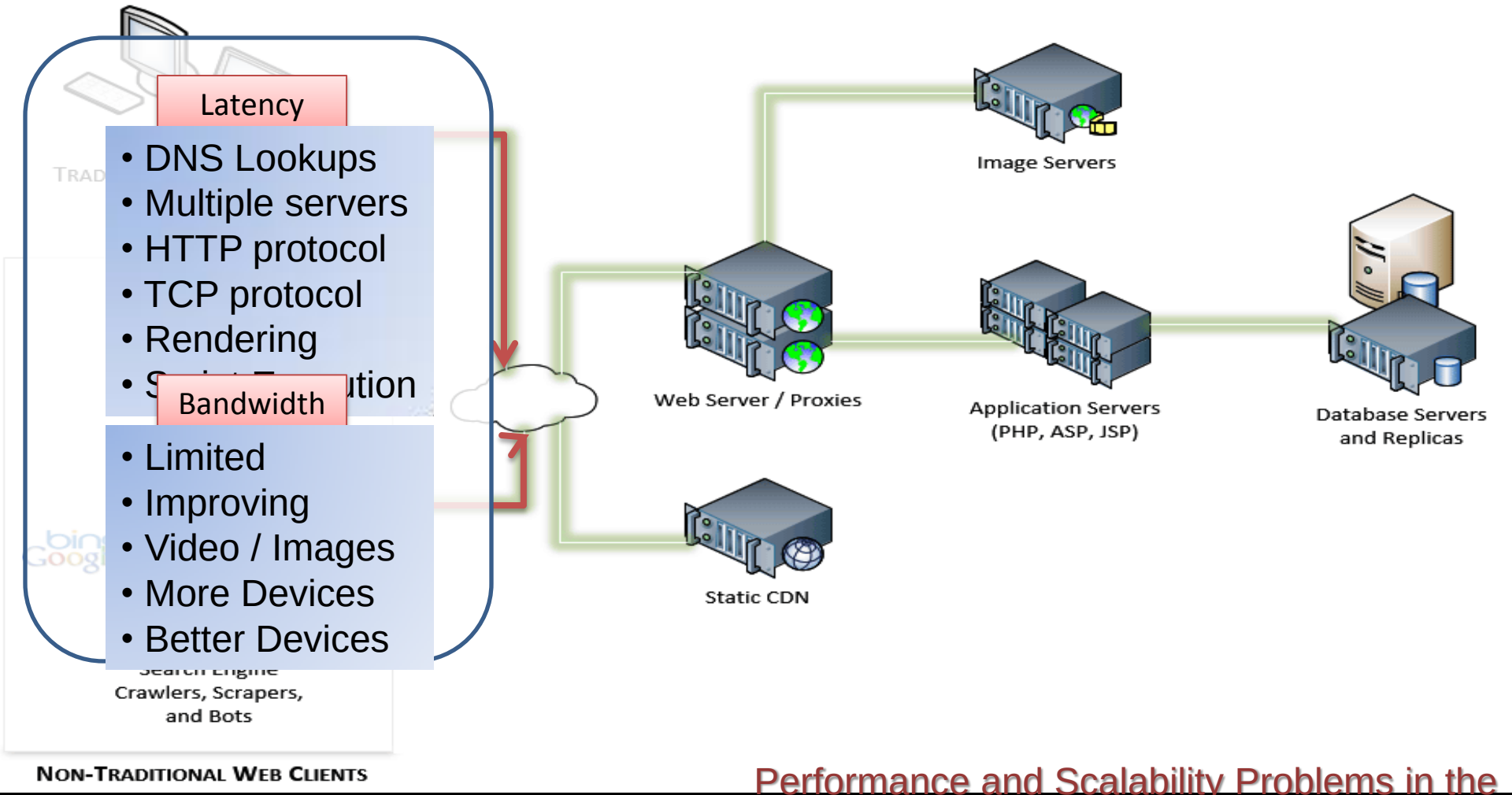
Application Servers
(PHP, ASP, JSP)

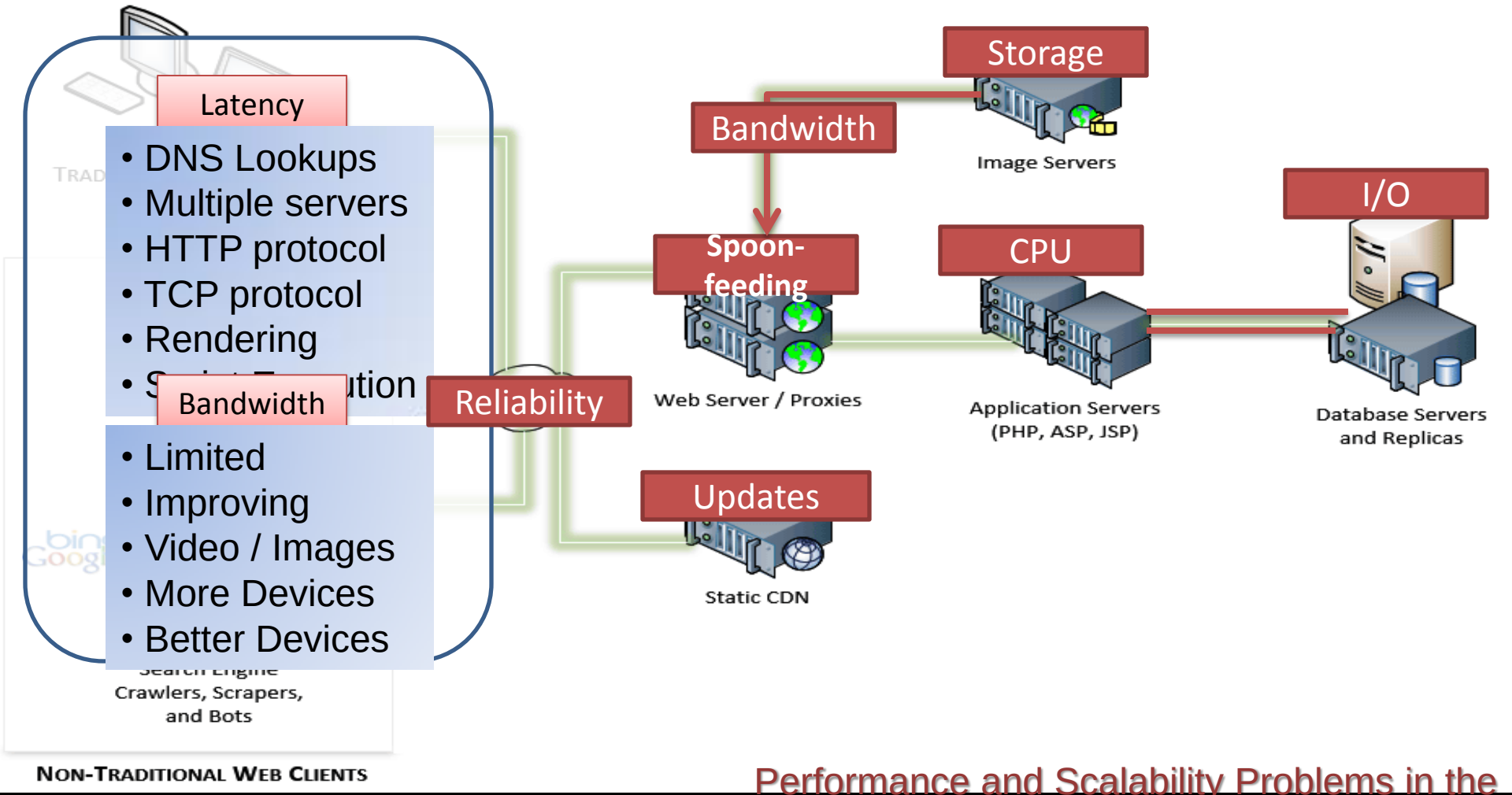
Database Servers
and Replicas

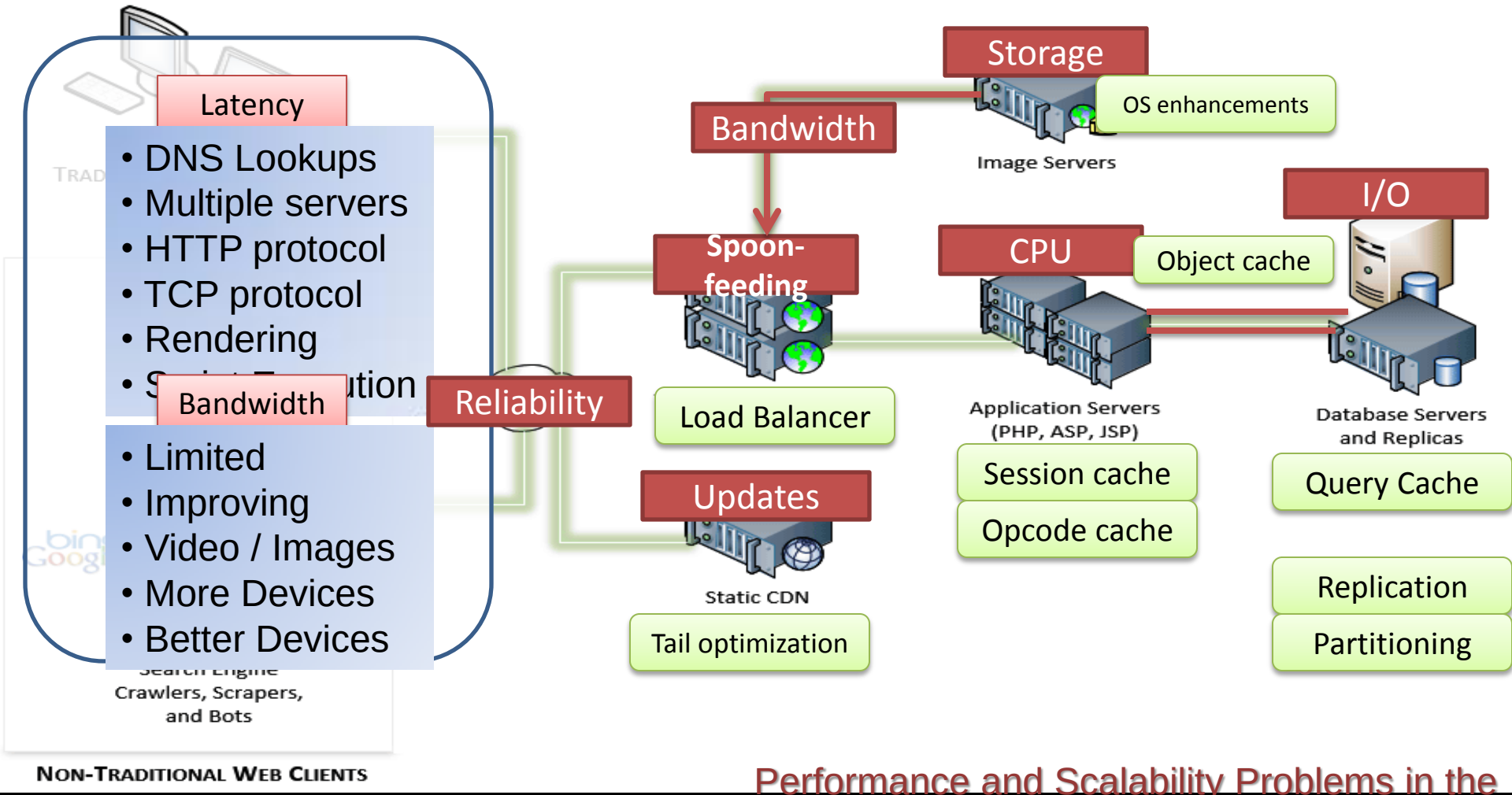
Static CDN

Server Side

Web Application Ecosystem


Latency
• DNS Lookups
• Multiple servers
• HTTP protocol
• TCP protocol
• Rendering
Bandwidth
• Limited
• Improving
• Video / Images
• More Devices
• Better Devices
Image Servers
Web Server / Proxies
Application Servers
(PHP, ASP, JSP)
Database Servers
and Replicas
Static CDN
Crawlers, Scrapers,
and Bots
NON-TRADITIONAL WEB CLIENTS


