A Hierarchical Characterization of a Live Streaming Media Workload

Eveline Veloso
Virgílio Almeida
Wagner Meira
Computer Science Department
Federal University of Minas Gerais
Brazil

Azer Bestavros
Shudong Jin
Computer Science Department
Boston University
USA

This paper appears in:
Networking, IEEE/ACM Transactions on

Publication Date: Feb.2006
Volume: 14, On page(s): 133- 146
ISSN: 1063-6692

A Hierarchical Characterization of a Live Streaming Media Workload

- Introduction
- Live Streaming Workload
- Client Layer Characteristics
- Session Layer Characteristics
- Transfer Layer Characteristics
- Representativeness of findings
- Synthesis of live media workloads
- Summary and conclusion

Introduction

- Motivation
  - Characterization and synthetic generation of streaming access workloads -> Fundamental Importance
  - Have been small number of studies but: pre-recorded, stored streams... NON LIVE-STREAM
  - This paper provides a characterization using:
    - Unique data
    - Hundred of thousand of sessions
    - Thousand of users
    - “Reality Show” in Brazil
- Diferences Stored/Live streaming
  - Server overload
    - Stored: Reject new connects / Live: Impossible
  - Bad QoS
    - Stored: Stop and continue later / Live: Impossible
  - Media access patterns
    - Stored (user driven): user decides what to access and when
    - Live (object driven): user just join or leave

Live Streaming Workload I

- Source of the Workload
  - Logs from one month
  - Server: Microsoft Media Server
  - Clients: audio/video from 48 cameras
- Characterization Hierarchy and Terminology
  - Hierarchy of layers
    - Lowest layer: Server receive requests from multiple clients
    - Level up: Request from individual client grouped into sessions
    - Top level: Sessions from individual clients grouped into client behaviours.
  - Characterizing at levels of abstraction
    - 3 levels: client, session, individual transfers
    - Get characterization of:
      - Arrival processes (interarrival times, level of concurrency
      - Access patterns (ON/OFF times)
      - Other (popularity)

Live Streaming Workload II

- Characterization Hierarchy and Terminology
  - Client layer
    - Top layer
    - Focuses client population
    - Characteristics: Nº of clients accessing, interarrival times, relationship between client´s interest and frecuency of access
  - Session layer
    - Individual client
    - Focuses variables governing client session
    - Client session: Interval of time when client request/receive within a Toff (Max time of inactivity
    - Client access patter: ON/OFF periods
  - Transfer layer
    - Bottom layer, zooming an ON session
    - Focuses on individual data transfers
    - ON/OFF: Served/Not served lived objects
    - Characterization: transfer length, Nº of concurrent transfers, interarrival times

Live Streaming Workload III

- Characterization Hierarchy and Terminology

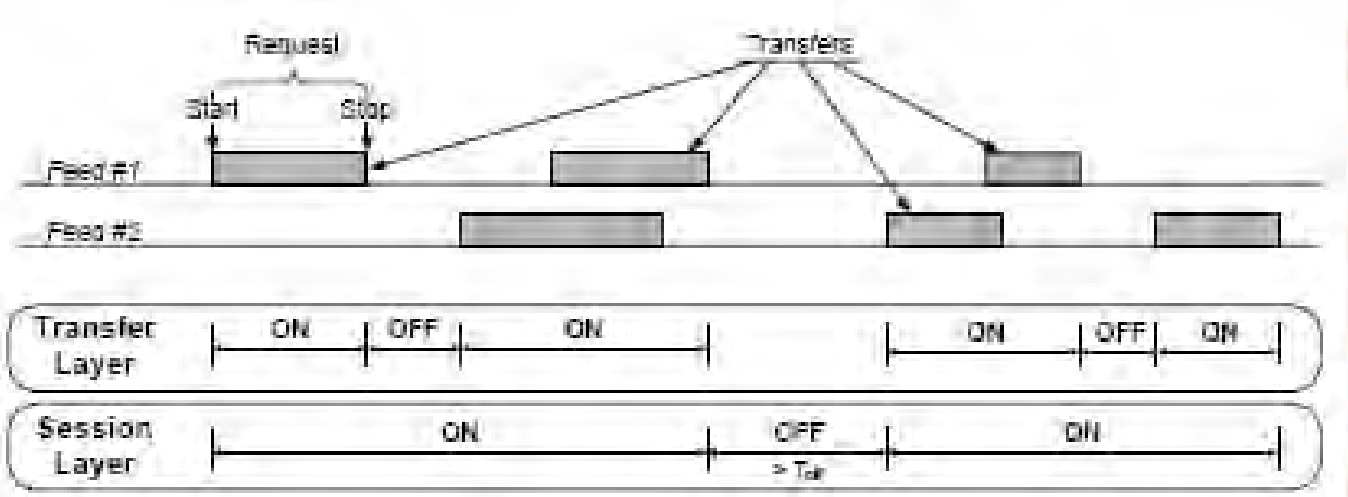

Figure 1. Relationship between client activities and ON/OFF times at the session and transfer layers

