It’s Not the Cost, It’s the Quality!

Ion Stoica

Conviva Networks and UC Berkeley
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
  - USA
  - NBC
  - ABC
  - USTREAM
  - MLB
  - Vancouver 2010
  - NFL
  - ESPN360
  - MSNBC

- CDNs
  - Akamai
  - Limelight Networks
  - Level3 Communications
  - AT&T
  - Highwinds

CONVIVA
Trends
Trends: CDN Pricing

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

[Graph showing a downward trend in CDN pricing from 2006 to 2010.]
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, 2\textsuperscript{nd} half)
How Does HTTP Chunking Work?

origin

CDN1

CDN2

ISP A

ISP B

http cache

Viewer

Viewer

Viewer
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**
Quality Matters
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

- **Short clips:** [2min, 3min]
- **Medium clips:** [9min, 11min]
- **Full Episodes:** [42min, 45min]
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

Average Playing Time (Minutes)

<table>
<thead>
<tr>
<th>Quality</th>
<th>Average Playing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>98&lt;=Q&lt;=100</td>
<td>1.8</td>
</tr>
<tr>
<td>95&lt;=Q&lt;98</td>
<td>1.6</td>
</tr>
<tr>
<td>90&lt;=Q&lt;95</td>
<td>1.5</td>
</tr>
<tr>
<td>75&lt;=Q&lt;90</td>
<td>1.4</td>
</tr>
<tr>
<td>50&lt;=Q&lt;75</td>
<td>1.3</td>
</tr>
<tr>
<td>0&lt;=Q&lt;50</td>
<td>1.0</td>
</tr>
</tbody>
</table>
Short Clip (2-3min) Analysis

![Bar chart showing the number of sessions for different quality ranges.]

- **98<=Q<=100**: 300,000 sessions
- **95<=Q<98**: 100 sessions
- **90<=Q<95**: 100 sessions
- **75<=Q<90**: 200 sessions
- **50<=Q<75**: 500 sessions
- **0<=Q<50**: 1,000 sessions
Viewer Hour Gain

- **D**: Average duration of sessions with high quality ($0.98 \leq \text{quality} < 1$)
- **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

Viewer hour loss for 9-11min clips: 2%
Full Episodes (42-45min) Analysis

- Viewer hour loss for episodes: **5.3%**
- Viewer hour loss for all content: **4.8%**
Large Scale Live Event

Viewer hour loss: 30%!
Large Scale Live Event: Engagement Funnel

Half people leave due to quality issues
Another Case Study: Live Event

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

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

Average Play Time (minutes)

- 10-15% increase

- 500 kbps
- 1500 kbps
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
Root Cause Analysis
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)
  - playing time/(playing time + buffering time)
- Rendering quality (RQ)
  - rendering frame rate/encoded frame rate

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

- Good: 74%
- Low BQ: 21.5%
- Low RQ: 4% (CPU issue)
- Network/CDN issues

Legend:
- Good
- Low BQ
- Low RQ
Quality Issues: High Period

- Good: 62%
- Low BQ: 31.5%
- Low RQ: 6.2% (CPU issue)
- Network/CDN issues: 31.5%

Breakdown of Quality Issues:
- Good: 62%
- Low BQ: 31.5%
- Low RQ: 6.2% (CPU issue)
- Network/CDN issues: 31.5%
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

- Unqualified: 46%
- CDN: 33%
- ISP: 16%
- Unknown: 5%
Buffering Quality: Heavy Period

- Unqualified: 36.7%
- CDN: 39.7%
- ISP: 8.2%
- Unknown: 15.4%
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

<table>
<thead>
<tr>
<th>Quality Issues Classification</th>
<th>Solution</th>
</tr>
</thead>
<tbody>
<tr>
<td>CDN</td>
<td>Resource switching</td>
</tr>
<tr>
<td>(7-12% of total sessions)</td>
<td></td>
</tr>
<tr>
<td>End-Host CPU</td>
<td>Bit-rate switching</td>
</tr>
<tr>
<td>(4-6% of total sessions)</td>
<td></td>
</tr>
<tr>
<td>ISP</td>
<td>Localize traffic, bit-rate switching</td>
</tr>
<tr>
<td>(2-3% of total sessions)</td>
<td></td>
</tr>
<tr>
<td>Unqualified</td>
<td>Mitigated by above</td>
</tr>
<tr>
<td>(9-11% of total sessions)</td>
<td></td>
</tr>
<tr>
<td>Unknown</td>
<td>N/A</td>
</tr>
<tr>
<td>(1-4% of total sessions)</td>
<td></td>
</tr>
</tbody>
</table>
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!