

The background is black with a large, faint, light-gray graphic of concentric circles and a power button symbol (a circle with a vertical line) in the center. The text is centered over this graphic.

It's Not the Cost, It's the Quality!

Ion Stoica

Conviva Networks and UC Berkeley

CONVIVA

A Brief History

- ⌚ Fall, 2006: Started Conviva with Hui Zhang (CMU)
- ⌚ Initial goal: use p2p technologies to reduce distribution costs and improve the scale
- ⌚ Slowly, realized our customers (content premium producers & aggregators) value more quality than cost
- ⌚ Today: maximize distribution quality, distribution management, and provide real-time analytics


Where is the Data Coming From?

Content Providers and Aggregators



CDNs





Trends

CONVIVA

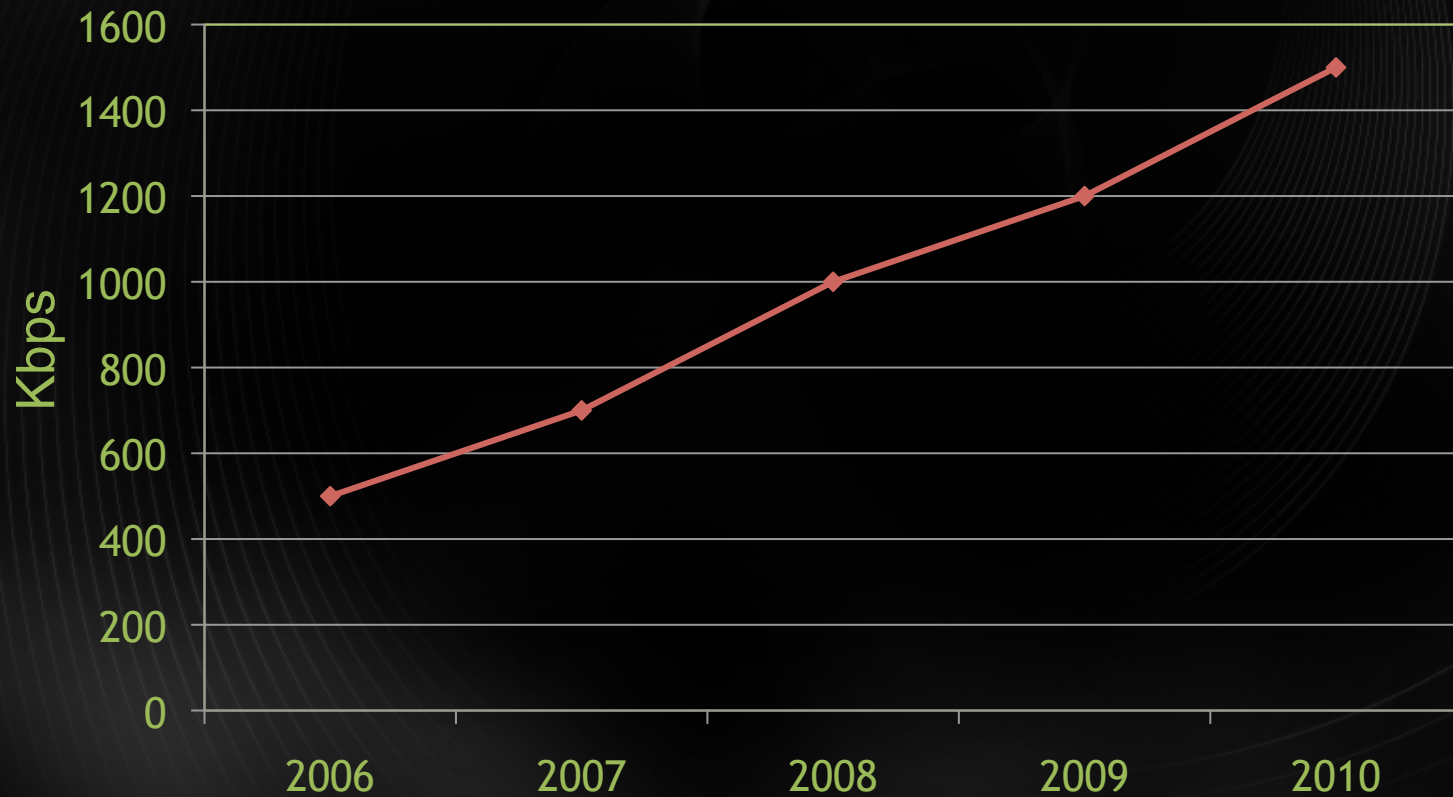
Trends: CDN Pricing

- ⏻ CDN pricing has decreased x1.5-2 every year over the last 5 year



Trends: Streaming Rate for Premium Content

- ⏻ Average streaming rate has increased 20-40% every year



Trends: Per-hour Streaming Cost

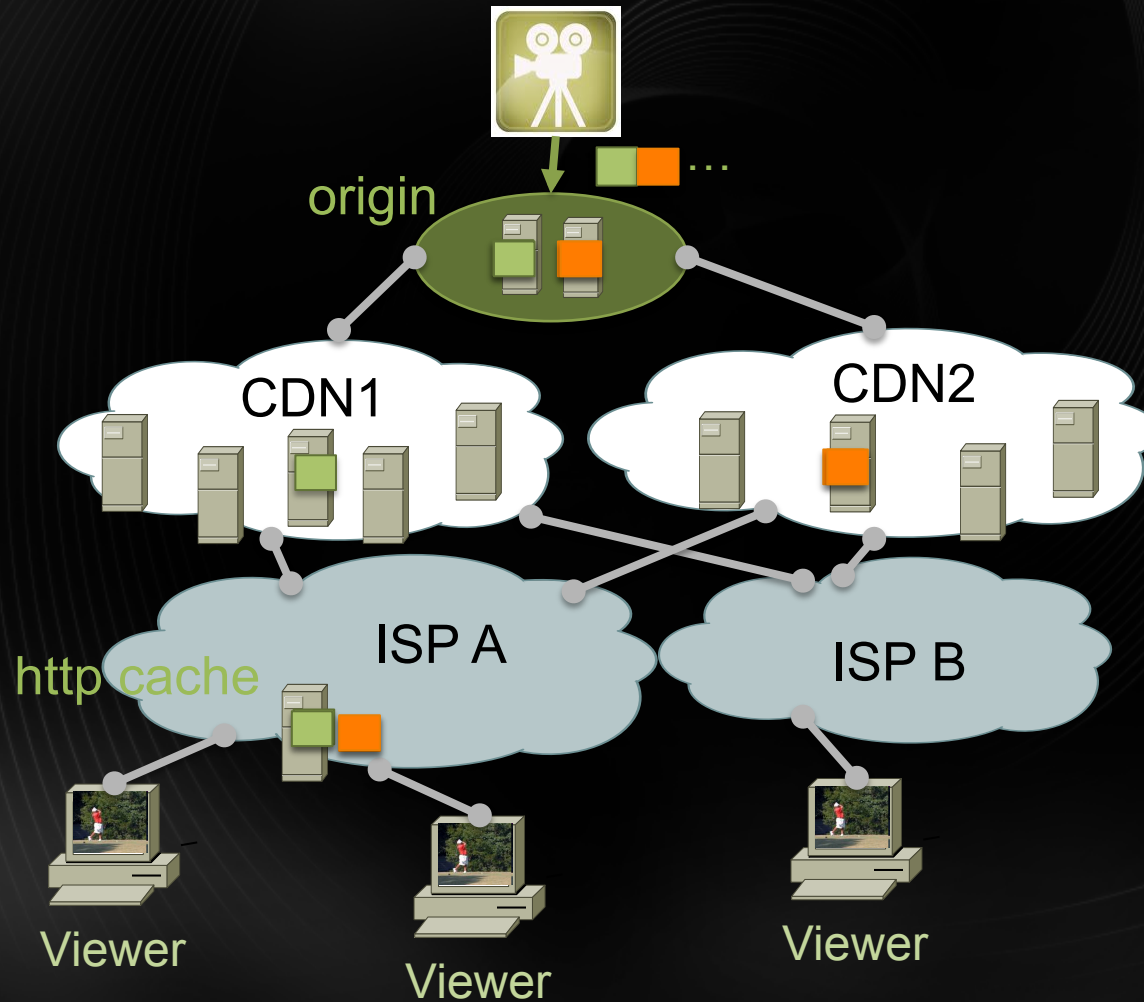
- ⏻ Per-hour streaming cost has decreased 15-35% every year



HTTP Chunking

- Ⓢ Trend accelerated by switching from proprietary streaming technologies (e.g., Adobe's FMS) to HTTP Chunking:
 - Move Networks (2005)
 - Apple (2008)
 - Microsoft (2008/2009)
 - Adobe (2010, 2nd half)

How Does HTTP Chunking Work?



HTTP Chunking Advantages

- ⦿ Chunks: immutable, relative large objects (hundreds of KB)
 - Great for caching
- ⦿ Leverage existing HTTP infrastructure
 - CDNs
 - ISP deployed caches
 - Enterprise http proxies
- ⦿ Low cost and high scale

What Does this Mean?

