Hierarchical Characterization of
Live Streaming Media

Azer Bestavros

Joint work with

Eveline Veloso, Virgilio Almeida, Wagner Meira, and Shudong Jin

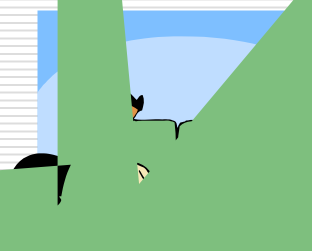
Computer Science Department
Federal University of Minas Gerais
Brazil

Computer Science Department
Boston University
USA

ACM Internet Measurement Workshop, Marseille, France
November 7, 2002

Live versus Stored Content

- How different are workloads resulting from clicking with a mouse versus surfing with a remote control?



n stored content access
, but none on live content!

Live versus Stored Content

- Live → Streaming (not vice versa)
  - Access to stored streaming media (e.g. movie clips, music, etc.) is not access to "live" content
  - Periodic rebroadcast of content (e.g. pay-per-view) is not access to "live" content
- Value of live content is in its spontaneity
  - Watching Brazil beat England "live" is intrinsically different from watching it on tape
- Internet as live content delivery device
  - Enables bypassing of editorial controls (e.g., user chooses which feed to watch)

Primary Workload Considered

- Live Reality Show Workload from one of the top content providers in Brazil
- 24x7 live content complements a one hr/wk reality TV show (a la “big brother” in US)
- Web site offers users two live objects, each is a feed from one of 48 cameras mounted around a “house” where contestants live
- Content served over unicast with server adjusting rate to match client bandwidth