Latency
• DNS Lookups
• Multiple servers
• HTTP protocol
• TCP protocol
• Rendering
• Script Execution
Bandwidth
• Limited
• Improving
• Video / Images
• More Devices
• Better Devices
Crawlers, Scrapers, and Bots
Non-Traditional Web Clients
Reliability
Storage
Bandwidth
Image Servers
Spoon-feeding
Web Server / Proxies
CPU
Application Servers (PHP, ASP, JSP)
I/O
Database Servers and Replicas
Updates
Static CDN


Latency
• DNS Lookups
• Multiple servers
• HTTP protocol
• TCP protocol
• Rendering
Bandwidth
• Limited
• Improving
• Video / Images
• More Devices
• Better Devices
Reliability
Crawlers, Scrapers, and Bots
Non-Traditional Web Clients
Storage
OS enhancements
Image Servers
Bandwidth
Spoon-feeding
Load Balancer
CPU
Object cache
Application Servers (PHP, ASP, JSP)
Session cache
Opcode cache
I/O
Database Servers and Replicas
Query Cache
Replication
Partitioning
Updates
Static CDN
Tail optimization

TRADITIONAL WEB CLIENTS

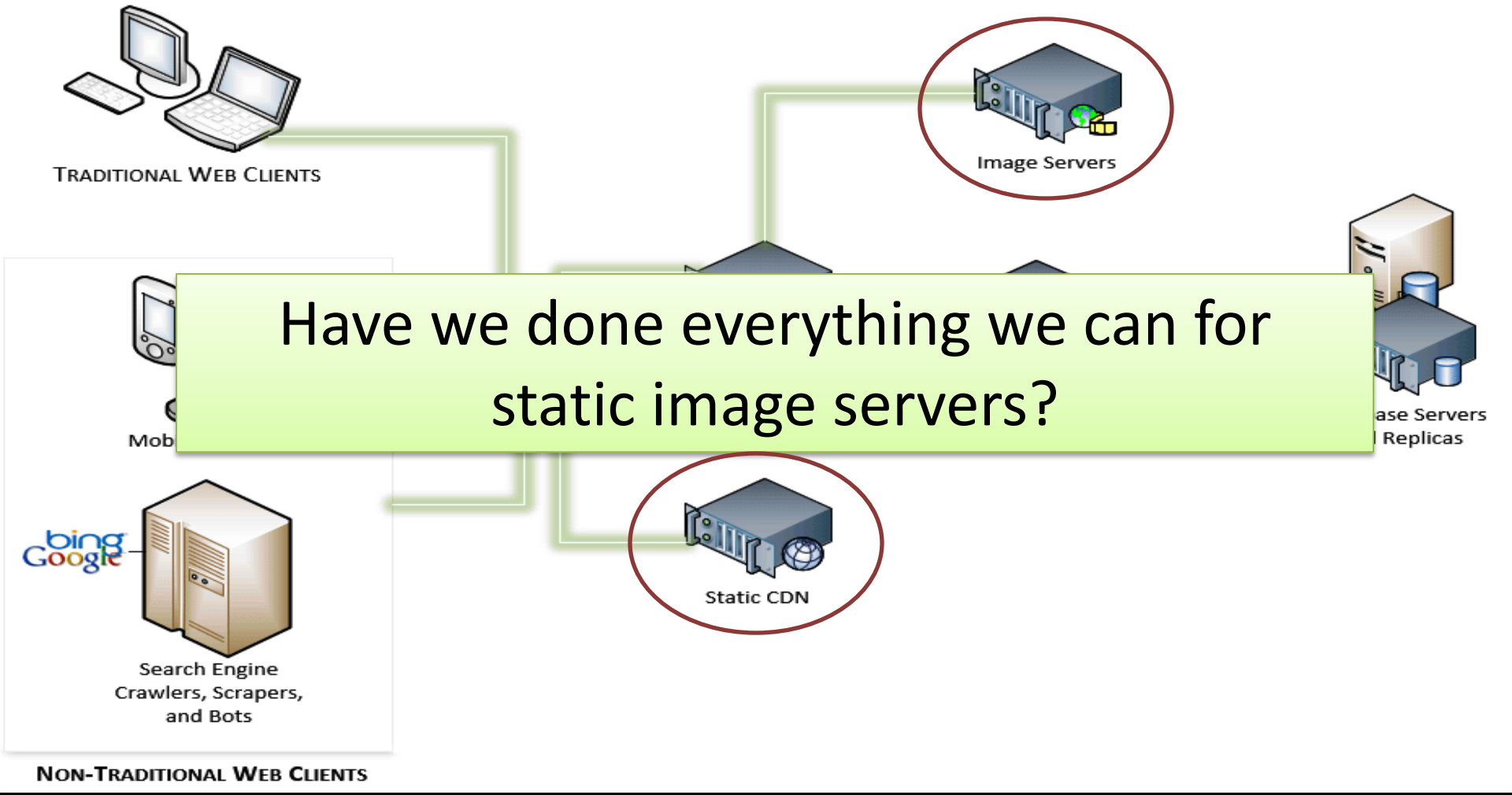
Image Servers

Resource Contention

Many opportunities to study systems of systems and their complex interactions, often resulting in Pareto-optimal (multi-objective tradeoff) scenarios.

Search Engine
Crawlers, Scrapers,
and Bots

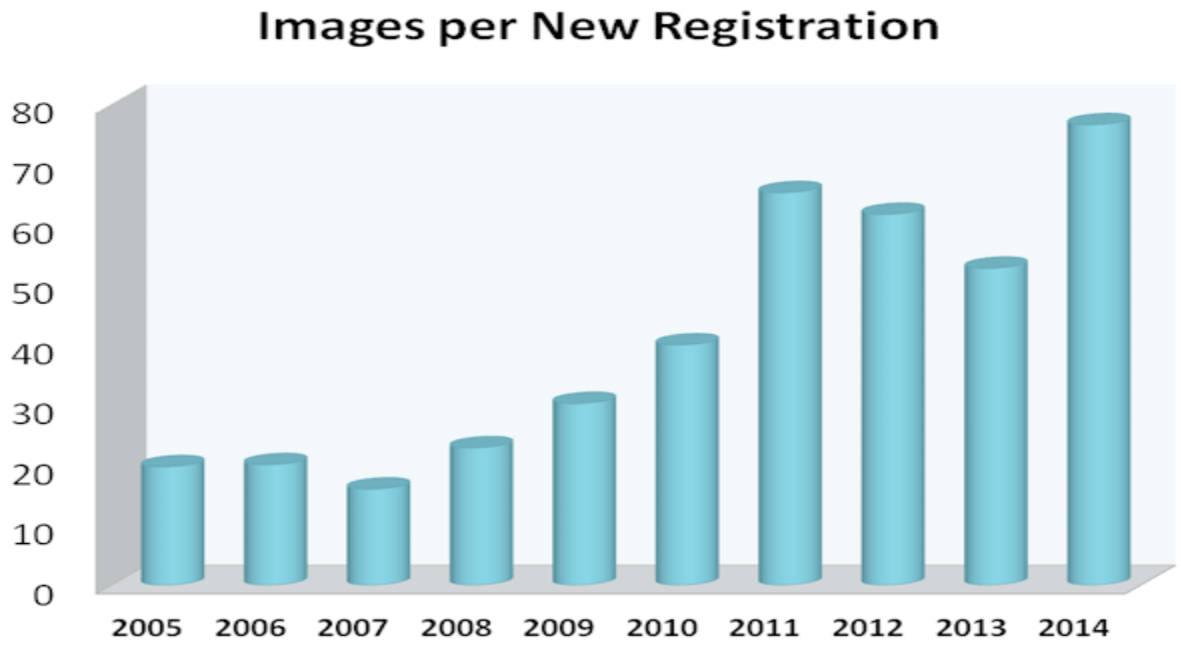
NON-TRADITIONAL WEB CLIENTS


Traditional Web Clients
Image Servers
Have we done everything we can for static image servers?
Static CDN
bing
Google
Search Engine
Crawlers, Scrapers,
and Bots
Non-Traditional Web Clients

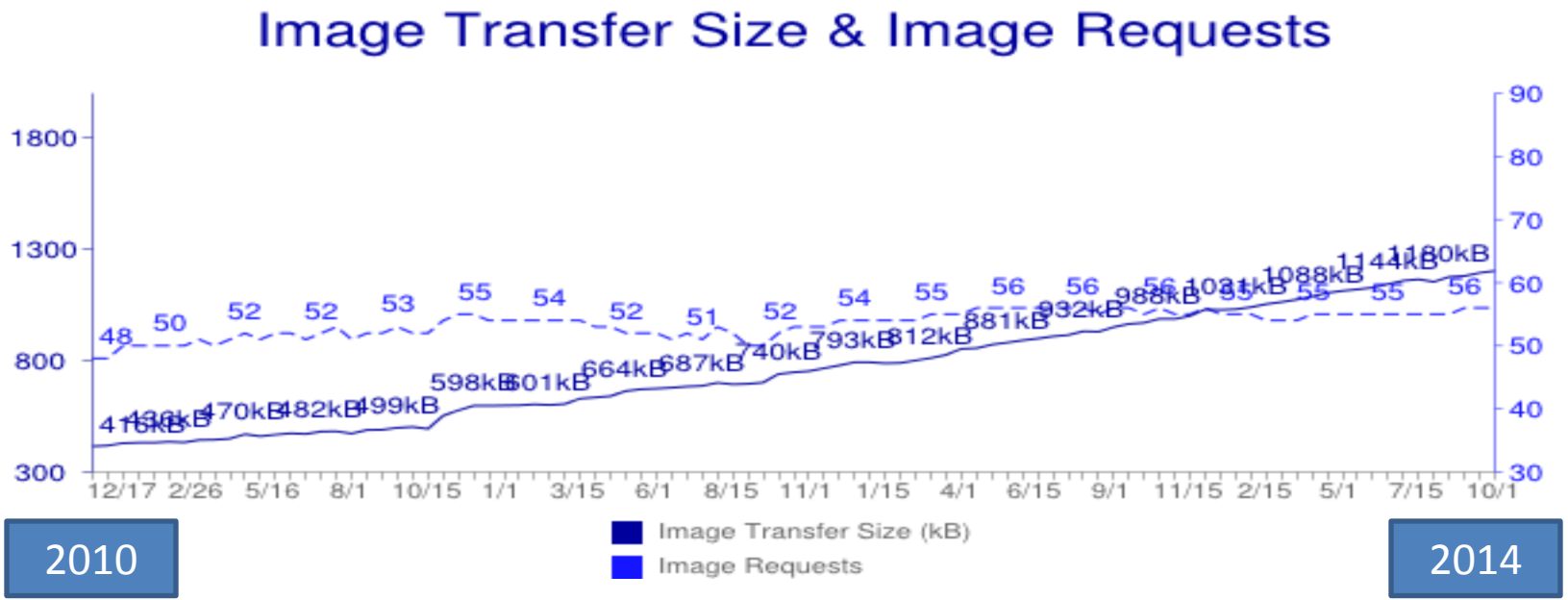
Social Media

- **Blogs and social media**
  - Social networks and online communities
  - Exploding with multimedia content
  - 73% of the world population now use
- Facebook
  - 2nd busiest site globally
  - 1.28 billion monthly users
  - 609 million mobile – daily

Social Media Images



Images per new registration on the author's online community.