- 🔌 Ad supported premium content
 - CPM (cost per thousand of ad impressions) for premium content has reached: \$20-\$40
 - **One ad covers one hour of streaming!**
- 🔌 Paid content
 - \$0.99 episode, distribution cost < **3%**
- 🔌 Subscription based premium content
 - Distribution, usually a few percents of total cost
 - It costs \$1.6 per month to stream content to an user watching 2 hours per day
- 🔌 Production & rights costs dominate

The background is a dark, almost black, gradient. It features a series of concentric, slightly irregular circles that create a tunnel-like or ripple effect. In the center of these circles, there is a faint, dark silhouette of a person's head and shoulders, facing forward. The overall aesthetic is modern and minimalist.

Quality Matters

CONVIVA

Quality Matters

⌚ Better quality

- Increase viewing time → more ad opportunities
- Increase retention rate
- Protect brand

⌚ Quality

- Join time
- Buffering ratio
- Rendering quality
- Streaming rate

Analysis

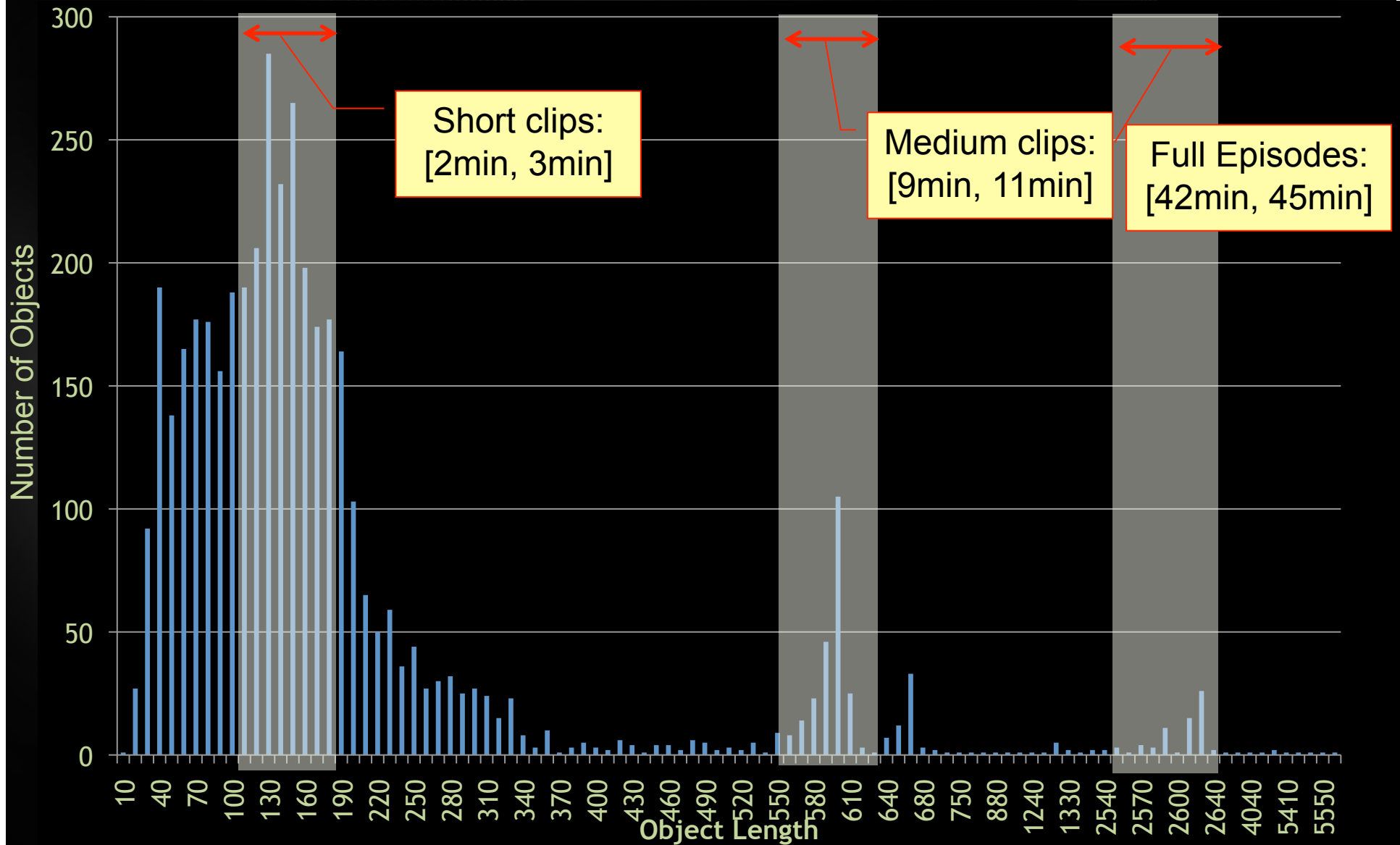
⏻ Load:

- Four channels of a premier video-on demand (VoD) content producer
 - Four days
 - Number of sessions (views): 1,176,049
- A large live event: ~250,000 concurrent viewers

⏻ Metrics

- Content length distribution
- Viewer Hour Loss (VHL): number of viewer hours lost due to quality issues

VoD Object Length Distribution



Quality Metrics

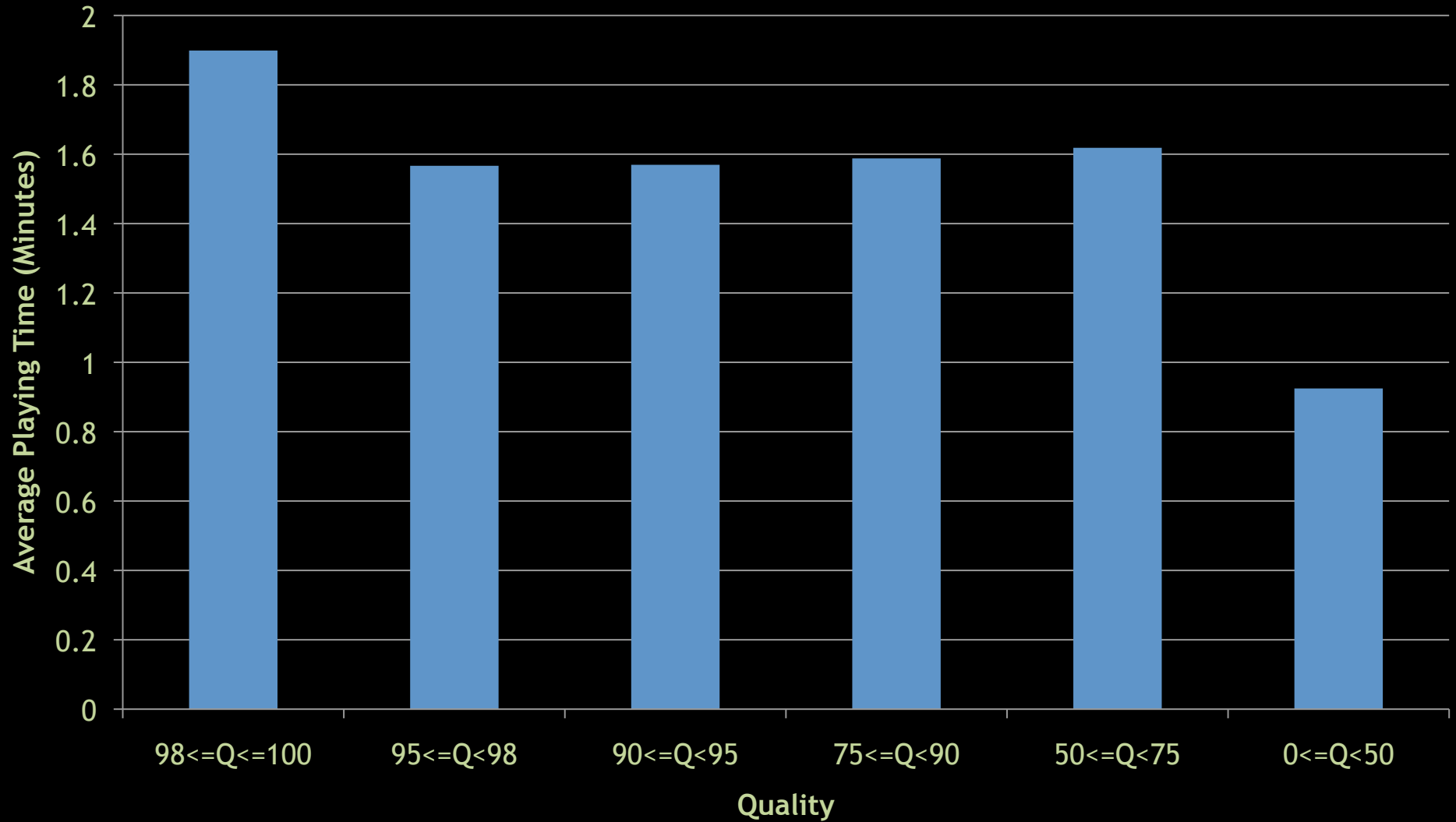
- ⦿ Buffering Quality (BQ):
 $\text{PlayingTime} / (\text{PlayingTime} + \text{BufferingTime})$
- Rendering Quality (RQ):
 $\text{RenderingRate} / \text{EncoderRate}$
- Good session
 - BQ > 95%
 - RQ > 60%