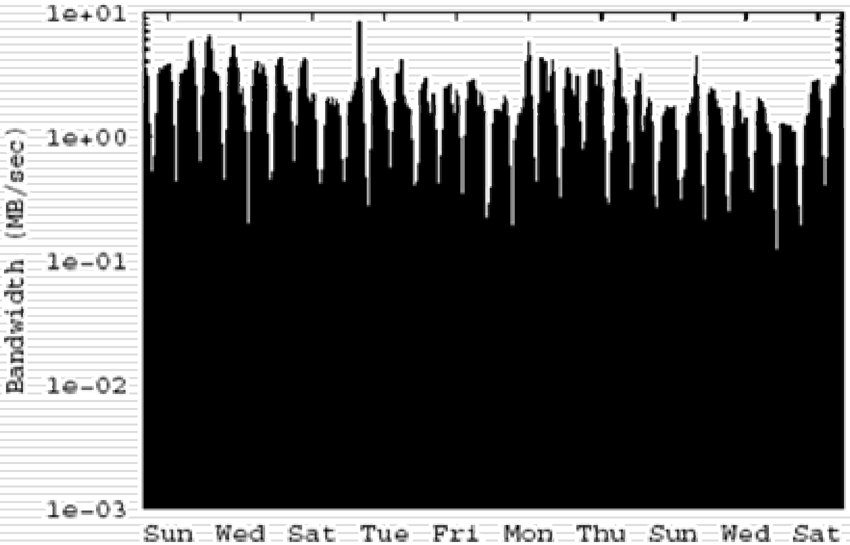
Basic Statistics

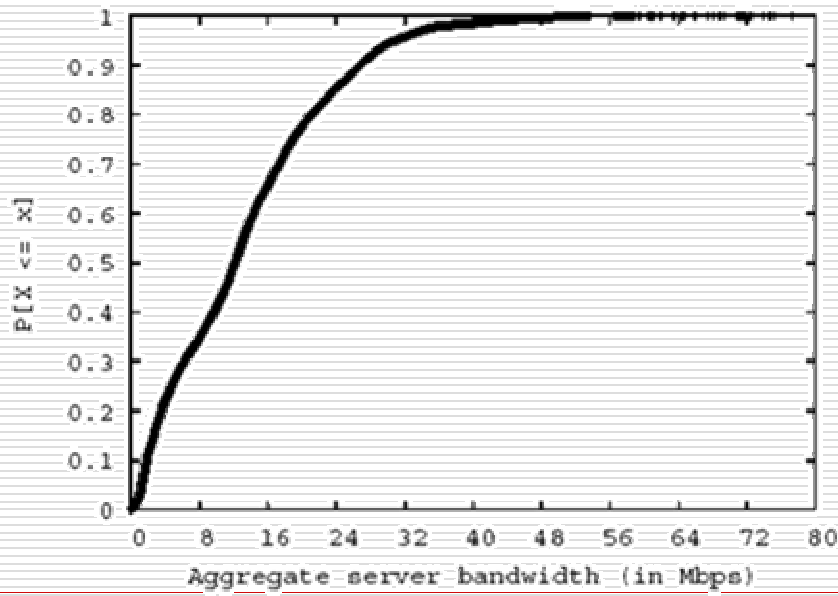
Log period	28 days in early 2002
Total # of live objects	2
Total # of client ASs	1, 010
Total # of client IPs	364, 184
Total # of users	691, 889
Total # of sessions	> 1, 500, 000
Total # of transfers	> 3, 500, 000
Total content served	> 8 TeraBytes

Note: Ratio of client:IP addresses is 2:1

Basic Statistics

- Peak 1-minute aggregate B/W ~ 80Mbps
- Server network/CPU not an issue—this is important to ensure characterization is not impeded by lack of server resources