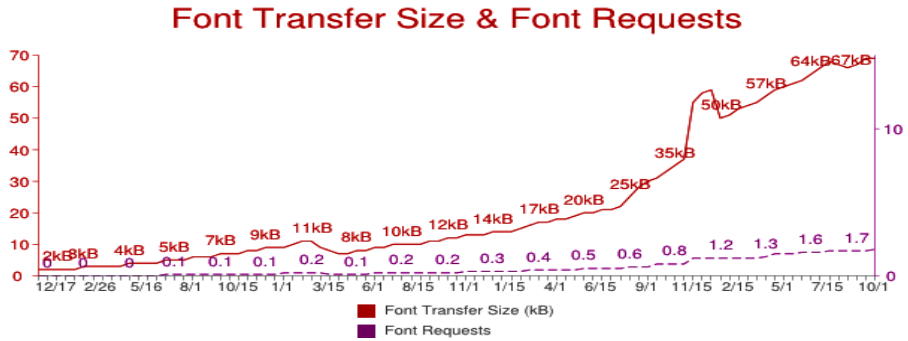


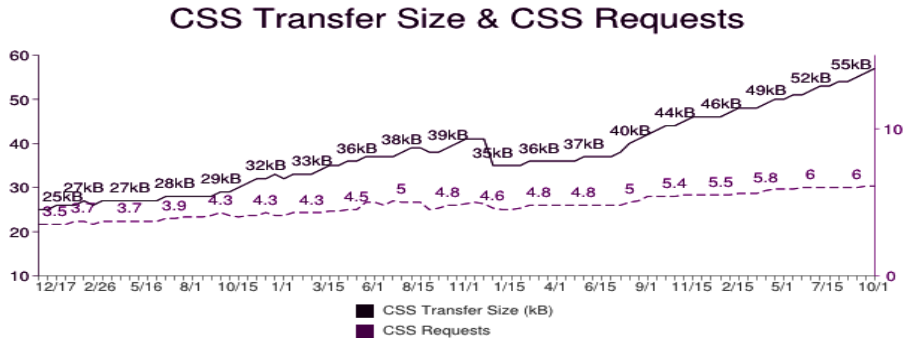
2010

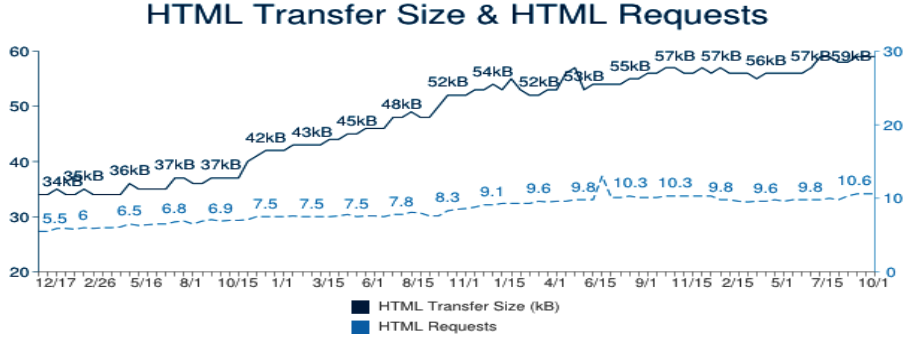
2014

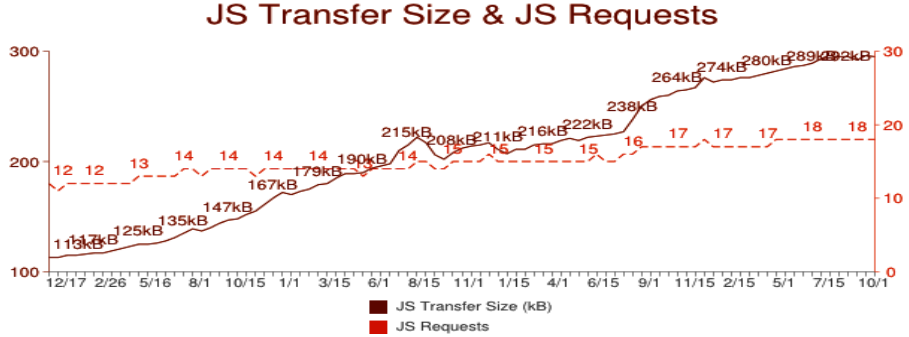
Source: httparchive.org crawl statistics

Image sizes have increased steadily



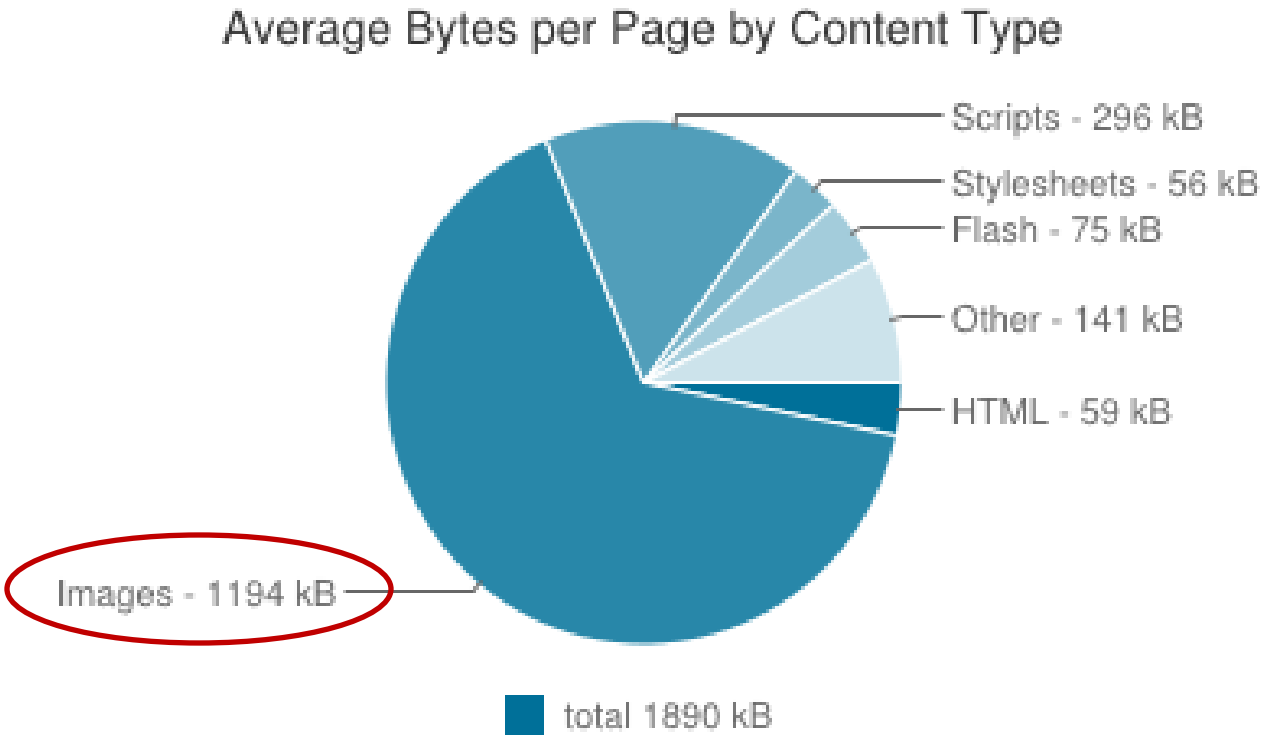






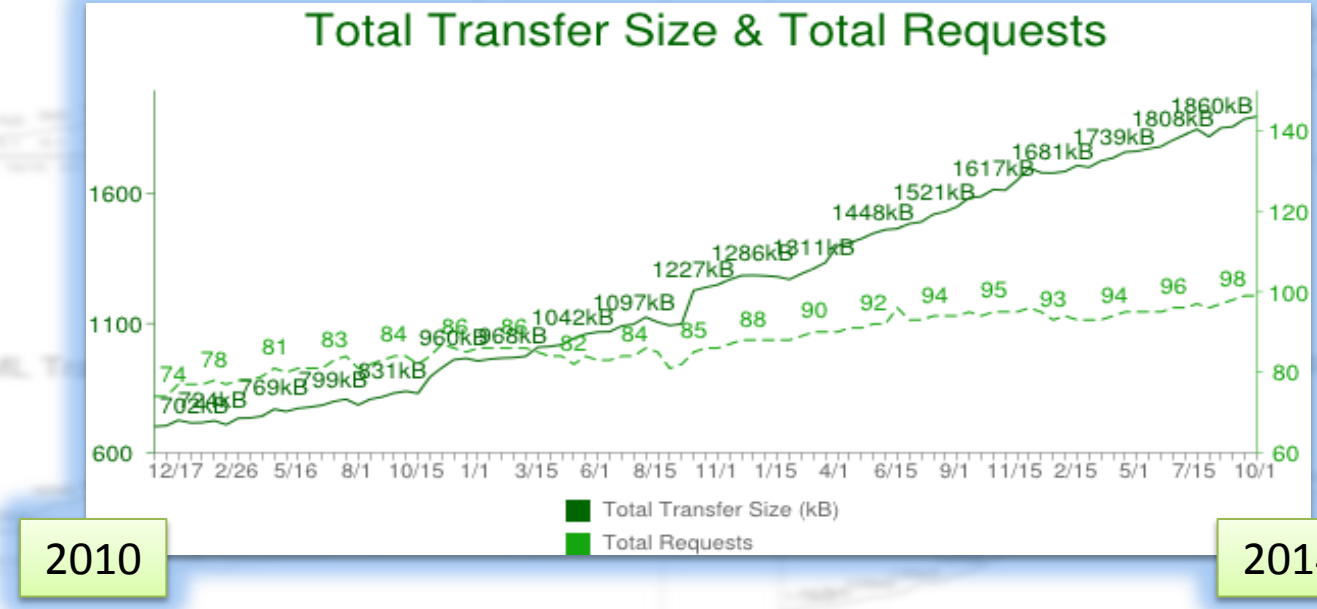
Source: httparchive.org crawl statistics

