Cache–Friendly Streaming Bitrate Adaptation by Congestion Feedback in ICN

Dinh Nguyen¹, Jiong Jin², and Atsushi Tagami¹
¹KDDI R&D Laboratories Japan, ²Swinburne Univ. of Tech. Australia
Content

1. Motivation
2. Fair and stable bitrate adaptation
3. Impact on caching
4. Evaluation
5. Conclusion and future work
Dynamic Adaptive Streaming over HTTP

Player (HTTP Client)

Source (HTTP Server)

Bitrate Adaptation Logic

TCP Receiver

Network

TCP Source
DASH over CCN

Player (CCN Consumer)

Source (CCN Repository)

Bitrate Adaptation Logic

Conventional bitrate adaptation!

The problem

- Bitrate adaptation results in low cache hit
  - Vanilla ICN caches admit every video segment
  - Non-repetitive requests from users due to unfair and unstable bitrate adaptation
  - Multiple representations of the same segment
Motivation

- Avoid the *side-effect* of bitrate adaptation to cache
  - High cache hit rate even with adaptive streaming
- Bitrate adaptation should be *friendly* to in-network caches
  - Repetitive requests: same bottleneck same bitrate
  - Less fluctuation: quickly settle at fair, stable bitrate
- To devise a streaming bitrate adaptation with
  - Fairness
  - Stability
Conventional bitrate adaptation

1. Estimate bandwidth in the last download

\[ \tilde{x} = \frac{\text{segment.size}}{\text{download.time}} \]

Estimate bandwidth

2. Smooth the estimated bandwidth using e.g. EWMA

EWMA(\(\tilde{x}\))

Smooth

3. Choose nearest bitrate as the target video bitrate

\(\hat{x} \leq \text{EWMA}(\tilde{x})\)

Quantize

4. Request next segment of bitrate \(\hat{x}\) after

\[ \Delta t = \begin{cases} 
0 & \text{if } B < B_{\text{max}} \\
\tau & \text{if } B \geq B_{\text{max}} 
\end{cases} \]

Schedule next request

/* \(B/B_{\text{max}}\): current/max playback buffer */

/* \(\tau\): segment length in second */
A simple test

- 8 conventional adaptive streams compete over 24Mbit/s bottleneck

Inaccurate bandwidth share estimation

Unfair, unstable bitrates
Inaccurate bandwidth estimation

\[ \hat{x} = X \]

\[ \hat{x}_1 = \hat{x}_2 = X \]

Unstable

Unfair

Copyright (c) KDDI R&D Labs. All Rights Reserved.
Utility fairness resource allocation

- Utility fairness optimization [Wang06]

\[
\max_{x \geq 0} \sum_{n \in N} \int_{m_n}^{x_n} \frac{1}{U_n(y)} dy
\]

s.t. \( \sum_{n:i,j \in L(n)} x_n \leq c_{i,j} \quad \forall i, j \)

- Utility-fair bandwidth share

\[
x_n(t + 1) = U_n^{-1}\left(\frac{1}{q_n(t)}\right)
\]

congestion feedback

Utility functions

- Step-wise utility functions
- Logistic approximation at each step for continuity
- Could be different for users requesting the same content
Utility–fair bitrate adaptation

1a. Measure and smooth congestion price feedback

\[ q_n = \text{EWMA}[q(D)] \]

/* \( q_n/q(D) \): smooth/current congestion price */

1b. Compute utility-fair bandwidth share

\[ \hat{x}_n = U_n^{-1} \left( \frac{1}{q_n} \right) \]

2. Smooth the estimated bandwidth

\[ \text{EWMA}(\hat{x}) \]

3. Choose nearest bitrate as the target video bitrate

\[ \hat{x} \leq \text{EWMA}(\hat{x}) \]

4. Request next segment of bitrate \( \hat{x} \) after

\[ \Delta t = \begin{cases} 
0 & \text{if } B < B_{\text{max}} \\
\tau & \text{if } B \geq B_{\text{max}} 
\end{cases} \]

/* \( B/B_{\text{max}} \): current/max playback buffer */
/* \( \tau \): segment length in second */
Impact on caching (ext. to Che’s)

- Fewer # representations in cache $R$
- Larger effective cache size

$$\bar{C} = \frac{C}{S \left( \sum_{k=1}^{R} b_k \right)}$$

- Longer characteristic time $T_{\bar{C}}$

$$\sum_{i=1}^{M} \left( 1 - e^{-p_i T_{\bar{C}}} \right) = \bar{C}$$

- Higher cache hit rate

$$P_{hit} = \sum_{i=1}^{M} p_i \left( 1 - e^{-p_i T_{\bar{C}}} \right)$$

Copyright (c) KDDI R&D Labs. All Rights Reserved.
Simulation

- Chunk-level CCN simulation
- Congestion signal: queuing delay
- LRU caches
- Chain topology with bottleneck
- Playback buffer: 30 seg.

Evaluate
  - Conventional adaptation
  - PANDA [Li2014]: AIMD bandwidth probe
  - Proposed utility–fair adaptation

Elastic background traffic

- 24 Mbit/s bottleneck
- 8 streaming sessions + 1 download
Stability and fairness

- 1Gbit/s bottleneck, chain topology
- 60% streaming + 40% downloading
- 2000 content objects (Zipf.8), Poisson arrival 0.8 req/s
- 20GB LRU cache @ all routers
Cache-friendliness

Much higher hit ratio

Fewer redundant representations
Conclusion

- Promising result on bitrate adaptation by congestion feedback in CCN
  - **Accurate bandwidth estimation**: using utility fairness framework
  - **Cache-friendliness**: fairness and stability in bitrate adaptation increase cache hit

- Future work
  - Testbed evaluation
  - Caching for video streaming
  - Exploiting better congestion feedback signal available in CCN/NDN
Thank you!