Live Streaming Workload IV

- Basic Log Statistics and Server Configuration

Log period	28 days in early 2002
Total # of live objects	2
Total # of client ASs	1,010
Total # of client IPs	364,184
Total # of users	691,889
Total # of sessions	> 1,500,000
Total # of transfers	> 3,500,000
Total content served	> 8 TeraBytes

Table 1. Basic statistics of the trace used in this paper

- Provided Information
  - Client Identification (IP address, player ID)
  - Client environment specification (OS version, CPU)
  - Requested object identification (URI of stream)
  - Transfer statistics (loss rate, average bandwidth)
  - Server load statistics (server CPU utilization)
  - Other information (referer URI, HTTP status)
  - Timestamp in seconds of when log entry was generated

Live Streaming Workload V

- Log Sanitization
  - Server Overloads
    - Slow-down user activities -> problems detecting user interarrivals
    - Turn away users -> problems detecting concurrency
  - Not in this test
    - Server utilization below 10% in 99,9% of time
    - Server load below 10% in 99,9% of time

Client Layer Characteristics I

- Characteristics
  - Level of concurrency
  - Relationship: frecuency of access / interest in one object
  - Client population in general
- Client Topological and Geographical Distribution
  - Over 1000 diferent Autonomous Internet Systems
  - Zpif-like distribution profile
- Client Concurrency Profile
  - At time t, c(t) number of active clients
  - Factors of variability
    - Diurnal effect: no interesting between 4a.m./11a.m.
    - Day of the week
    - Lag increase/decrease

Client Layer Characteristics II

- Client interarrival times
  - t(i) arrival time for i^{th} session
  - a(i)=t(i+1)-t(i) interarrival time of the i^{th} and $(i+1)^{th}$
  - i, i+1 belongs to different clients
  - Marginal distribution of a(i): Pareto
- Client arrival process
  - Process not stationary-> Periodic nature?
  - Prior works: Consistent with Poisson arrivals, but maybe just in shor times...
  - Experiment: Generate arrivals with non stationary piece-wise-stationary Poisson process... That's it!!
- Client Interest Profile
  - (Re)visit of content: Zipf- like function
  - Popularity:
    - Stored streaming: Frecuency of access by various clients
    - Live streaming: Frecuency one client access live content

Session Layer Characteristics

- Number of sessions
  - Traces not identifies delimeters
  - Have to decide Toff (3600 seconds)
- Session ON time
  - l(i): ON time for session i
  - Lognormal distribution
  - Highly variability due to fundamental property of the interaction between user and live content
- Session OFF time
  - i,j consecutive sessions belonging to the same client
  - f(i)=t(j) – t(i) – l(i): OFF time
  - Revisits to show daily, or every day...
  - Exponential distribution
- Transfers per session
  - Pareto distribution
  - Variability due to client interactions with live content
- Interarrivals of session transfers
  - Lognormal distribution

Transfer Layer Characteristics I

- Number of concurrent transfers
  - At time t, number of active transfers between server/clients
  - Very similar distribution to number of concurrent clients
- Transfer interarrivals
  - t(i): starting time for i^{th} transfer
  - a(i)=t(i+1)-t(i): interarrival time of i^{th} and $(i+a)^{th}$ transfers
  - Distribution: 2 distinct Pareto
    - Interarrivals up to 100 seconds (popular times)
    - Interarrivals larger than 100 seconds (unpopular times)
  - Not stationary
- Transfers length and Client Stickiness
  - Length of time of individual transfers
  - l(j), length for the jth transfer: Prob[l(j)>x] -> lognormal distribution
  - Variability: Stored streaming: object size characteristics
    Live streaming: Willingness to ‘stick’ to a transfer

Transfer Layer Characteristics II

- Number of concurrent transfers
  - Periodic Variability
  - Two modes:
    - Client-bound
    - Congestion-Bound

Representativeness of findings I

- Findings are unique to the workload or representative?
- Second live streaming server: News and sport radio station
  - 28.558 requests
  - 12.867 clients
  - 2 weeks period
- Similar Findings (next table)
- Differences in interarrivals due to the nature of interactions between clients and the two kinds of objects.