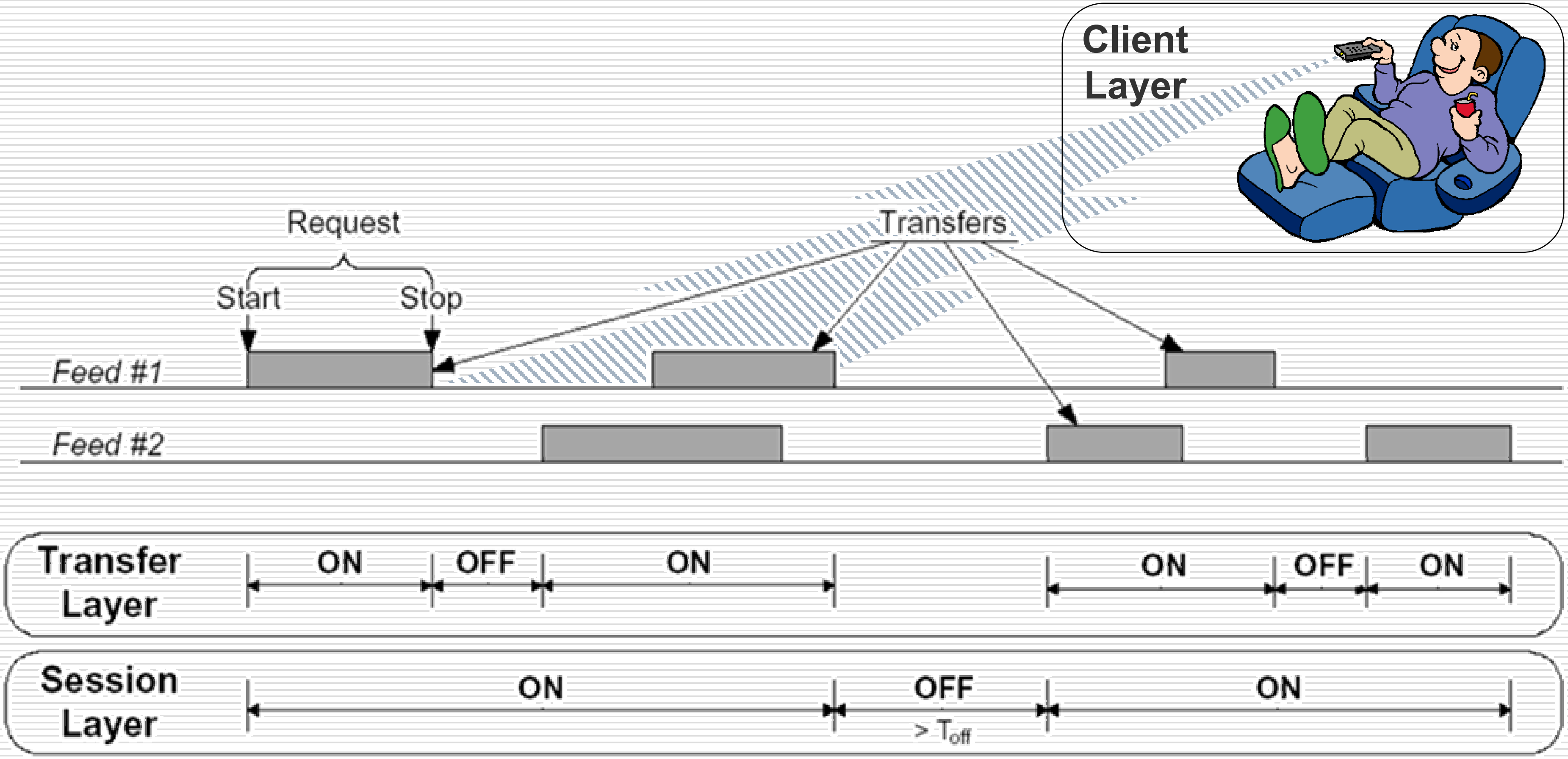




What did we log?

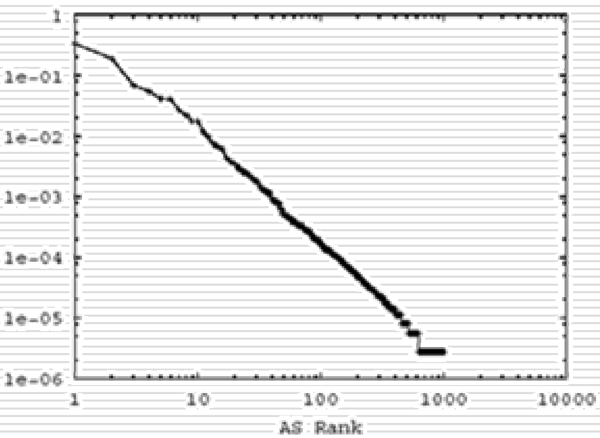
- Client Info:
  - ID, IP address, DNS, CPU, OS, language, …
- Access Info:
  - Object URI, start and stop times, codex, …
- Transfer Stats:
  - Packets sent, received, recovered, …
- Server Info:
  - CPU load, # of sessions, configuration, …

Characterization Hierarchy

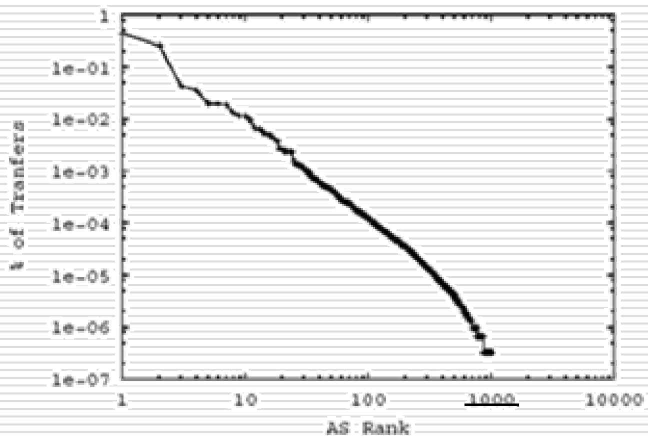


Client Layer: Basic Statistics

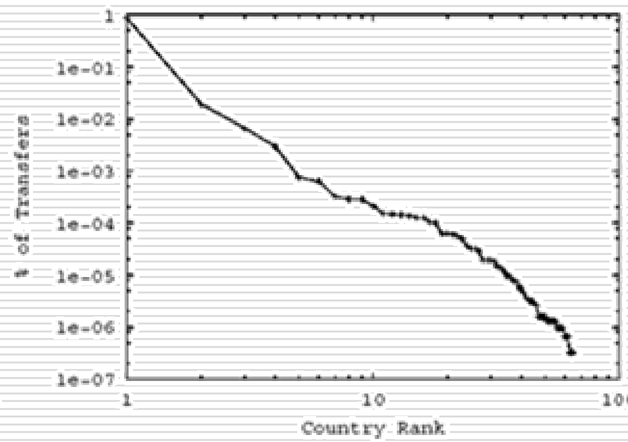
- Very diverse client population—spanning 65 countries and over 1,000 AS'es