In fact, web pages are just getting larger overall



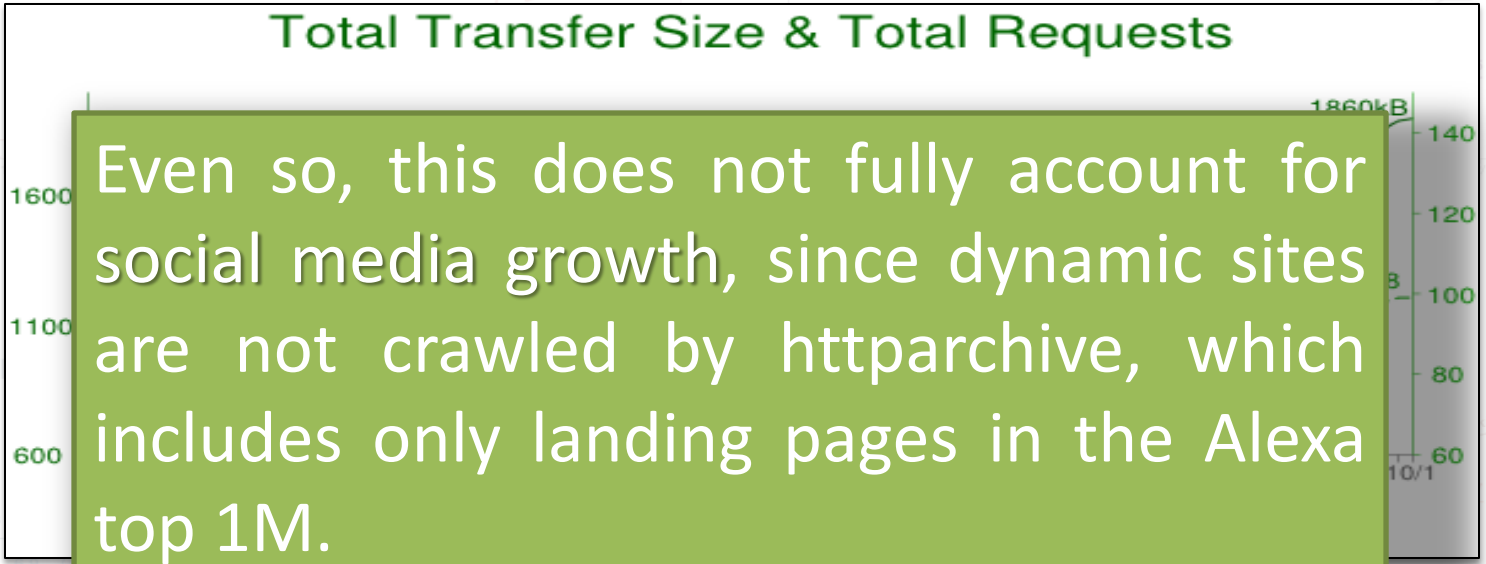
Source: httparchive.org crawl statistics

Images account for 63% of byte transfer



Source: httparchive.org crawl statistics

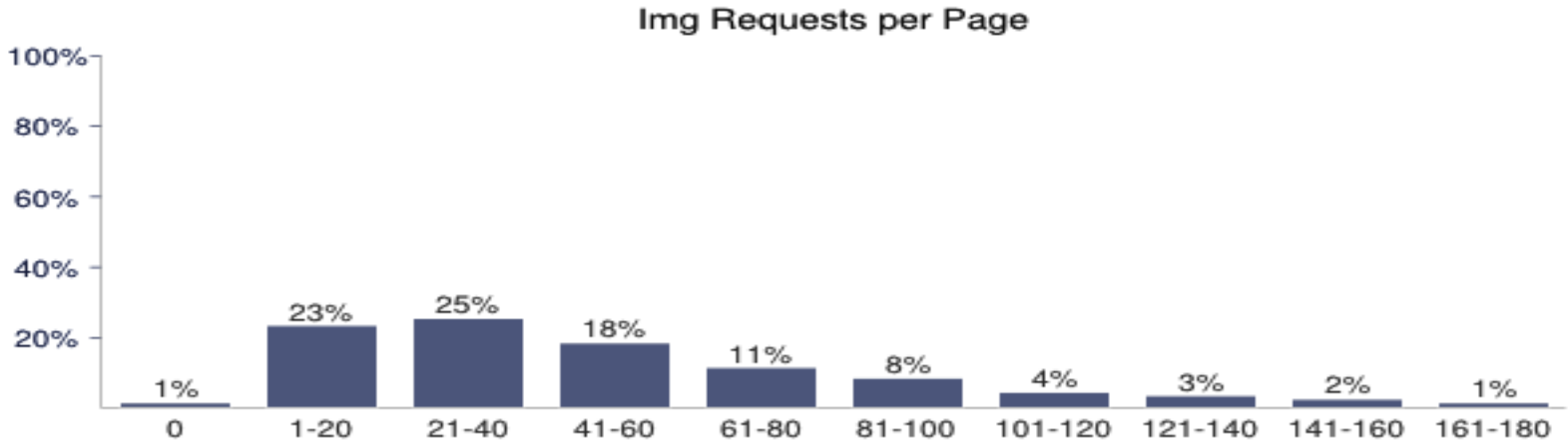
In fact, web pages are just getting larger overall

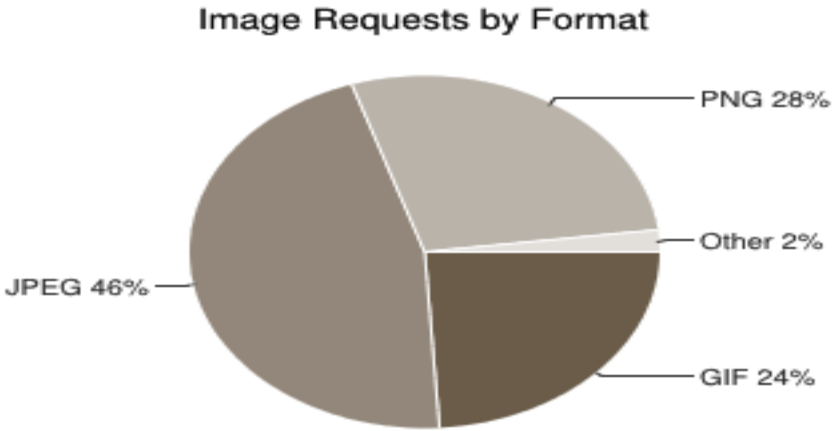


Even so, this does not fully account for social media growth, since dynamic sites are not crawled by httparchive, which includes only landing pages in the Alexa top 1M.

Source: httparchive.org crawl statistics

In fact, web pages are just getting larger overall





These extensions are typical of promotional and navigational graphics.

Source: httparchive.org crawl statistics

JPEG is much more dominant in social media

File Extensions

Extension	Proportion of Uploads
JPG	96.5 %
CDR	0.94 %
GIF	0.70 %
PNG	0.41 %
ALL OTHERS	2.7 %

Jpeg is the most prevalent uploaded to social networks, due in part because of their suitability for photos, good compression, large palette, and most common availability by mobile devices and consumer cameras.

Distribution of file extensions in the author's online community.

File Extensions

Extension	Proportion of Uploads
JPG	96.5 %
PNG	0.41 %
ALL OTHERS	2.7 %

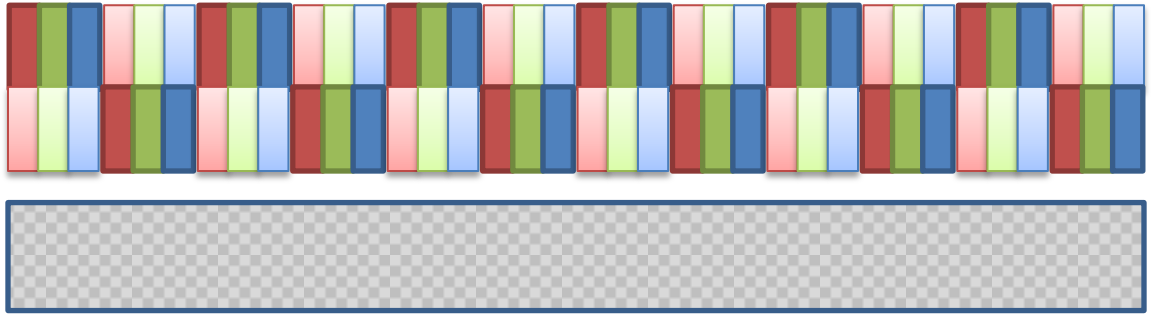
How can we optimize for this?

Distribution of file extensions in the author's online community.

Lossy Compression

(an application of *Psychophysics*…)

- RGB Color Space
  - (Red, Green, Blue) Pixel Intensity between 0-255



  - This representation is costly: 1024x1024 = 3.1MB

RGB Color Space

- **Luminance**
  - Changes in pixel intensity (brightness)

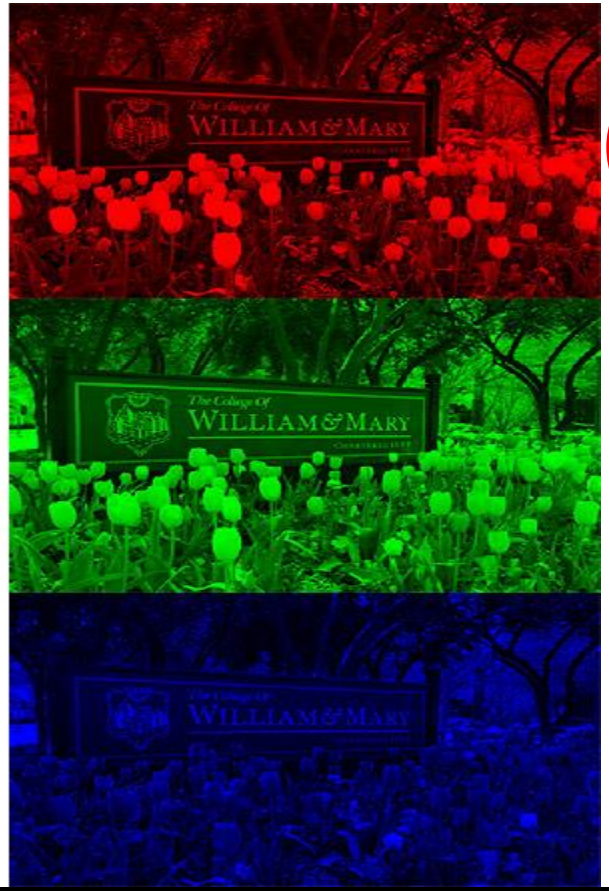
- **Chrominance**
  - Color information within the image

- **Human Visual System** (HVS) is much more sensitive to luminance than chrominance.