Analysis Underestimates Quality Impact

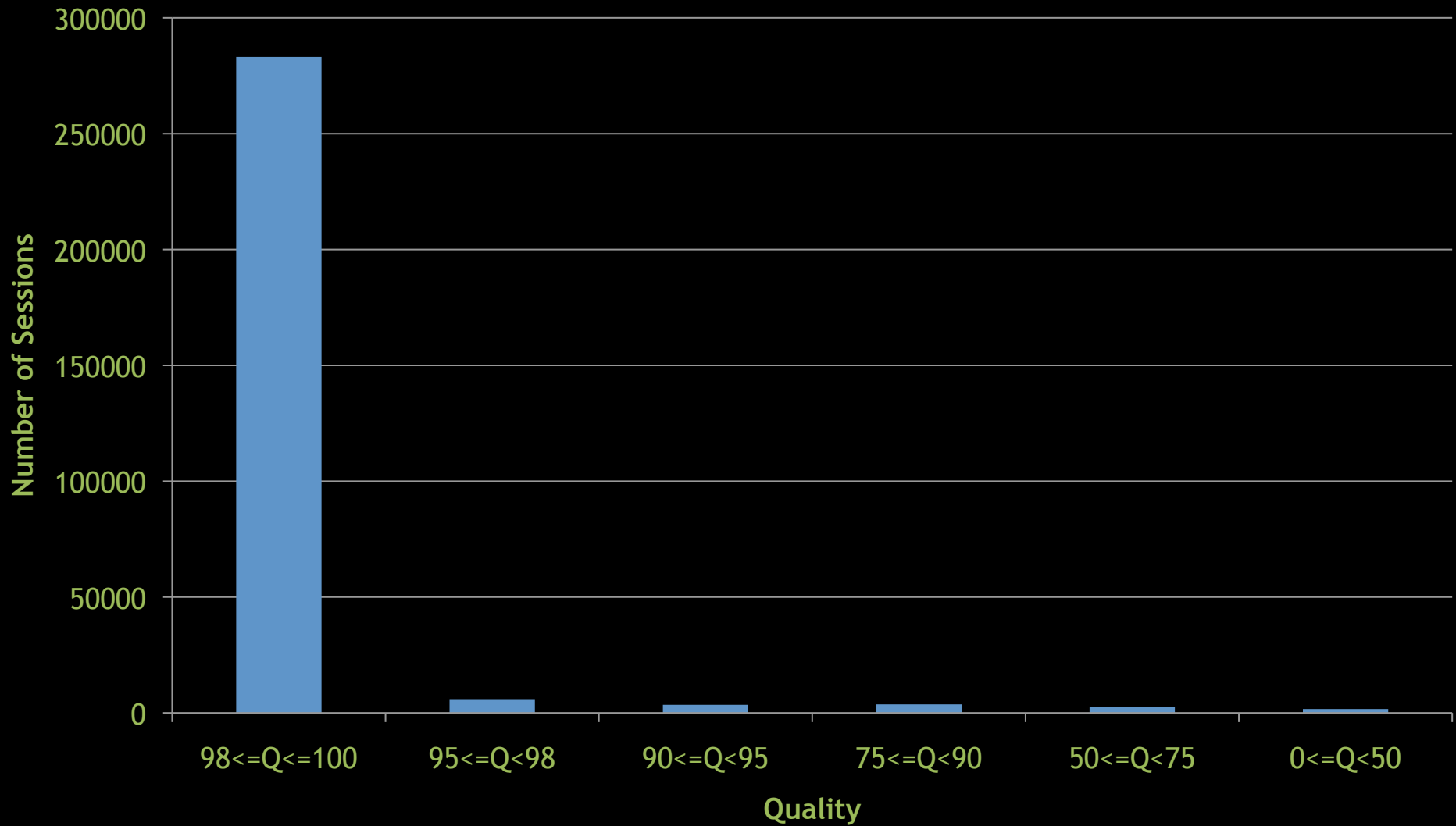
- ⦿ For most analysis use BQ only
 - RQ only a small part of quality issues due to low bit rate (500-700Kbps)