IP Address / AS



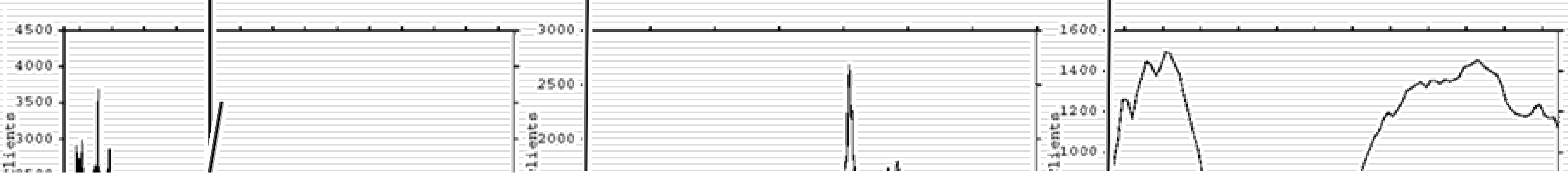
Transfers / AS



Transfers / Country

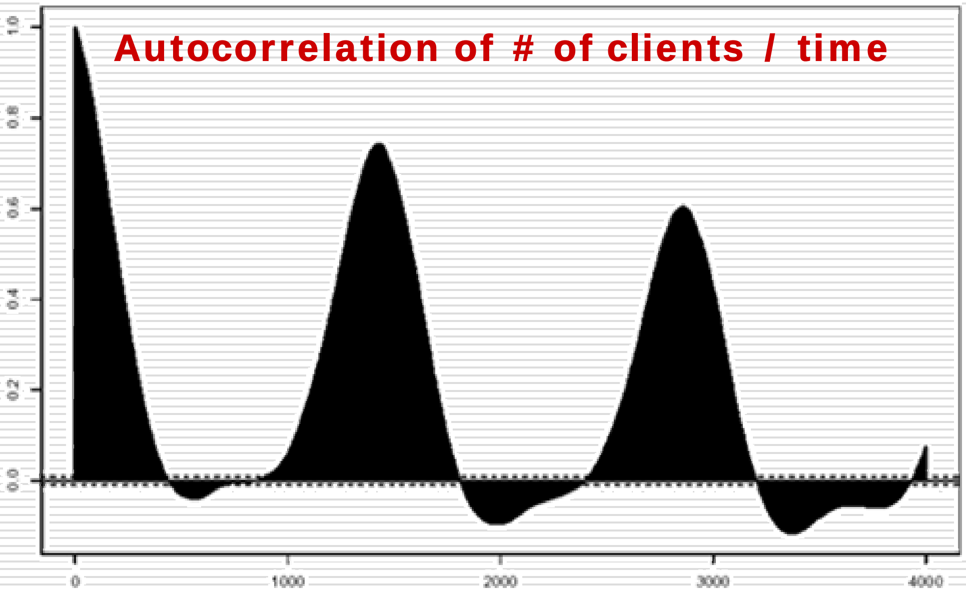
Client Layer: Concurrency Profile

- Clear periodic patterns (diurnal/weekly)
- Marginal distribution fits an exponential



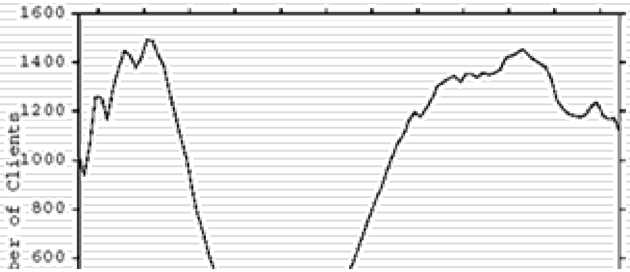
Live content "synchronizes" clients

- Diversity in client time zones does not translate to smoother diurnal patterns—it's when the contestants (not the client) "sleep" that matters