YCbCr Color Space
= Luminance (Y)
+ Chrominance B (Blue-Yellow)
+ Chrominance R (Red-Green)

RGB

YCbCr



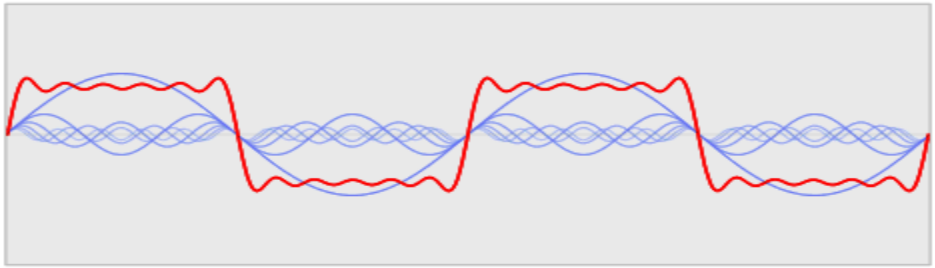
Most detail is in the Y (luminance) component

- Step 1: Convert from RGB to YCbCr
  - Lossless
- Step 2: Chrominance Subsampling
  - Replace each 2x2 block of values with an average.
  - Maintain integrity of luminance channel.
  - 75% of chrominance information discarded.
- Step 3: Discrete Cosine Transform
- Step 4: Quantization
- Step 5: Entropy Coding

Lossy Compression (JPEG)

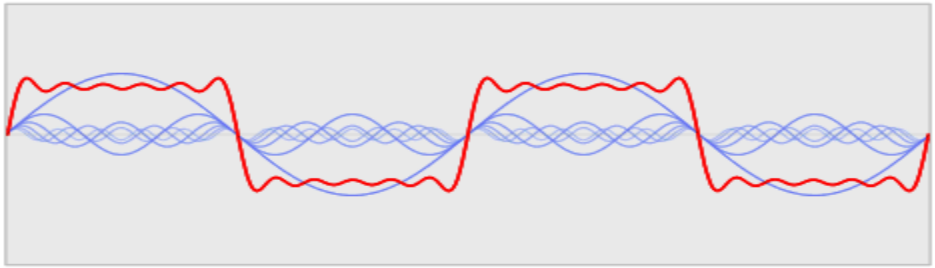
- Step 1: Convert from RGB to YCbCr
  - Lossless
- Step 2: Chrominance Subsampling
  - Replace each 2x2 block of values with an average.
  - Maintain integrity of luminance channel.
  - 75% of chrominance information discarded.
- Step 3: Discrete Cosine Transform
- Step 4: Quantization
- Step 5: Entropy Coding

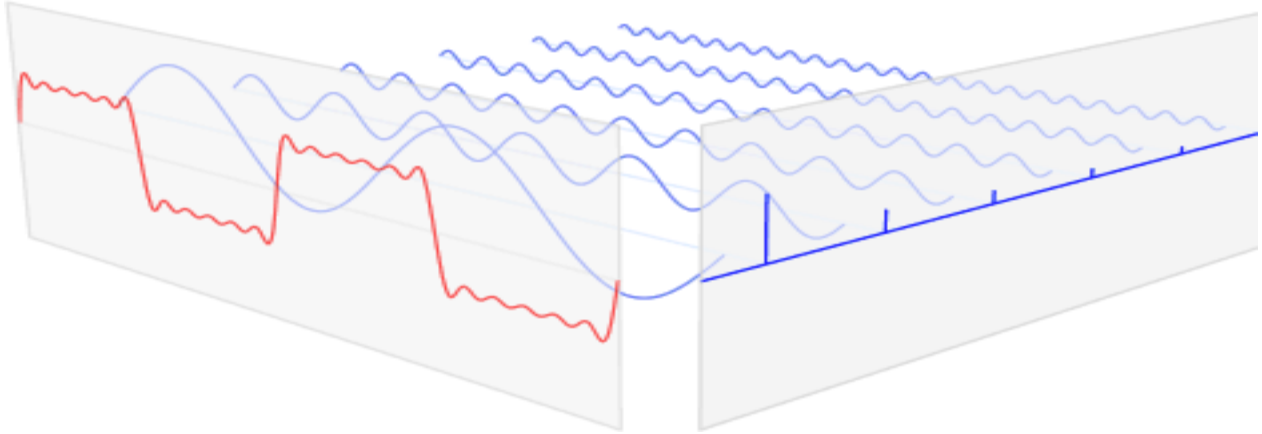
Lossy Compression (JPEG)



- DCT
  - Like FFT, converts a signal from **time** domain to **frequency** domain
  - Uses superposition of cosines
- Red wave is a superposition of blue waves.
  - Each wave has a fixed frequency.
  - Each wave has an amplitude (coefficient) that corresponds to the wave's influence.

Step 3: Discrete Cosine Transform

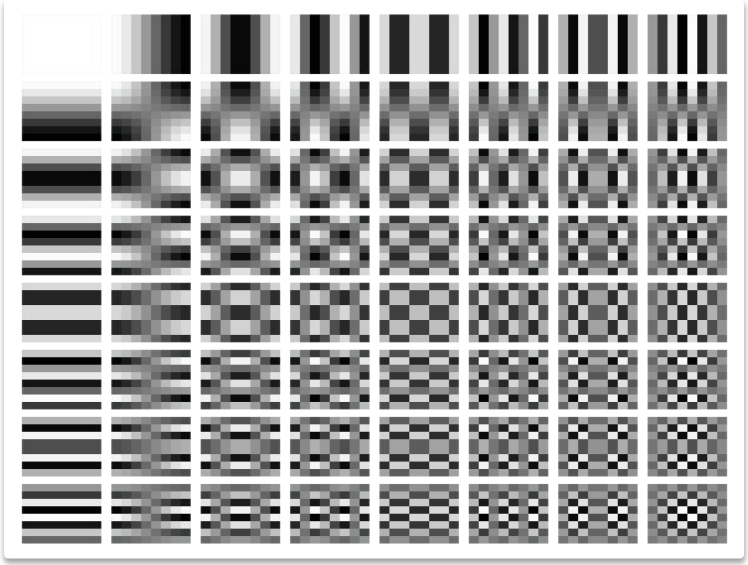




Lower frequency (large) waves are more important to the shape of the signal.

Step 3: Discrete Cosine Transform

- Set of 64 basis images
  - JPEG expresses each 8x8 block in an image as a linear combination of these 2D basis functions, each basis image gets assigned a coefficient.
- Observation
  - HVS is more sensitive to weights in the upper left, and less sensitive to lower right (high frequency).



Each channel, YCbCr, is processed using the DCT separately, to produce an 8x8 matrix of coefficients.

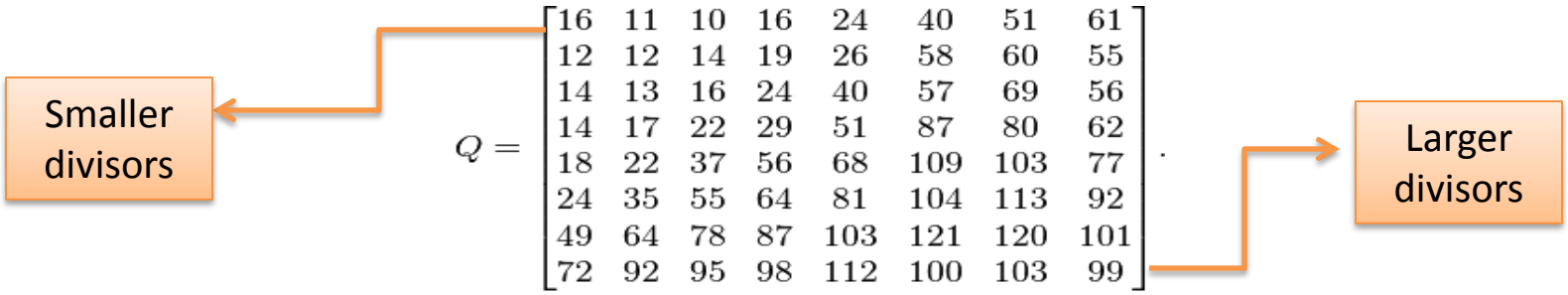
$$G = \begin{bmatrix} -415.38 & -30.19 & -61.20 & 27.24 & 56.12 & -20.10 & -2.39 & 0.46 \\ 4.47 & -21.86 & -60.76 & 10.25 & 13.15 & -7.09 & -8.54 & 4.88 \\ -46.83 & 7.37 & 77.13 & -24.56 & -28.91 & 9.93 & 5.42 & -5.65 \\ -48.53 & 12.07 & 34.10 & -14.76 & -10.24 & 6.30 & 1.83 & 1.95 \\ 12.12 & -6.55 & -13.20 & -3.95 & -1.87 & 1.75 & -2.79 & 3.14 \\ -7.73 & 2.91 & 2.38 & -5.94 & -2.38 & 0.94 & 4.30 & 1.85 \\ -1.03 & 0.18 & 0.42 & -2.42 & -0.88 & -3.02 & 4.12 & -0.66 \\ -0.17 & 0.14 & -1.07 & -4.19 & -1.17 & -0.10 & 0.50 & 1.68 \end{bmatrix} \begin{matrix} \overrightarrow{u} \\ \\ \\ \\ \\ \\ \\ \downarrow v. \end{matrix}$$

2D Discrete Cosine Transform on 8x8 blocks of pixels

- Step 1: Convert from RGB to YCbCr
  - Lossless
- Step 2: Chrominance Subsampling
  - Replace each block of 4 pixels with an average
  - Maintains luminance.
  - 75% of chrominance information discarded.
- Step 3: Discrete Cosine Transform: lossless
- Step 4: Quantization
- Step 5: Entropy Coding

Lossy Compression (JPEG)

- Quantization
  - A quantization matrix controls how much information is discarded.



  - Dividing our coefficients by Q causes high-frequency components (lower right) to become mostly zero.

$$B = \begin{bmatrix} -26 & -3 & -6 & 2 & 2 & -1 & 0 & 0 \\ 0 & -2 & -4 & 1 & 1 & 0 & 0 & 0 \\ -3 & 1 & 5 & -1 & -1 & 0 & 0 & 0 \\ -3 & 1 & 2 & -1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

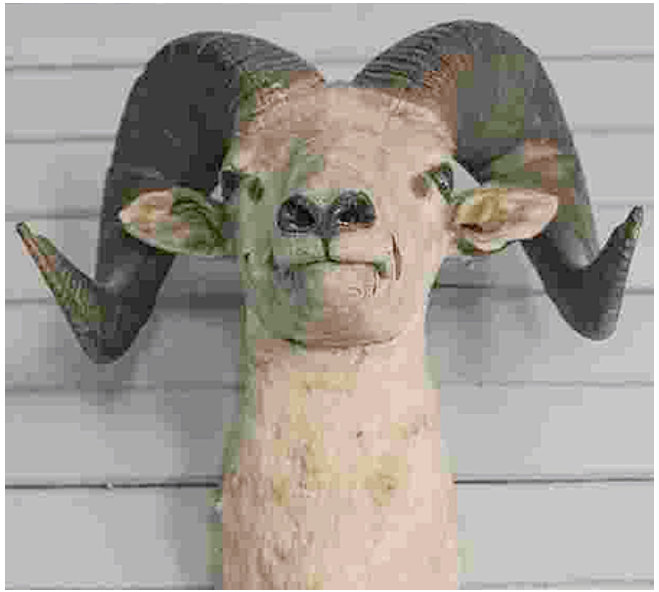
New matrix is easier to compress using entropy (Huffman) coding.