- ⦿ Ignore connection failures

Short Clip (2-3min) Analysis



Short Clip (2-3min) Analysis

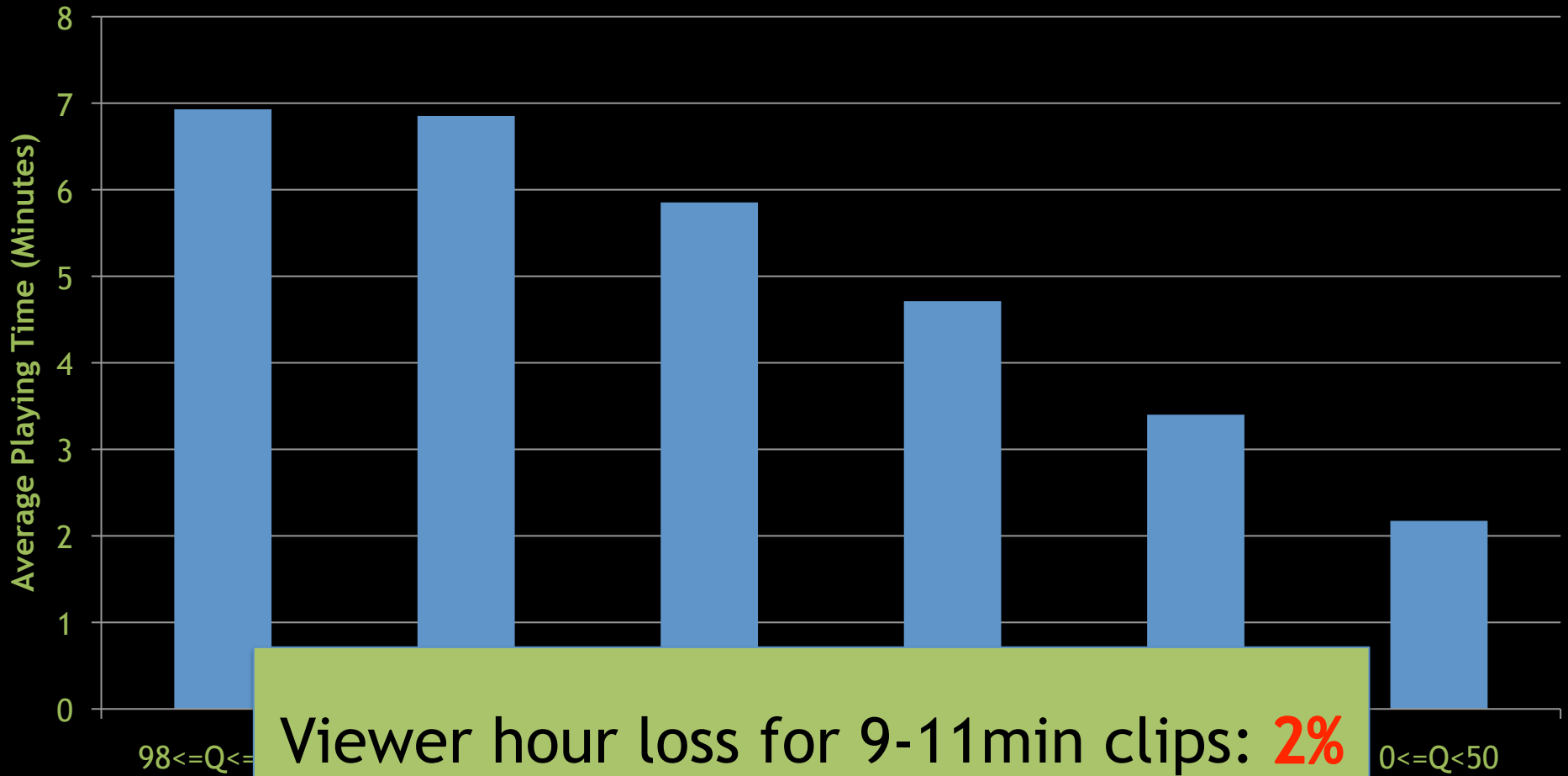


Viewer Hour Gain

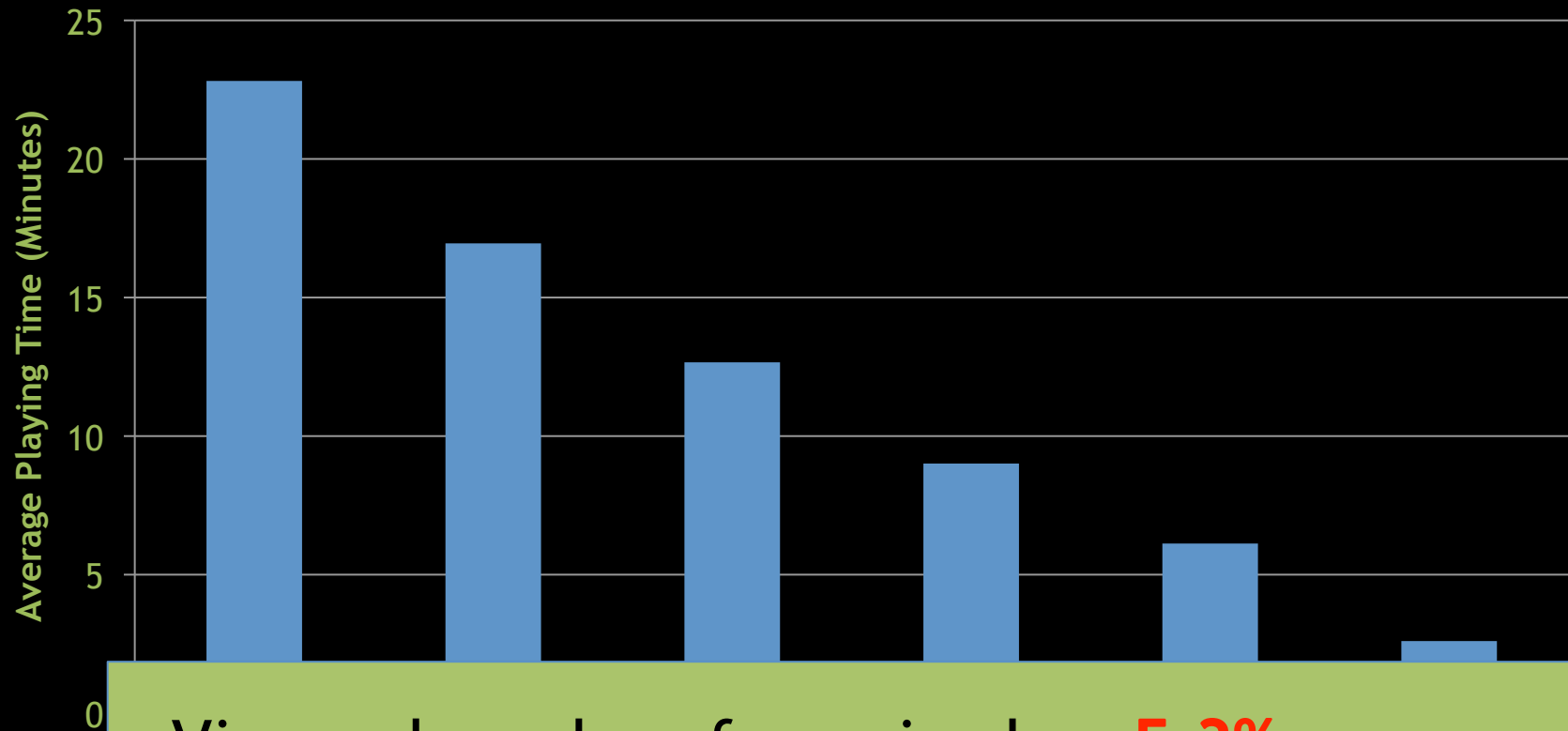
- Ⓞ D : Average duration of sessions with high quality (0.98 ≤ quality < 1)
- Ⓞ D_q : Average duration of sessions with quality = q
- Ⓞ N_q : Number of sessions with quality = q
- Ⓞ Viewer hour gain for sessions with quality q
$$N_q \times (D - D_q)$$
- Ⓞ Total viewer hour gain