Representativeness of findings II

	Live Reality Show		Live News & Sports	
Workload Variable	Distribution	Parametrization †	Distribution	Parametrization †
Client Interest (transfers)	Zipf	$\alpha = 0.719$, $\beta = 0.006$	Zipf	$\alpha = 0.609$, $\beta = 0.011$
Client Interest (sessions)	Zipf	$\alpha = 0.470$, $\beta = 0.001$	Zipf	$\alpha = 0.504$, $\beta = 0.005$
Number of Active Clients	Exponential	$\lambda = 0.0019$	Exponential	$\lambda = 0.0463$
Client Interarrival Times	Pareto	$a = 2.520$, $b = 1.550$	Lognormal	μ=3.59, σ=1.52
Number of Transfers per Session	Pareto	$a = 1.43$, $b = 0.62$	Pareto	$a = 1.68$, $b = 0.39$
Session ON Time	Lognormal	μ=5.19, σ=1.44	Lognormal	μ=5.74, σ=2.01
Session OFF Time	Exponential	λ = 5.025e-06	Exponential	λ = 6.008e-06
Session Transfer Interarrival Times	Lognormal	μ=4.93, σ=1.26	Exponential	$\lambda = 0.00114$
Number of Concurrent Transfers	Exponential	$\lambda = 0.0029$	Exponential	$\lambda = 0.0496$
Transfer Length	Lognormal	μ=4.29, σ=1.28	Lognormal	μ=5.08, σ=2.03
Transfer Interarrival Times	Pareto	$a = 2.54$, $b = 0.989$	Lognormal	μ=3.09, σ=1.43

† The exponential distribution is of the form $\lambda e^{-\lambda x}$. The Zipf distribution is of the form $\frac{\beta}{x^{\alpha}}$. The Pareto distribution is of the form $\frac{ab^a}{x^{a+1}}$. The lognormal distribution is of the form $\frac{1}{\sqrt{2\pi}\sigma x} e^{-(\log(x)-\mu)^2/2\sigma^2}$.

Table 2 Summary of the distributional characteristics of the "Reality Show" and "News & Sports" live streams.

Synthesis of live media workloads I

- A generative model for live Media Workloads
  - Which variables are going to be used? -> Generative Model
- Generative Model
  - Client Arrivals
    - When: Non-stationary Poisson process
    - Which: Associated with a given arrival: Session frecuency interes profile
  - Session Length
    - How many transfers within a session?: Marginal distribution of number of transfers per session
  - Transfers
    - When starts? Distribution of the interarrival time of intra-session transfers
    - How long? Distribution of transfers length

Synthesis of live media workloads II

Variable	Distribution	Parameters / Settings
Mean Client Arrival Rate $f(t)$	Periodic over p	$p = 24$ hours
Client Arrival Process	Piece-wise-stationary Poisson	$\lambda = f(t)$
Client Interest Profile	Zipf	$\alpha = 0.470$, $\beta = 0.001$
Transfers per Session	Pareto	$a = 1.43$, $b = 0.62$
Interarrival of Session Transfers	Lognormal	$\mu = 4.93$, $\sigma = 1.26$
Transfer Length	Lognormal	$\mu = 4.29$, $\sigma = 1.28$

Summary of the variables retained for the synthesis of live streaming media workloads in GISMO

- There are diferences (periodicity) between Reality show overload and soccer program, but can be easily adjusted

Synthesis of live media workloads III

- GISMO: Generator of Internet Streaming Media Objects and Workloads
  - What is a GISMO workload?
    - Set of objects (with popularity distribution, size distribution...)
    - Sequence of user sessions
- Need to extend GISMO for live media workloads
  - Add non-stationary arrivals (reflecting diurnal effect)
  - Frecuency of access: allow the association of sessions to clients to follow a particular distribution (Zipf-like)

Summary and Conclusion

- Presented the fist characterization of live streaming media delivery on the internet
- 3 layers: clients, sessions and transfers
  - Client layer
    - Arrival: Piece-wise stationary Poisson process
    - Identity: Zipf-like distribution
  - Session layer
    - ON-time: lognormal distribution
    - OFF-time: exponential distribution
    - Number of transfers within a session: Pareto distribution
  - Transfer layer:
    - Arrival: Similar to client arrival
    - Length: lognormal distribution (session ON time distribution)
    - Bandwith: Determined by client connection speeds. 10% of transfers limited by network resources

Xabier Nicuesa Chacón

A Hierarchical Characterization of a Live Streaming Media Workload

by **Eveline Veloso, Virgílio Almeida, Wagner Meira, Azer Bestavros, Shudong Jin**

Program:

Tecnologías para la gestión distribuida de la información

Course

Servicios web y distribución de contenidos

May 3th, 2007