- **Quantization**
  - This is where the “lossy” compression occurs.

- **JPEG Quality**
  - *Quality factor* scales the quantization matrix. [0..100]%
  - The image on the right is at 10% quality, note the sky.
  - Manual process.
  - Q optimization is combinatorially hard.

JPEG Quality Factor

- Google introduced a new WEBP image format for the web (based on VP8).
  - More smoothing at lower quality settings
  - Produces smaller files than JPEG
  - Uses a similar quality factor (0..100)



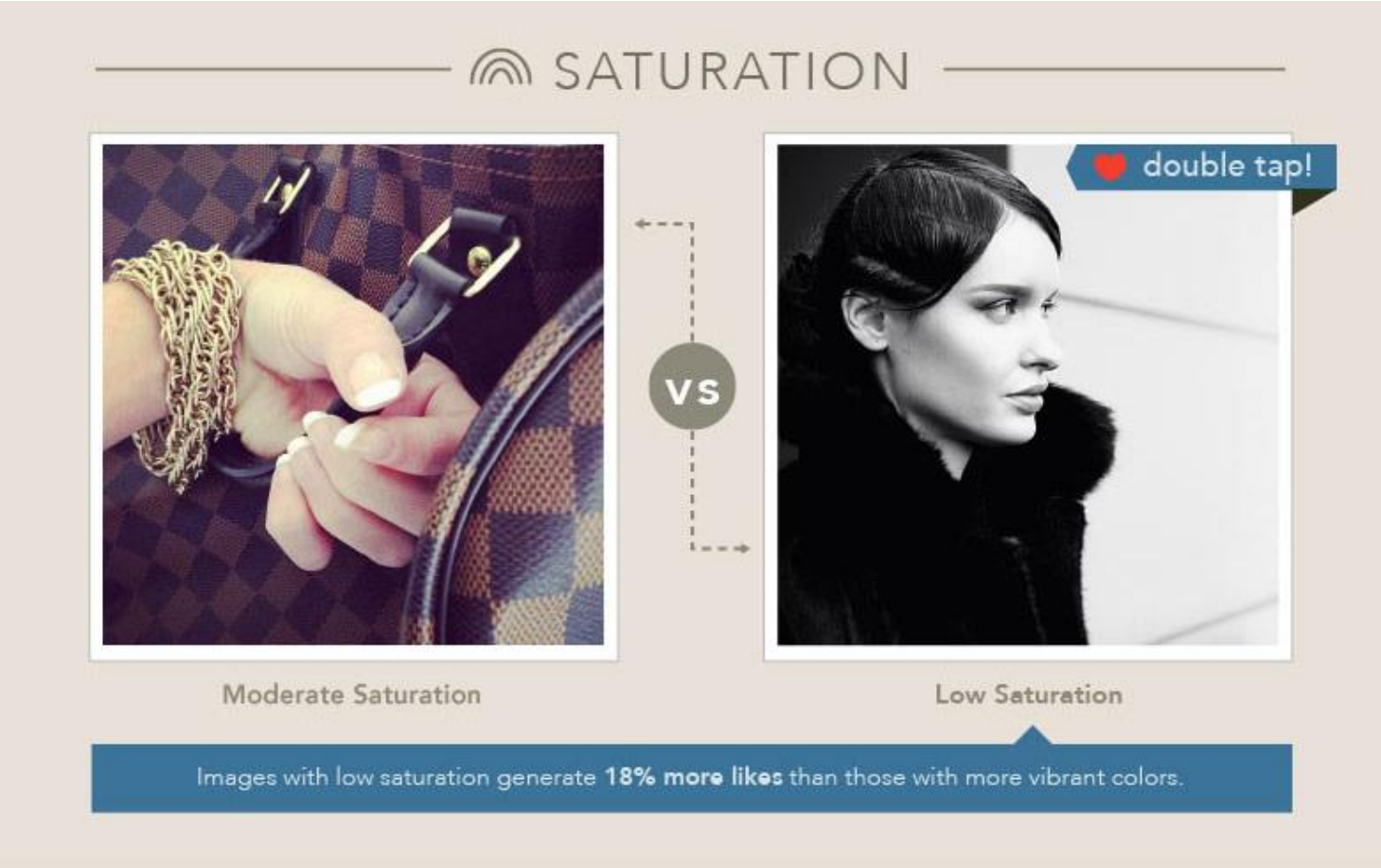
JPEG vs WEBP

Psychovisual Enhancements

- Curalate – Visual Marketing Insights
  - 8,000,000 Instagram photos
  - Data science team correlated features to “likes”

Source: Curalate Inc. social media analysis 2014

Curalate – Analysis of Instagram



Source: Curalate Inc. social media analysis 2014

Curalate – Analysis of Instagram

Source: Curalate Inc. social media analysis 2014

Curalate – Analysis of Instagram

- JPEGMini — SPIE'11
  - Uses experimentally acquired distributions to reduce quality factor to psychovisual threshold.
  - Works best for large photography archives.
- What makes images popular? — WWW'14
  - Analysis of color, gradients, social cues
- JPEG quantization table selection by the firefly algorithm (swarm intelligence) — ICMCS'14
- Google+ image enhancement — Google
- Facebook image compression — Facebook
- Curalate – Social Media Image Analysis — Curalate Inc.
- Smush.it — Yahoo
  - Removes EXIF metadata, lossless.

...

[Related Work]

Empirical Observation

Can we introduce psychovisual enhancements to make lossy compression more effective?

Research Question

Psychovisual Enhancements

- 15 images — Image Set A
  - Sourced from random selection of photos
  - Indoor and outdoor, scenes and close-ups
  - No human subjects
- 125,000 images — Image Set B
  - Randomly selected from online uploads

- Explored two psychovisual enhancements:
  - Increase saturation
  - Lighten shadows

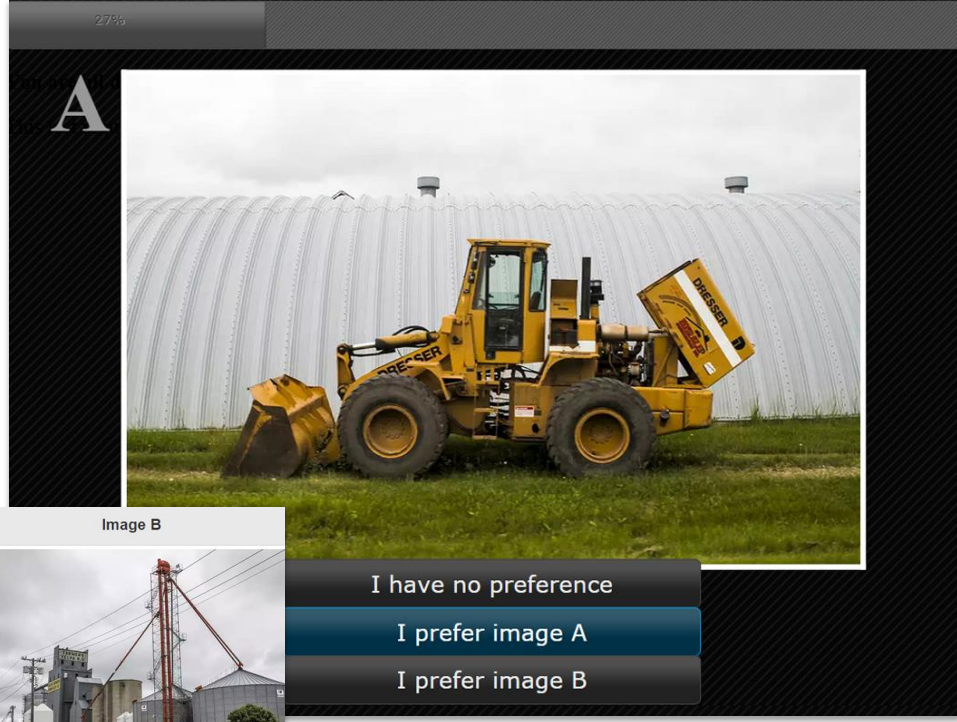
- Explored 6 quality settings
  - 10, 20, 30, 40, 50, and 60
  - 85 is a typical setting used today
  - Many web designers are reluctant to go below 95