Viewer hour loss for 1-2 minute clips: **1.2%**

Medium Clip (9-11min) Analysis

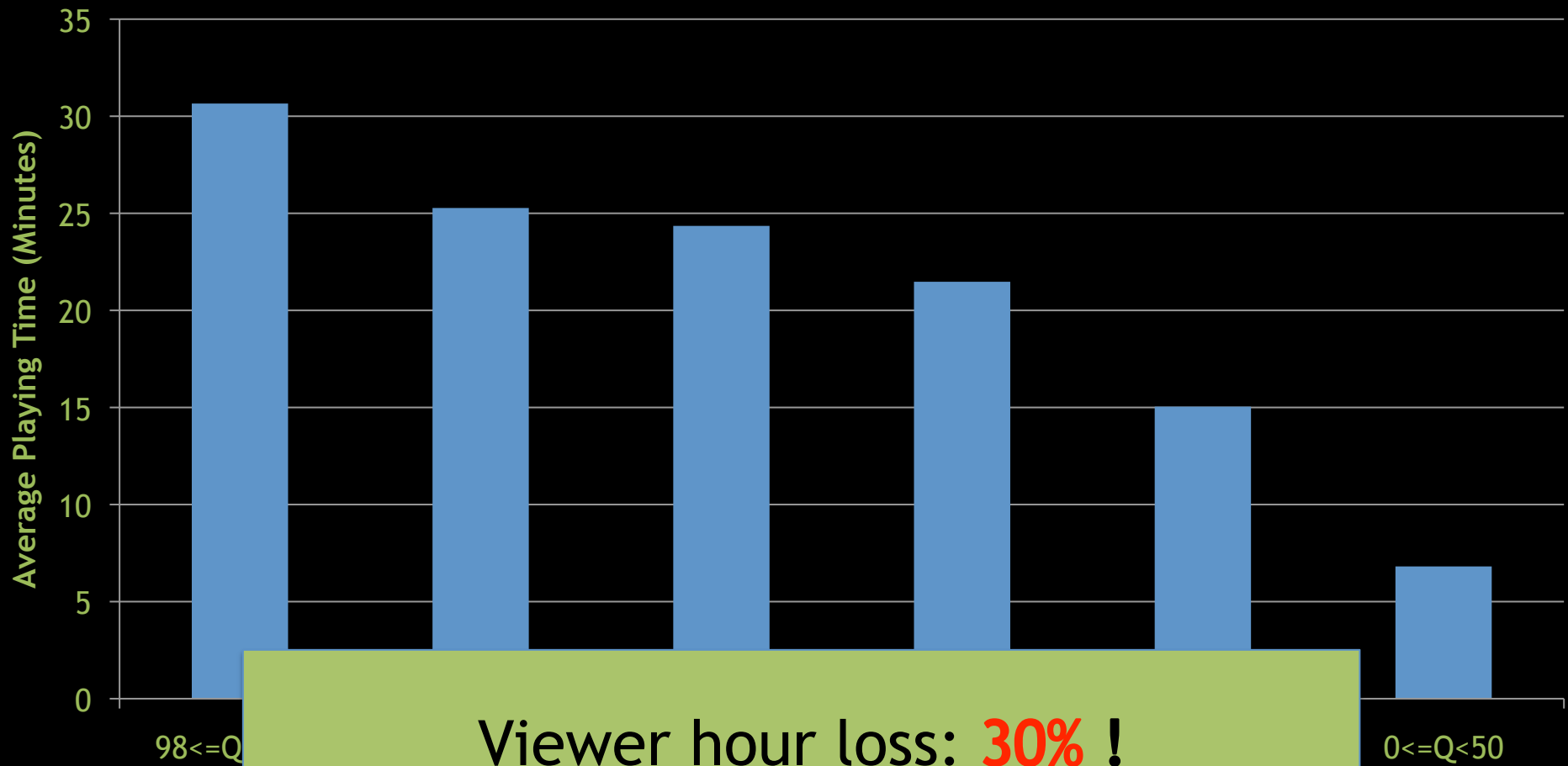


Full Episodes (42-45min) Analysis

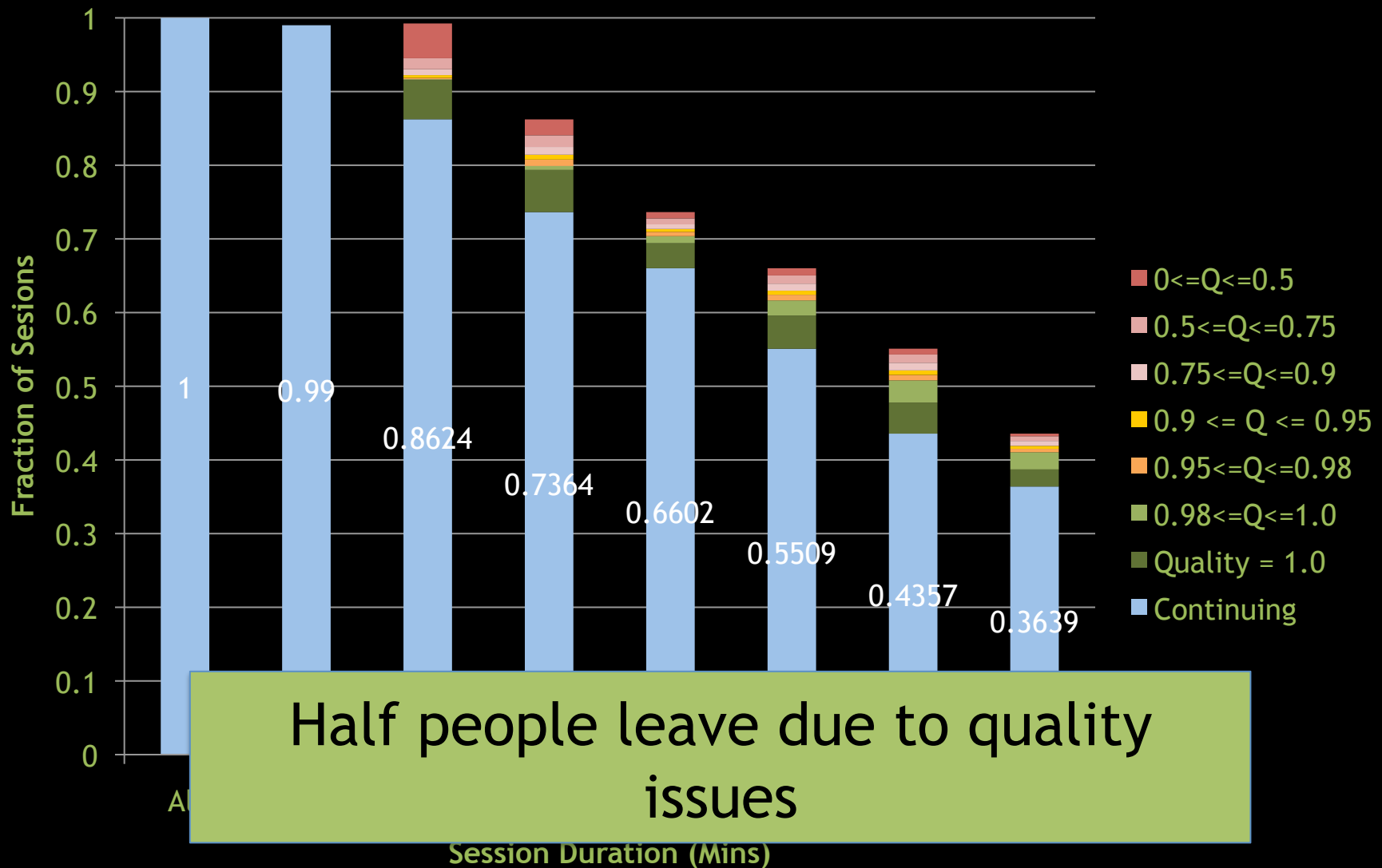


- Viewer hour loss for episodes: **5.3%**
- Viewer hour loss for all content: **4.8%**

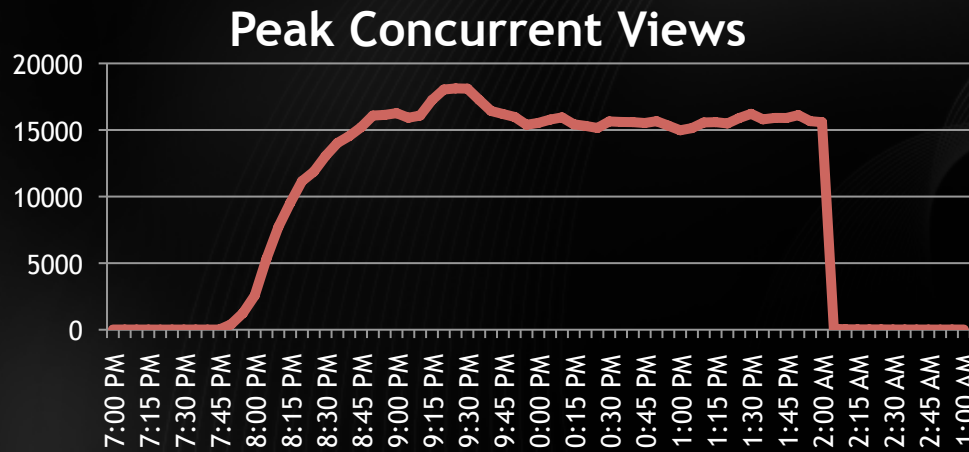
Large Scale Live Event



Large Scale Live Event: Engagement Funnel



Another Case Study: Live Event



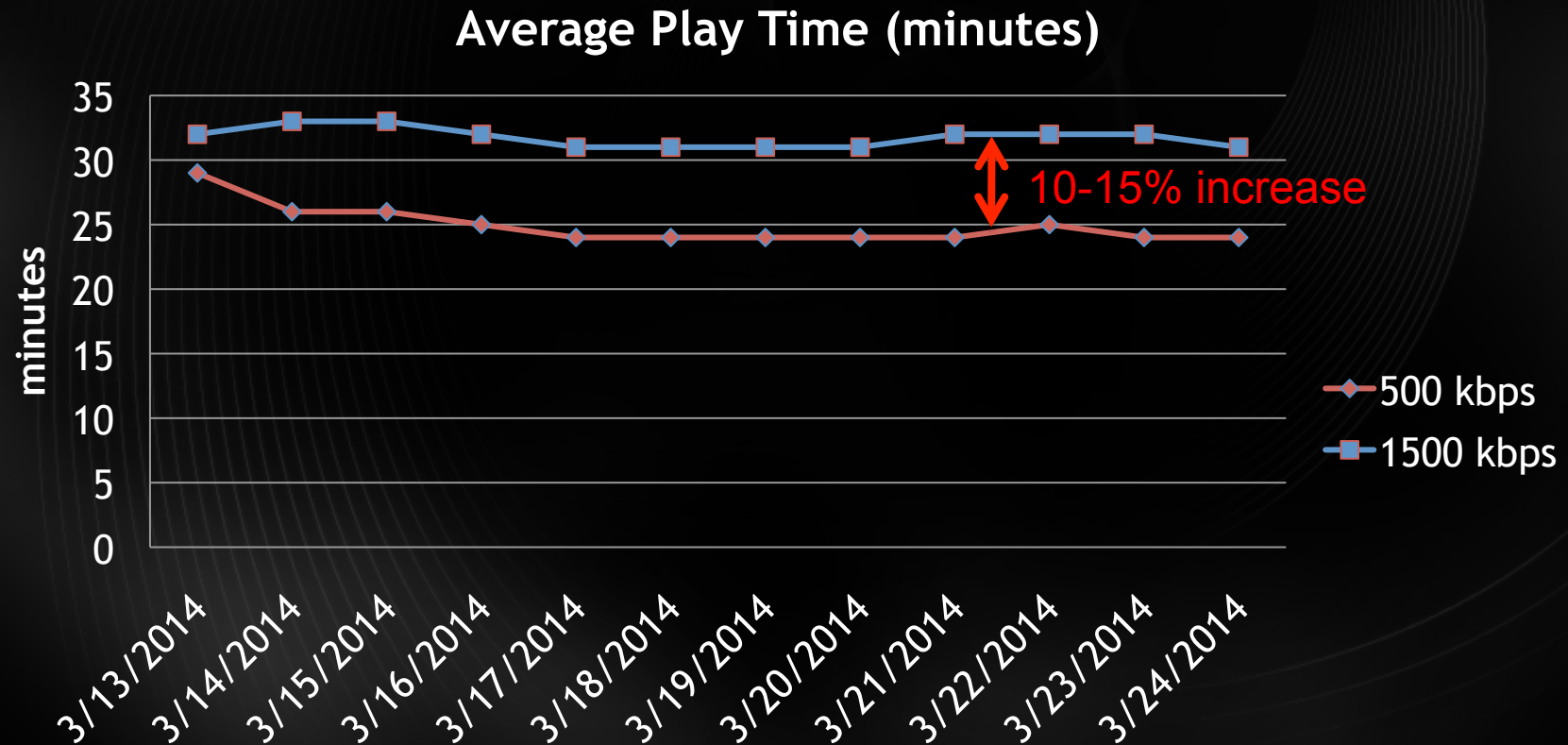
Total sessions	151,980
Unique viewers	73,942
Sessions per viewer	1.9
Total viewer hours	58,436

	Quality	Engagement
Total views	151,980	25 minutes
Failed views	13,815 (9%)	0 minutes
Quality impacted views	21,584 (14%)	16 minutes
Good views	116,581 (77%)	27 minutes
Unique viewers	75,328	48 minutes
Failed viewers	1,386 (2%)	0 minutes
Quality impacted viewers	14,309 (19%)	30 minutes
Good viewers	59,633 (79%)	51 minutes
Total viewer hours		58,436 hours
Lost viewer hours		5,134 hours (9%)

Viewer with poor quality watch 41% less minutes!

Does High Bit Rate Video Help?

- Comparing Engagement of low and high bitrates
 - Viewers watch longer on average on 1500Kbps



Summary

🕒 Quality impact:

- BQ can impact viewer engagement by up to 40%
- Higher bit-rates can increase viewer engagement by up to 15%

