Caching with Named Data Networking: A Case Study of Youku

Sen Wang, Jun Bi, Jianping Wu
Institute for Network Sciences and Cyberspace
Tsinghua University
Motivation

• In this paper, we attempt to transparently caching HTTP traffic within a NDN network
  – In particular, we conduct a case study of caching the video traffic of Youku
    • The largest website for user generated video in China
  – Analyze a large real trace from an access network
• Provide relatively large-scale testbed and real traffic
• Combine the new design with the traditional Internet
• Create Incentive to deploy NDN network
System design

• As an adaptor of NDN and HTTP, the gateway must be both compatible with NDN protocol and HTTP protocol.
• The HTTP-NDN gateway consists of two parts
  – Ingress gateway and Egress gateway
GET method

- A GET request
  - Transformed into an interest packet in IG and recovered in EG, then it is sent to the web server.
- HTTP response
  - Transformed into a bunch of Data packets and travels back to IG, where the Data packet is extracted and sent to the client.
Name Translation

• Requests for video chunk
  – Account for 99.14% of the total traffic
  – Each video chunk of *youku* can be identified by a unique string in the URL
    – Name the video chunk with this unique string instead of the URL
  – Extend the Interest of NDN with a new attribute named *metadata* to carry the complete URL and the headers

• Requests for others
  – Do not cache
  – Name them with URL and all the headers with it
Cut video chunks into pieces

• A video chunk of youku
  – About 7 minutes and tens of Mbytes
• Four ways to request a part of a video chunk
  – A header in http request
    • "Range: bytes=5794837-5853076"
  – Parameters in URL
    • "xxxx.flv?start=4"
    • "xxxx.flv? preview_ts=12"
    • "xxxx.flv?ts_start=10&ts_end=30 &ts_seg_no=1"
• Cut video chunks of youku into ccnx chunks
  – According to informations of both time and location of key frame inside the video
  – ccnx://video_chunk/03000201004F759257B7F300946C19B8D94FF0-563F-4D8D-A687-35BEBCA497B4.flv/second/10/ccn_chunk/1
• We need a translation from the time-based naming to byte-based request
Adapt HTTP caching into NDN

- **NDN**
  - A requester can indicate whether this Interest can be satisfied by a still-fresh cached copy
    - *AnswerOriginKind*
  - An age-based way to estimate the freshness of a cached copy
    - *FreshnessSeconds*

- **HTTP**
  - An similar age-based way
  - A “validation” mechanism to reduce network bandwidth requirements
    - Send a validating request to the server
    - Meaningless for NDN with much smaller chunk

- **Cache consistency for youku traffic**
  - The video traffic of *youku* offers two headers
    - “Last-Modified”
    - “ETag”
  - No similar mechanism in NDN
  - Estimate the updating interval for every object at EG with “Last-Modified”.
    - Set the *FreshnessSeconds* according to the esitimation
Preliminary results

- Two video types
  - Flv and Mp4
- Average size
  - Entire Video Chunk
  - Partial Chunk

- Uniq ID Number
- ID requested once
- Request Number
Future work

• Compute the cacheability
  \[ Cacheability = \frac{\sum_{i=1}^{n}(f_i - 1) \cdot s_i}{\sum_{i=1}^{n} f_i \cdot s_i} \]

• Hit ratio vs Cache size
• Request Distribution
  – Its change over time
• Cacheability vs Population Size
Q&A

Thank you!