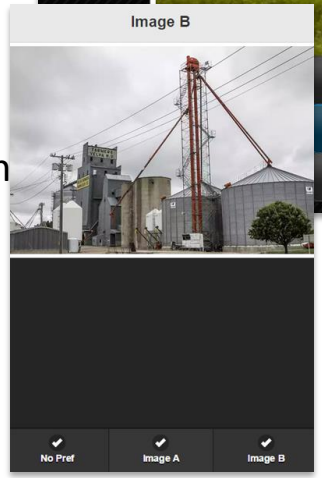
Experiment

- Implemented two interfaces
  - Desktop and Mobile

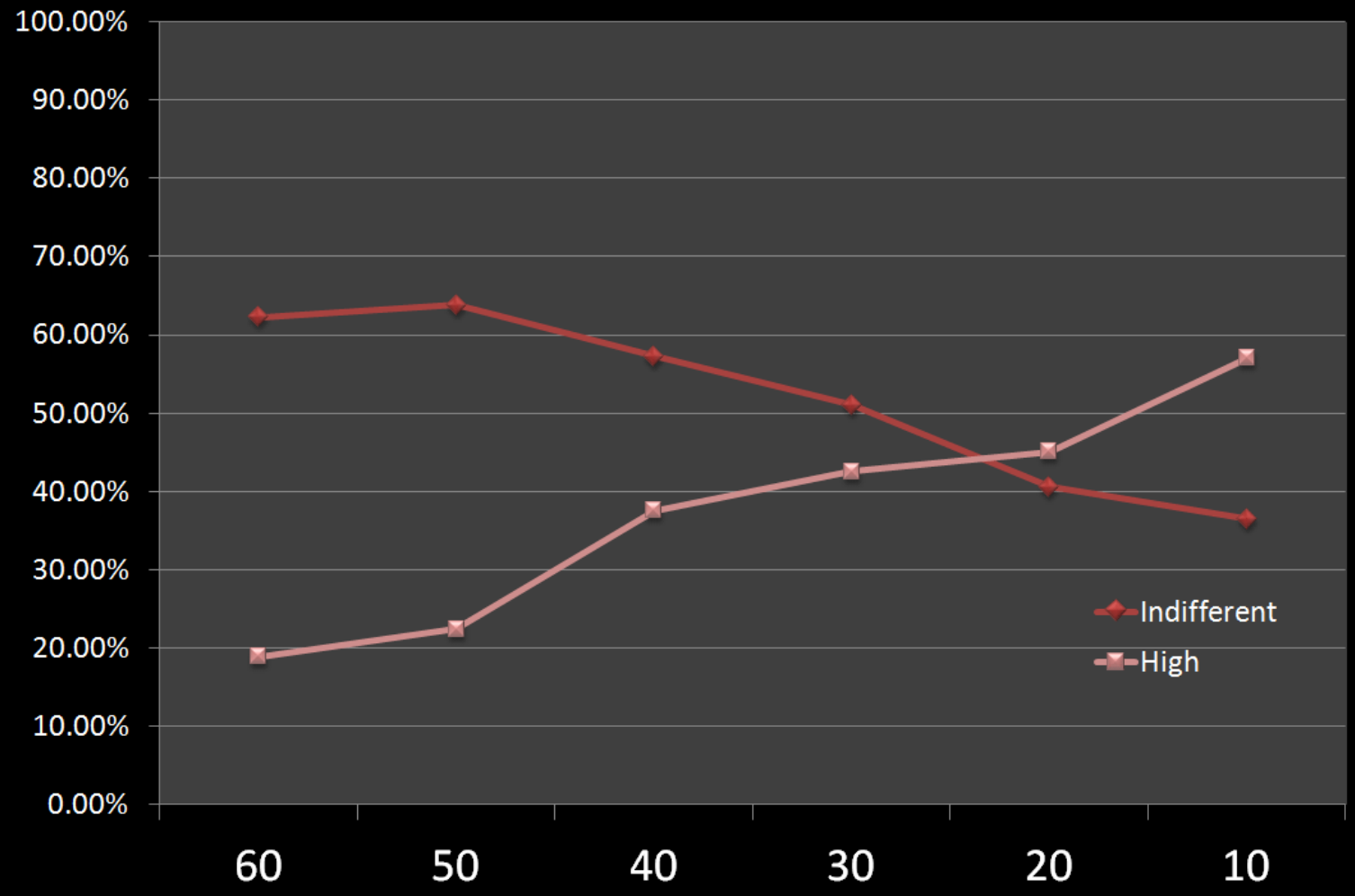
- Test subjects were not made aware of the nature of the test
- Order, quality, and sorting of images was randomized

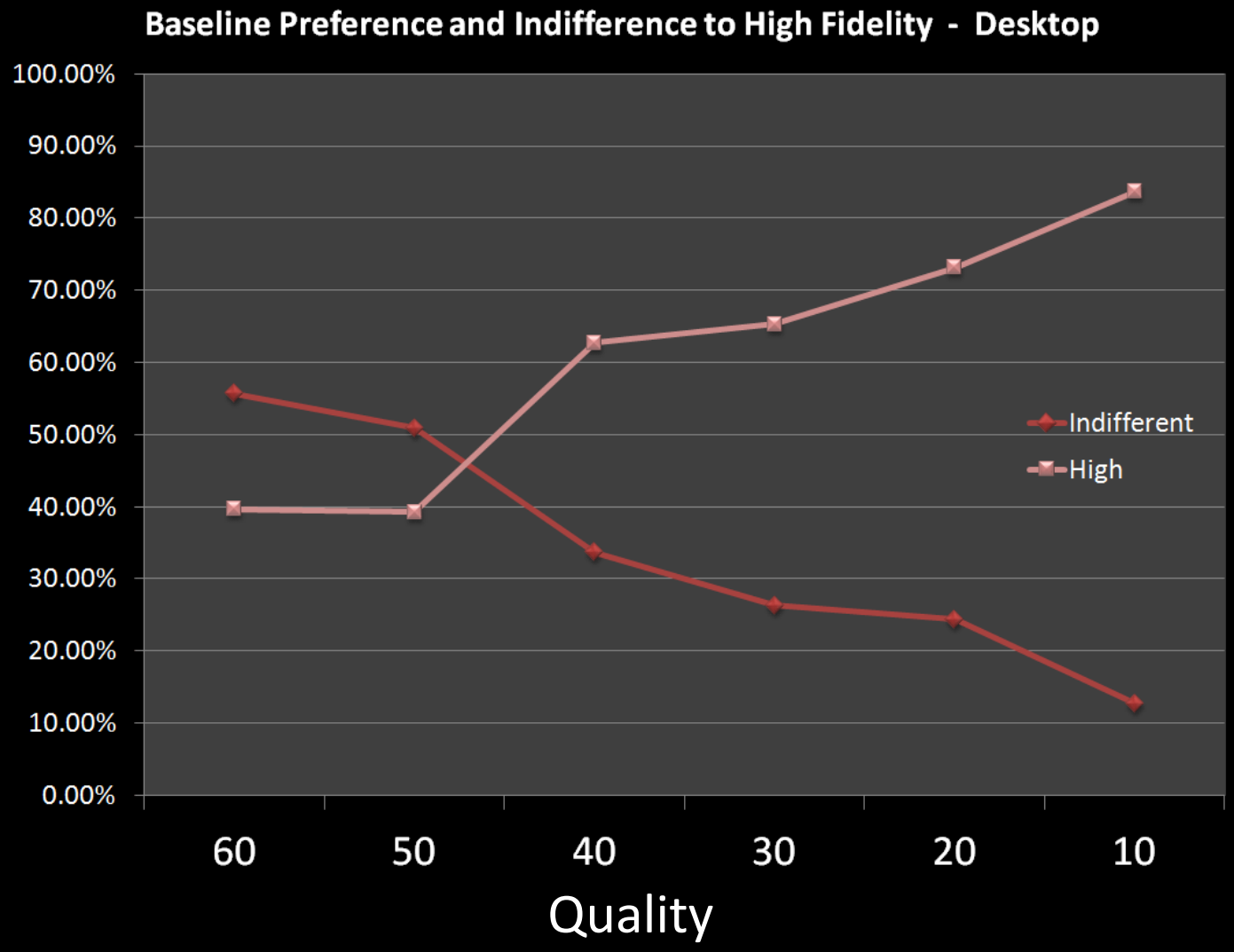
- Employed DSFCC randomized block design
  - Double Stimulus Forced Choice Comparison