🕒 Engagement loss due to quality issues: between 4 and 30%

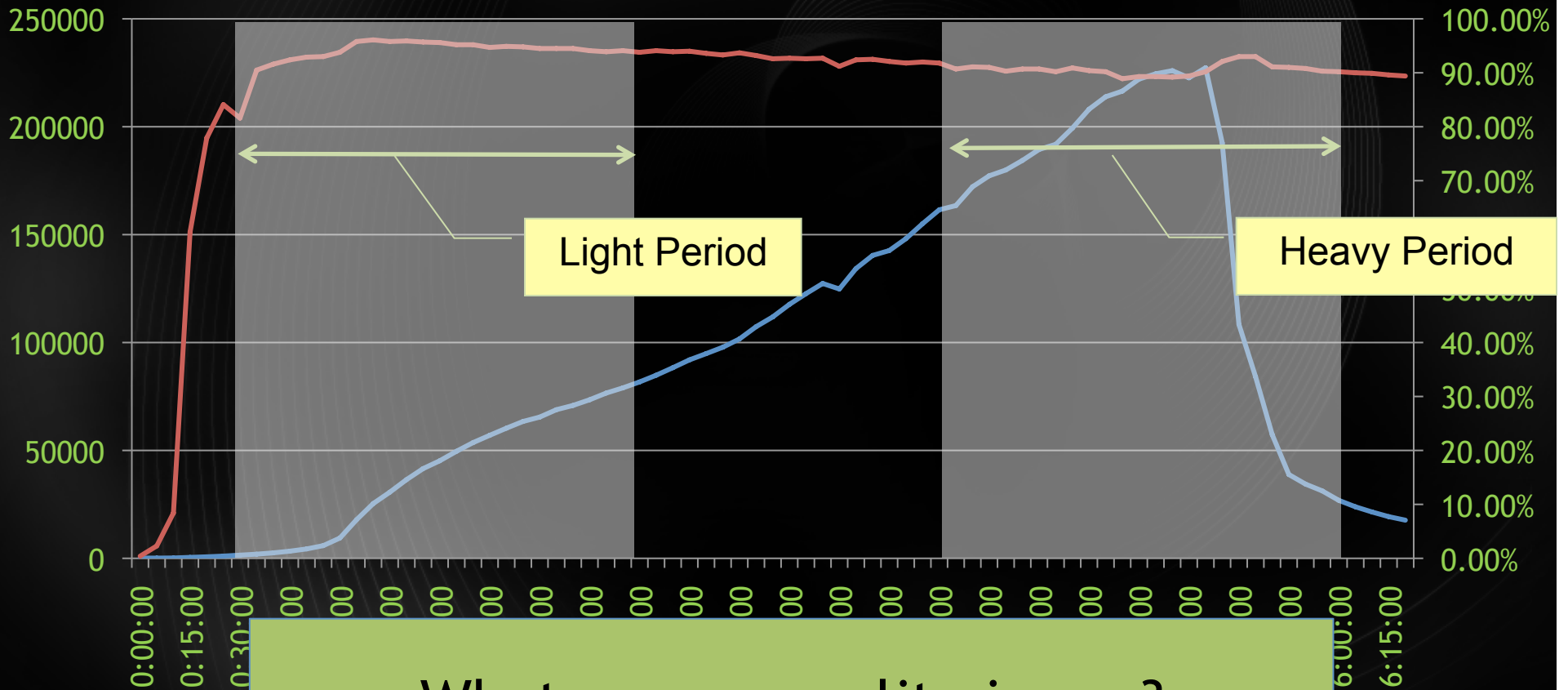
- Even a 4% improvement, may offset distribution costs
- Ignore other quality issues, like connectivity and media failures

The background is a dark, almost black, gradient. It features a series of concentric, slightly irregular circles that create a tunnel-like or ripple effect. In the center of these circles, there is a faint, dark silhouette of a person's head and shoulders, looking forward. The overall aesthetic is modern and technical.

Root Cause Analysis

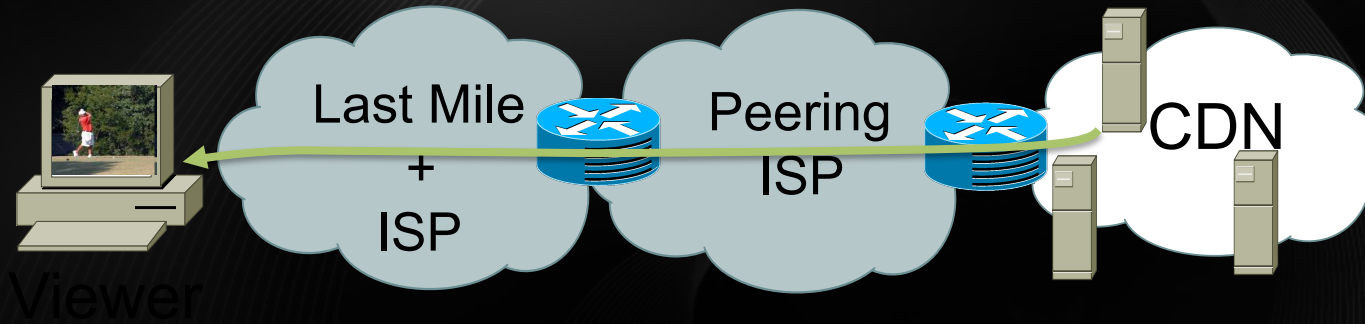
CONVIVA

Viewers vs. Buffering Quality



What causes quality issues?

Root Cause Analysis



⏻ Root cause a quality issue to:

- Viewer machine (CPU)
- Last mile + ISP (Autonomous System Number)
- CDN

⏻ Note:

- Cannot differentiate between edge and core ISPs
- Use only passive measurements, no IP traceroute

Metrics and Definitions

Quality metrics

- Buffering quality (BQ)
 - $\text{playing time} / (\text{playing time} + \text{buffering time})$
- Rendering quality (RQ)
 - $\text{rendering frame rate} / \text{encoded frame rate}$

Session classification:

- Good: $(\text{BQ} \geq 95\%) \text{ AND } (\text{RQ} \geq 60\%)$
- Low BQ: $(\text{BQ} < 95\%)$
- Low RQ: $(\text{BQ} \geq 95\%) \text{ AND } (\text{RQ} < 60\%)$

Methodology: Root Causing Viewer Machine

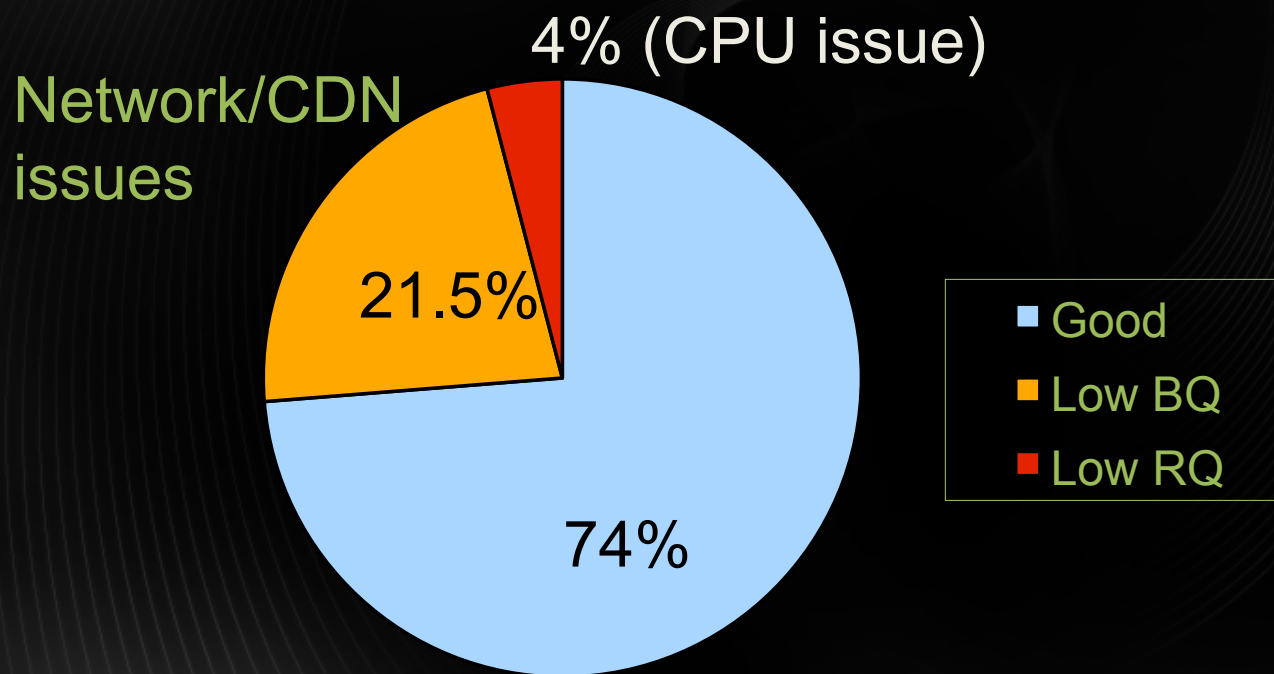
🔌 CPU likely to be the issue when:

