Anatomy of a Personalized Livestreaming System

Personalized Livestreaming

Travel

Hurricane Matthew

NFL

Interactivity is the Key!
Personalized Livestreaming is Hard

Livestreaming

Video → CDN → Viewer

Personalized Livestreaming

Video

Viewer

Low video delay

Interactions

Viewer

Is scalability going to be a problem?
Outline

- Scale, growth, usage pattern
  - Periscope and Meerkat

- Fundamental tradeoffs between scalability and interactivity
  - Dissection of Periscope

- Recent changes of Periscope
Available Data

- Interactions within broadcasts
  - Comments and hearts
  - Viewer joining
  - All events have timestamps

- Livestream video
  - Arrival time of each video frame

- Social graph
  - Follower/followee relationship
Data Collection Methodology

- Broadcast meta data
  - Script fully emulating Periscope and Meerkat API calls
  - Monitor public list for new broadcasts

- Interactions within broadcasts
  - Join new broadcasts and collect interactions

- Interacted frequently with Periscope and Meerkat
  - The companies were informed of the project
  - Identify security vulnerabilities
Broadcast Interaction Dataset Size

- **May 12 2015** to **May 15 2015**
  - All public broadcasts for 1+ month
  - 164K broadcasts from 57K broadcasters
  - 1.9M comments and 525K hearts

- **June 15 2015**
  - Meerkat

- **Aug. 20 2015**
  - All public broadcasts for 3+ months
  - 19.6M broadcasts from 1.8M broadcasters
  - 549M comments and 7.98B hearts

- **Periscope
  - Publicly available**
Scale of Livestreaming Services

- Large daily broadcast traffic

Lost 2 days of data due to a communication bug with the Periscope server.

Periscope: 472k daily broadcast in May 2016

Periscope’s volume tripled in 3 months

Meerkat’s volume halved in 1 month

Meerkat shut down in Sep. 2016

Daily Broadcasts (Periscope)
Broadcasts are Highly Interactive

- ~1.3M Periscope broadcasts: >100 comments & >1k hearts

- “Commenter limit” is constraining
  - Especially in popular broadcasts
  - 57.6k broadcasts: >1k viewers
  - Trending broadcasts: >700k viewers

- More hearts than comments
  - Caused by “100-commenter limit”
    - Only first 100 viewers can comment
    - ≠ viewers want to comment

This limit is correlated with Periscope’s CDN design
Outline

- Scale, growth, usage pattern
  - Periscope and Meerkat

- Fundamental tradeoffs between scalability and interactivity
  - Dissection of Periscope

- Recent changes of Periscope
Anatomy of Periscope

Video Channel

- **Long latency**
- **Broadcaster**
- **Wowza**
- **Fastly**
- **Viewer**
- **Viewer**
- **Viewer**

Messaging Channel

- **Near real-time**
- **Broadcaster**
- **PubNub**
- **Comment**
- **Heart**
- **Viewer**
- **Viewer**
- **Viewer**
- **Viewer**
- **Viewer**
Livestream Delivery Infrastructure

- **Sender** pushes each video frame to the receiver
- **RTMP** for smaller audience
- More viewers = High load
- **Receiver** polls the server for large video chunks
- **HLS** for larger audience

Sender pushes each video frame to the receiver. RTMP is used for smaller audiences. More viewers lead to high load. Receiver polls the server for large video chunks. HLS is used for larger audiences. Scalable but long delay.
• Only the first 100 viewers can comment. Why?

Implications of the Infrastructure

Limit: 100

Low latency
Able to comment

Long latency
Lagging comment
End-to-end Delay in Lab Settings

- 10 broadcasts in stable lab Wi-Fi

Which components of the system cause the long delay in HLS?
Delay Breakdown in Lab Settings

- **RTMP (Wowza)**
  - Upload (0.76s)
  - Buffering (0.016s)

- **HLS (Fastly)**
  - Chunking (1.16s)
  - Wowza2Fastly (3.01s)
  - Last Mile (1.17s)
  - Polling (5.71s)

- **Viewers**
  - Buffering (6.88s)
How to Reduce HLS Delay?

- Reduce major delay components
  - Chunking delay: smaller chunks
  - Polling delay: more frequent polling
  - Buffering: less buffer

- Use measurement to understand
  - How system parameters affect delay components?
  - What other effects on the system?

- Simulation on broadcasts in the wild
  - 16K randomly selected broadcasts
  - Scripts emulate as viewers
Optimizing Client-side Buffering

- Simulation: buffering delay vs. playback smoothness
  - Strategy: pre-buffer P seconds before playback
  - Stalling time: duration of video freezing

- Save 3s buffering delay with low impact on smoothness
Recent Updates of Periscope

- Allow more viewers to comment
  - Reduce HLS delay so HLS viewers can comment

- Major changes of Periscope
  - Smaller HLS chunks: 3s → 2s
  - Smaller polling interval: 3s → 1s
  - Smaller buffer: 9s → 6s

- CDN design and protocol design stay the same
Periscope Dataset Available

- 75 continuous days of Periscope public broadcasts
  - 13.9M broadcasts
  - 416M comments
  - 6.1B hearts

- In each broadcast
  - Meta data: title, time, broadcaster, location, etc.
  - Interactions: comments, hearts, join, invite, etc.

http://sandlab.cs.ucsb.edu/periscope
Today

- The expansion of the Google’ serving infrastructure
- A personalized livestreaming system
- A platform for crowdsourcing web QoE measurements