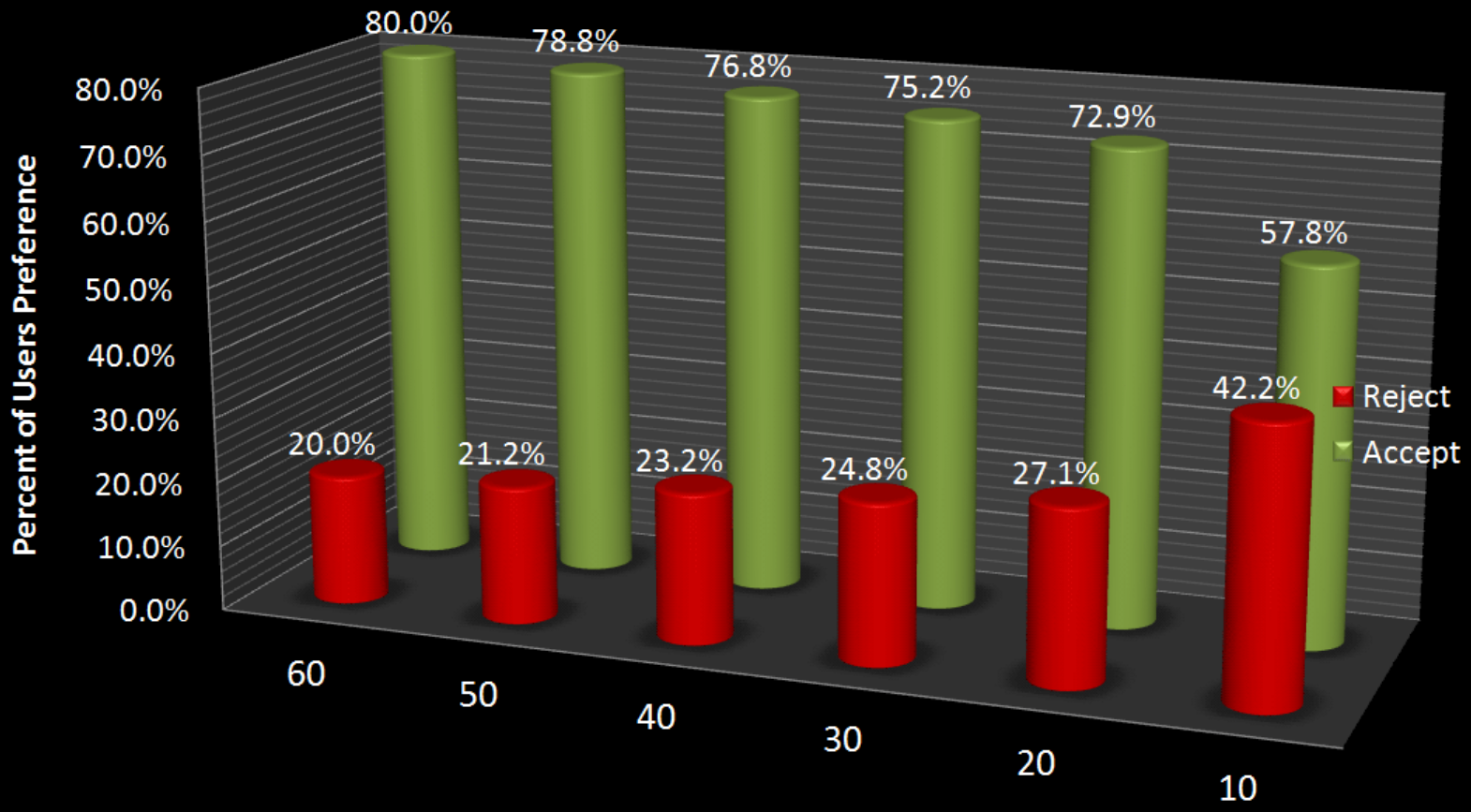


Experimental Design

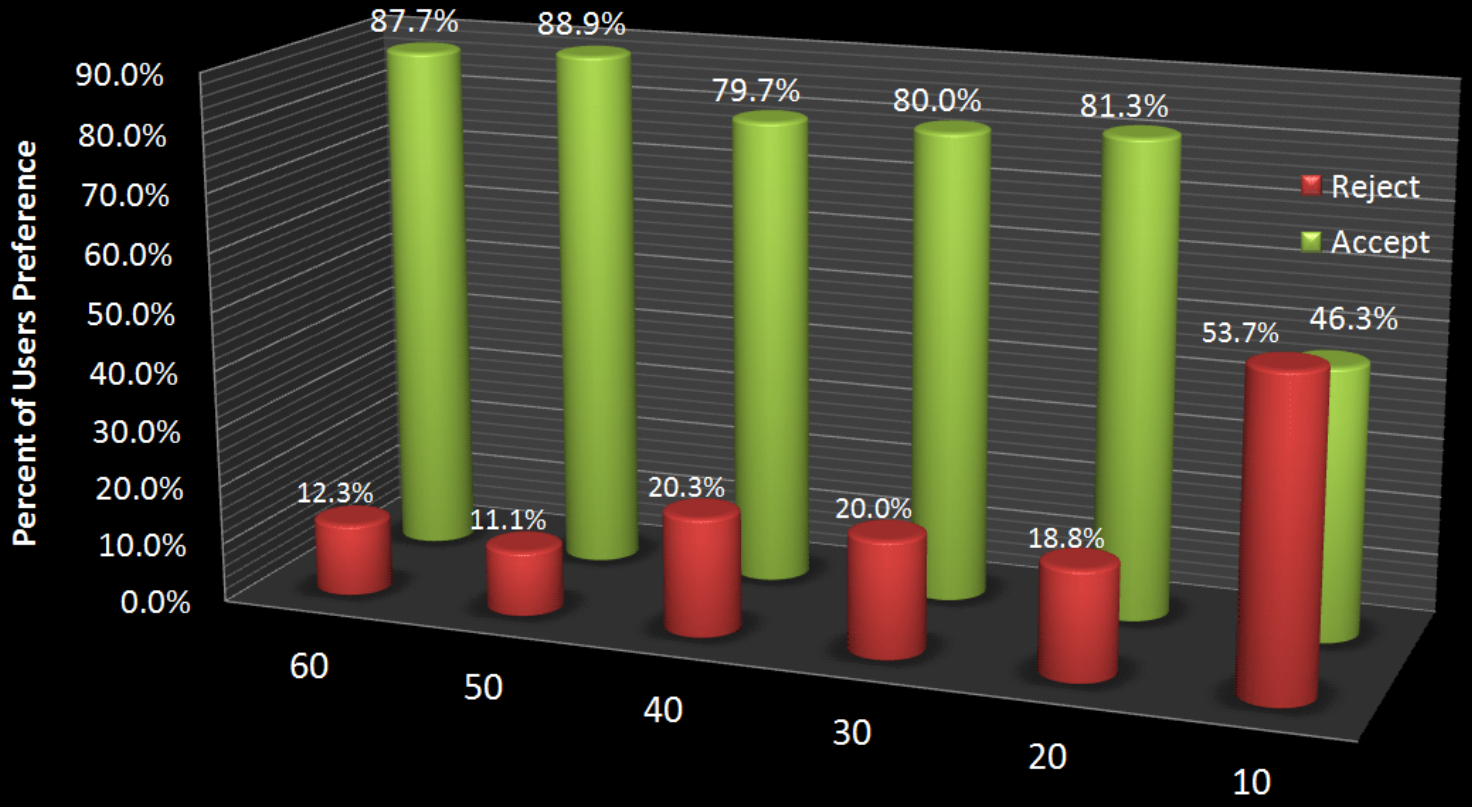

Baseline Preference and Indifference to High Fidelity
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
60
50
40
30
20
10
Indifferent
High

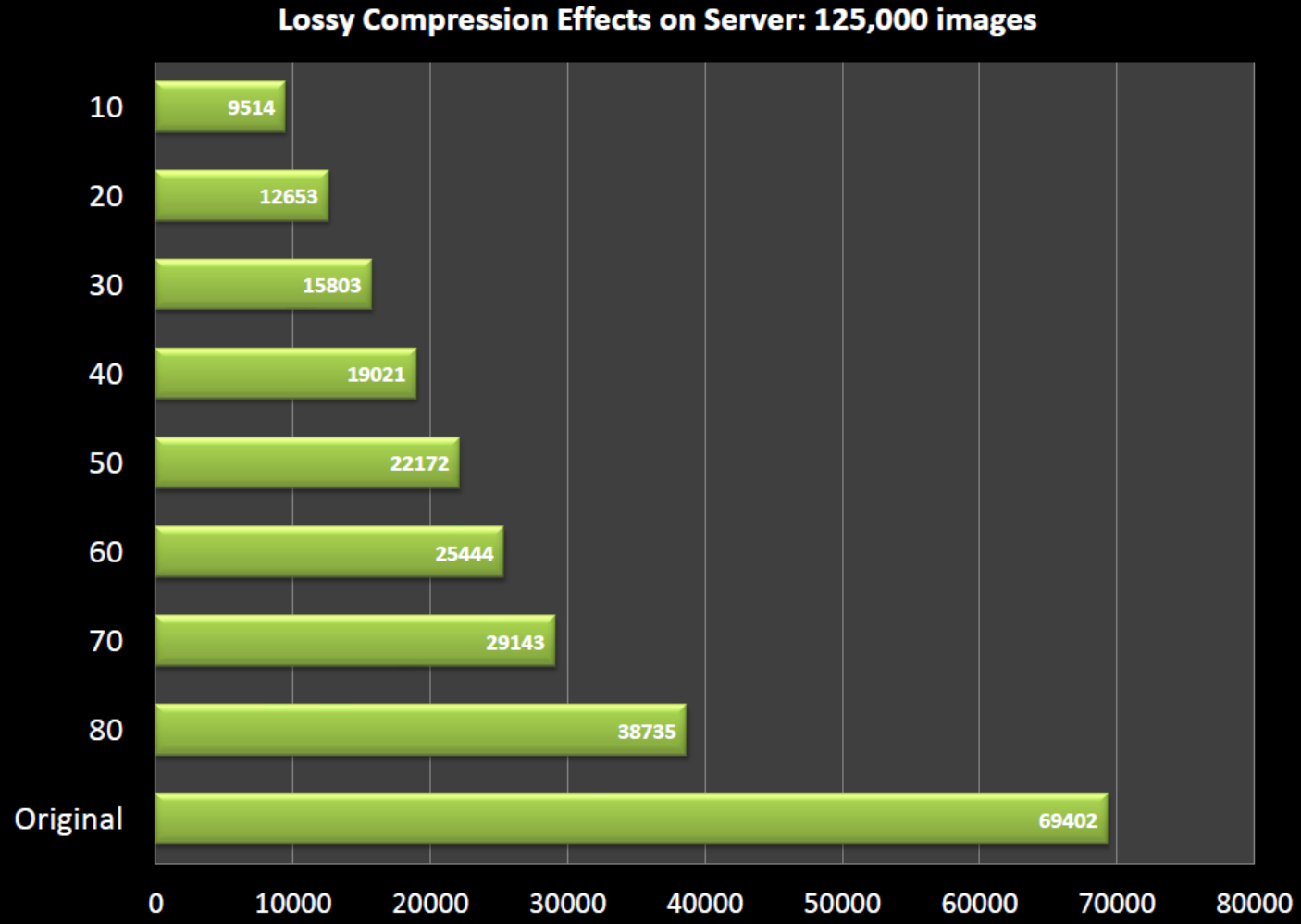

Baseline Preference and Indifference to High Fidelity - Desktop
100.00%
90.00%
80.00%
70.00%
60.00%
50.00%
40.00%
30.00%
20.00%
10.00%
0.00%
Indifferent
High
60
50
40
30
20
10
Quality

Scaling with Enhanced Reduced Fidelity (SERF)



Scaling with Enhanced Reduced Fidelity (SERF)



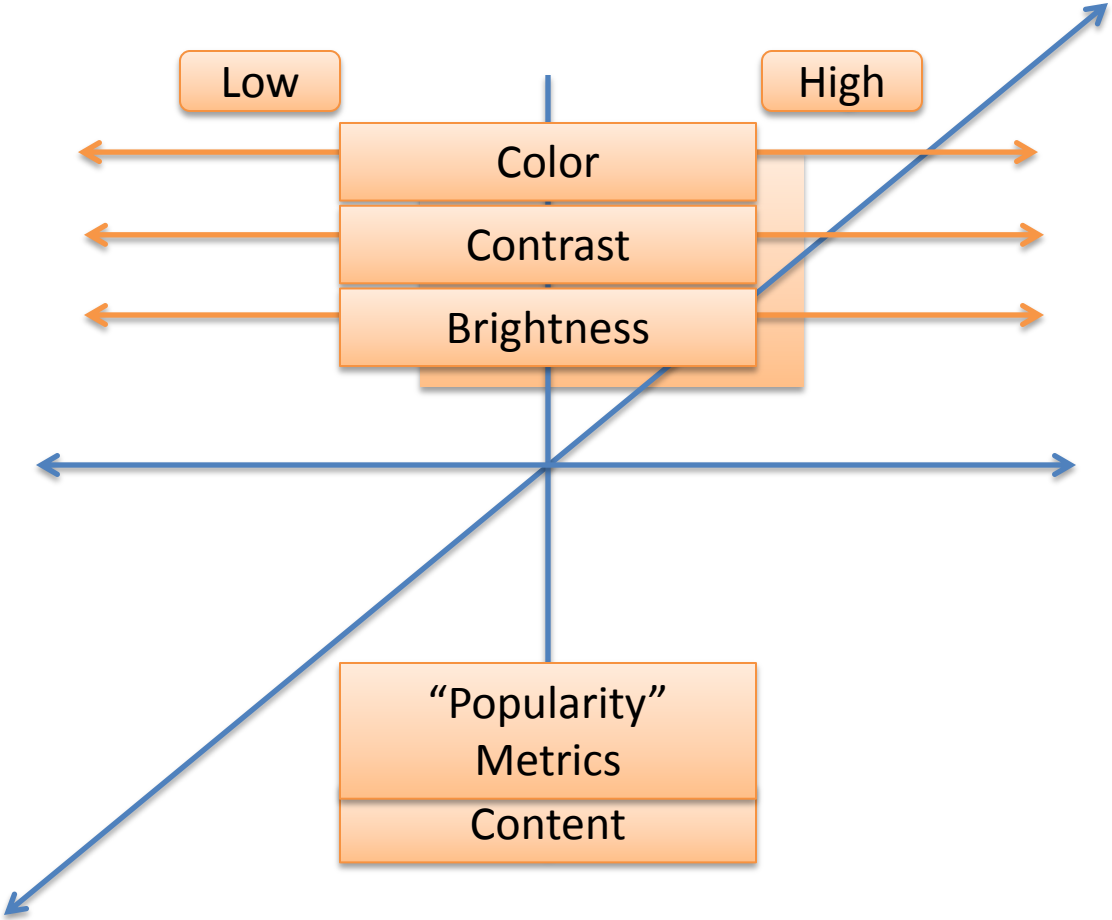

Lossy Compression Effects on Server: 125,000 images
10
9514
20
12653
30
15803
40
19021
50
22172
60
25444
70
29143
80
38735
Original
69402
0
10000
20000
30000
40000
50000
60000
70000
80000

- Image fidelity can be reduced when psychovisual enhancements are applied.
  - Only 7% point difference between q=60 and q=20 for mobile devices and desktop

- SERF can enable data centers to scale static image servers by compressing images below the psychovisual threshold.

Summary

- Many exciting directions…
  - What psychological effects are at play? Colors?
  - Do we apply enhancements indiscriminately, or are there diminishing returns?
  - Does the subject matter?
  - How does this fare with human subjects?
  - Are there stronger enhancements, or are these the best?



Many Dimensions…

Psychovisual Enhancements to Enable Lossy Compression

- Thank you.