- Rendering quality low
- Buffering quality high

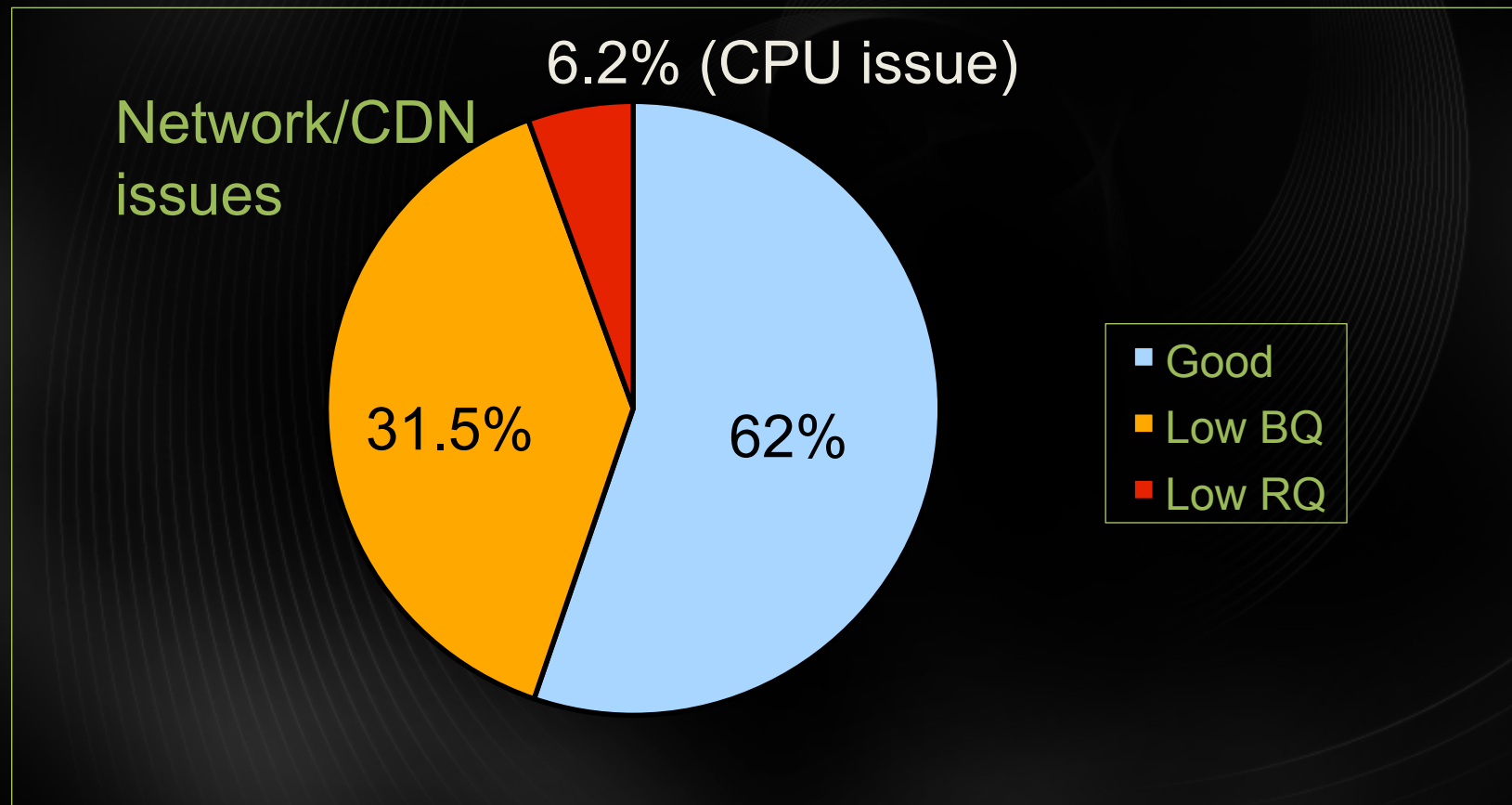
🔌 Conclude CPU is the issue when session's

- $RQ < 60\%$
- $BQ > 95\%$

Quality Issues: Light Period



Quality Issues: High Period

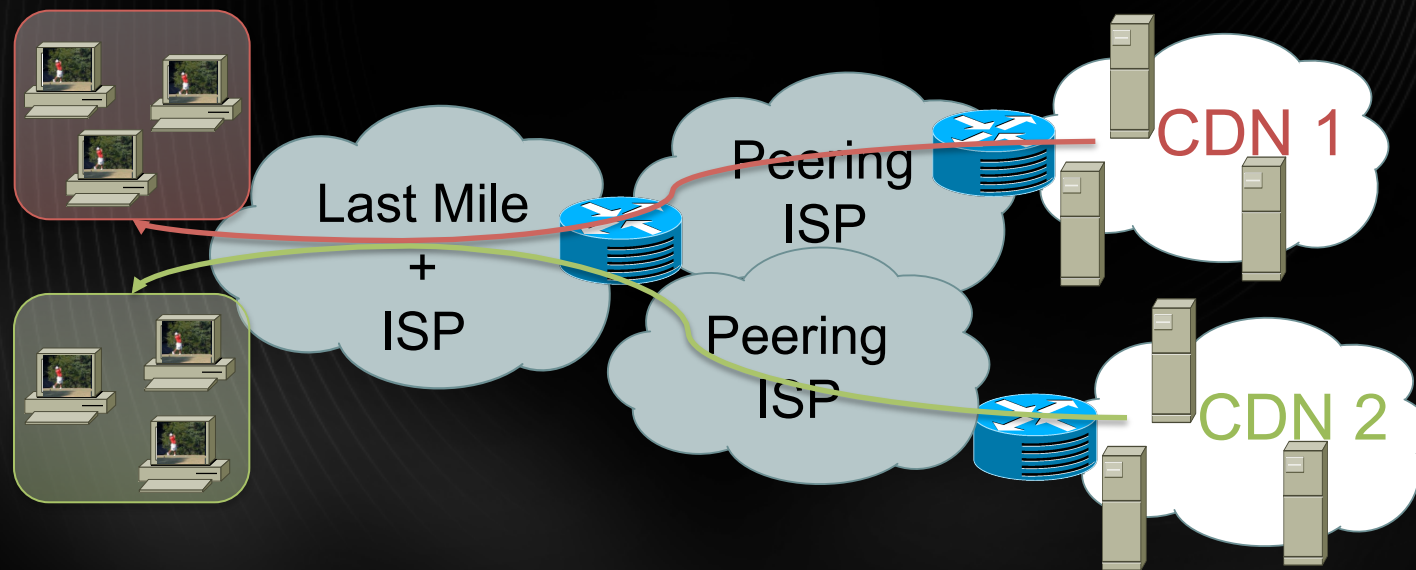


Explaining Buffering Issues

- ⌚ Assume buffering quality issues are either due to:
 - CDN, or
 - ISP
- ⌚ Recall: a session has buffering quality issues if
 - $BQ < 95\%$

Methodology: Root Causing CDN (1/2)

- 🕒 Viewers connected to same ASN but using two CDNs
- 🕒 Intuition: if quality experienced by CDN 1 viewers is significantly lower than of CDN 2 viewers for same ASN, CDN 1 has quality issues



Methodology: Root Causing CDN (2/2)

- 🔌 Select all ASNs who have more than 50 sessions for each CDN
- If difference between quality of viewers in CDN1 and CDN2 for same ASN is $> 10\%$
 - Lower quality CDN is root cause at current time

Methodology: Root Causing ASN/ISP

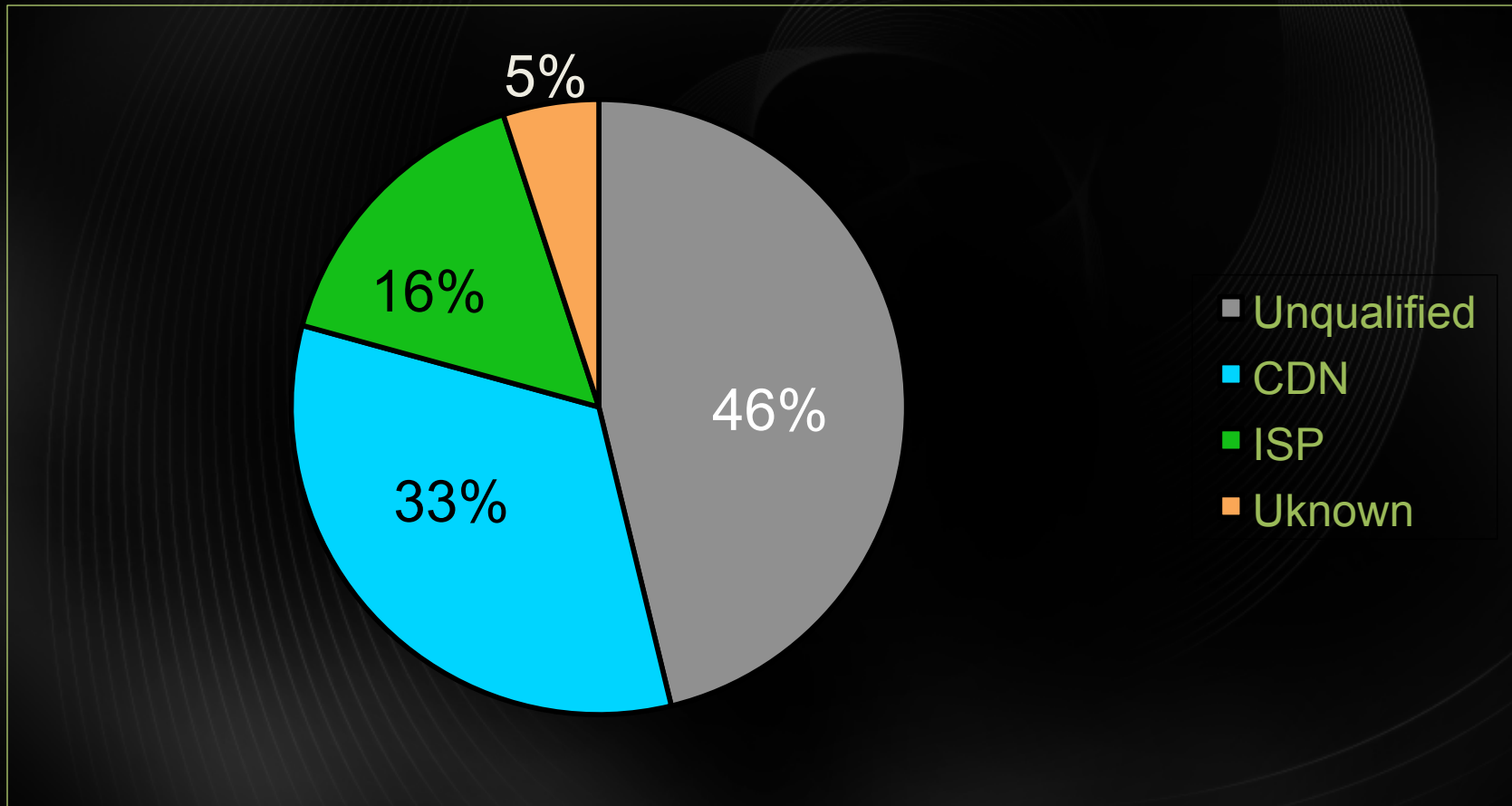
🕒 Two CDNs:

- Conclude ASN A has quality issues if ASN A's viewers connected to either CDN1 or CDN2 experience “bad quality”
- Average quality of viewers connected to other ASNs higher

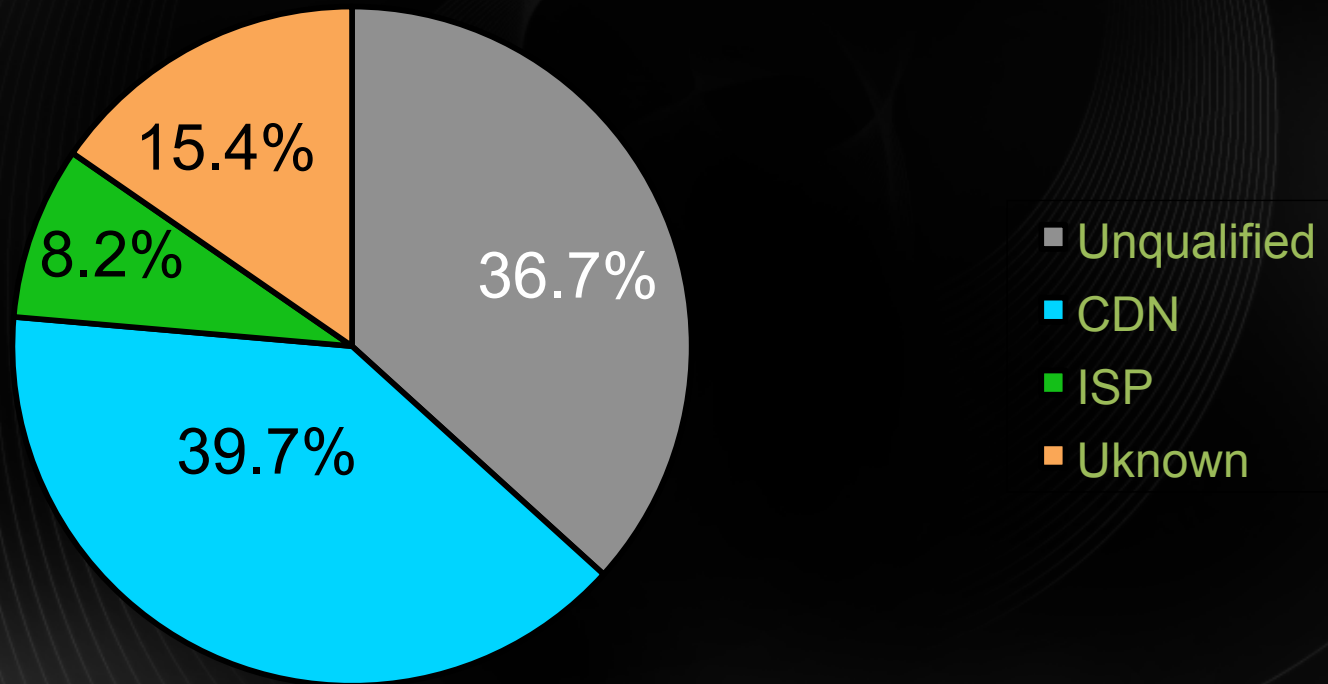
🕒 One CDN

- ASN A's viewers connected to CDN have much lower quality than the average quality of viewers connected to CDN

Buffering Quality: Light Period



Buffering Quality: Heavy Period



Some Findings

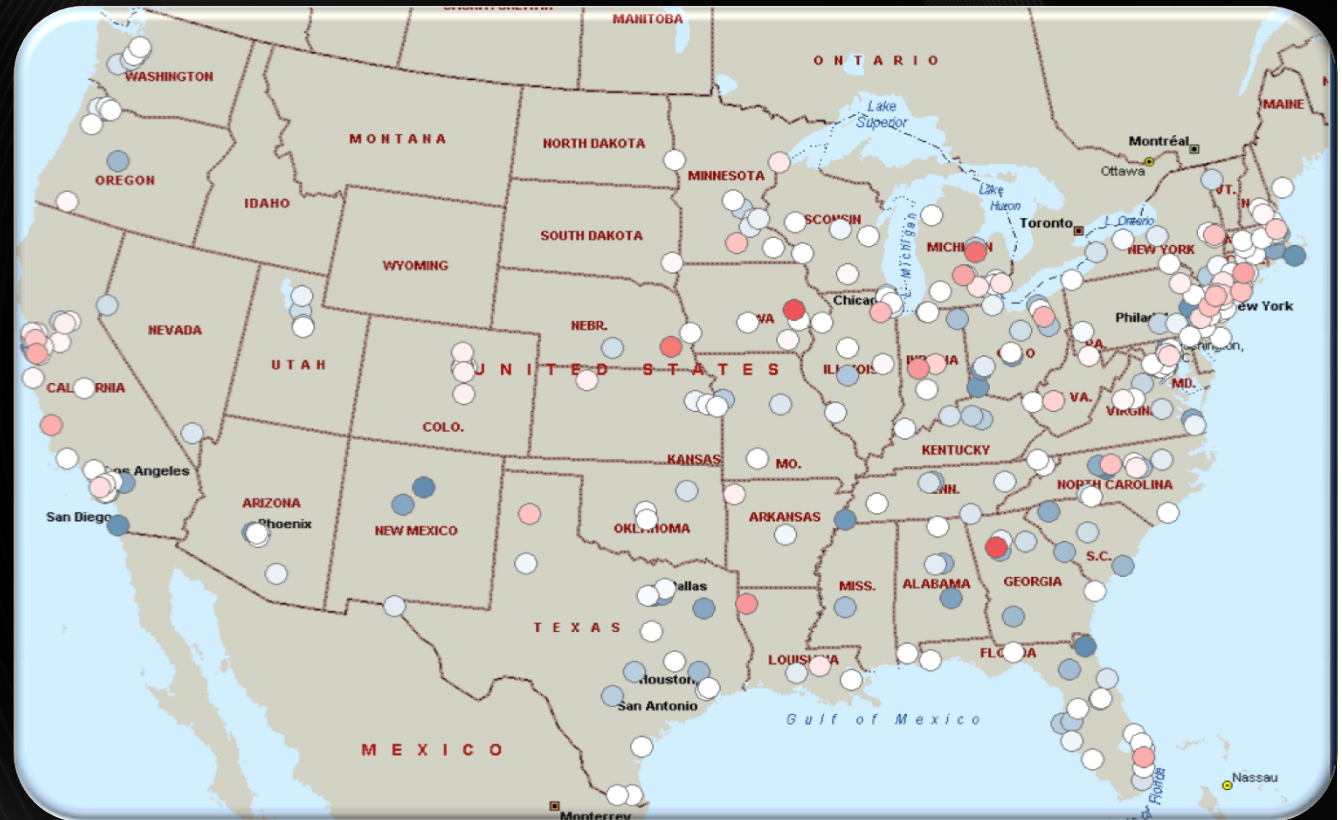
- ⌚ Most of ASNs who had quality issues were enterprise ASNs
 - Expected given that the large scale event was during the workday
 - One ASN had 44% buffering quality!
- ⌚ No CDN was uniformly bad
 - (see next)

CDN Comparison

- Ⓞ Quantify quality difference between CDNs
- Ⓞ Methodology:
 1. Select all ASNs which have more than 50 sessions on both CDNs
 2. Compute average quality for CDN1 and CDN2 viewers per ASN
 3. Order ASNs by difference in quality between CDN1 and CDN2

Internet delivery is more variable than realized...

- Content Delivery Networks **all** have problems sometime
- Even in the same viewer session the **best quality** changed many times during the event



● CDN 1 was best

● CDNs were even

● CDN 2 was best

Summary

24-38% of Total Sessions have Quality Issues	Quality Issues Classification	Solution
	CDN (7-12% of total sessions)	Resource switching
	End-Host CPU (4-6% of total sessions)	Bit-rate switching
	ISP (2-3% of total sessions)	Localize traffic, bit-rate switching
	Unqualified (9-11% of total sessions)	Mitigated by above
	Unknown (1-4% of total sessions)	N/A

Conclusions

- Ⓢ At least for premium content
 - Reducing cost is important, but...
 - ... improving quality is even more important
- Ⓢ P2P can play an important role
 - Localize traffic
 - Highly robust to source failures
- Ⓢ Great opportunity
 - Adobe has announced full p2p support for Flash Player 10.1
 - No need for client download!