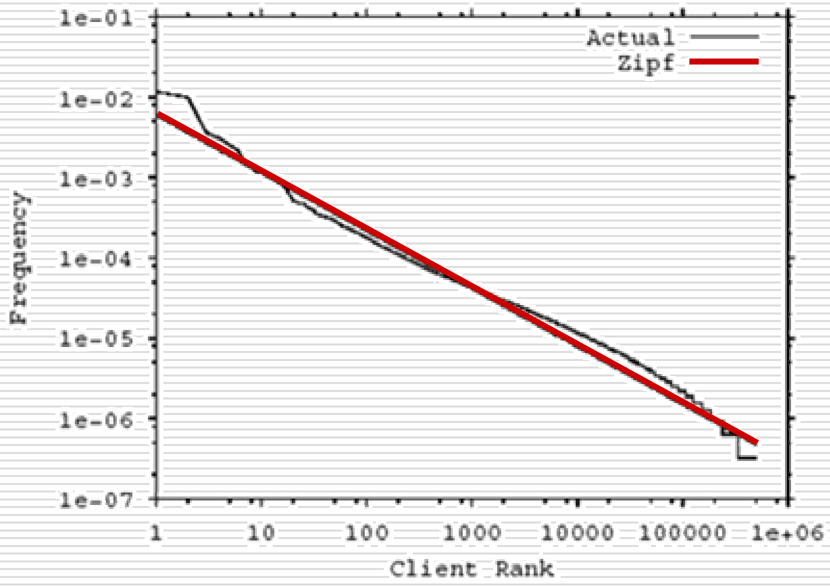
Client Layer: Arrival Process

- Generative arrival process at time *t* is Poisson with a periodic (diurnal) $\lambda(t)$
  - Resulting marginal IAT ~ Pareto
  - Good fit when $\lambda(t)$ is piece-wise constant over period < one hour



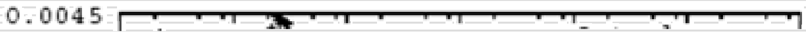
Client Layer: Interest Profile

- Number of client requests is a measure of client's interest in live content
- Interest vs client rank is Zipf; it's the dual of object popularity



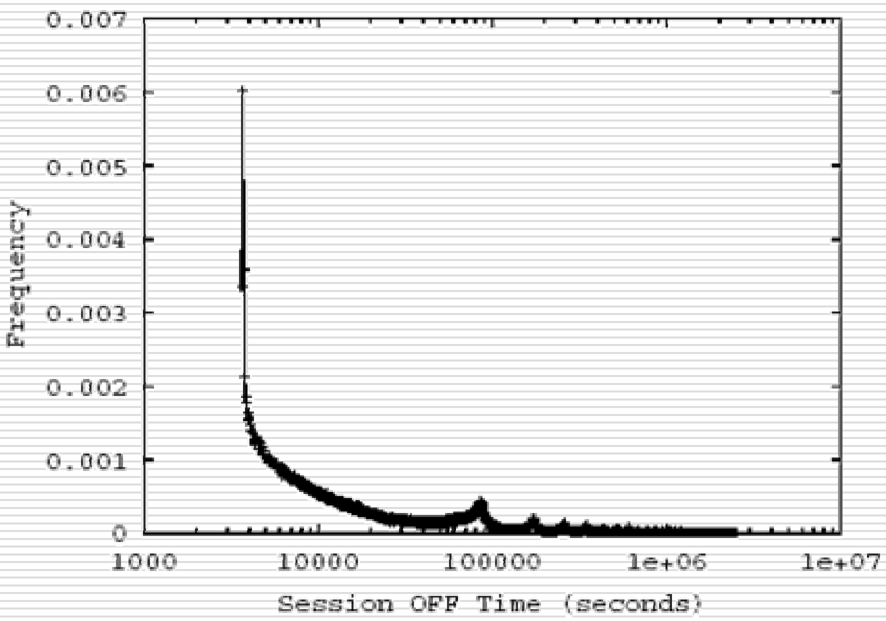
Session Layer: ON Time

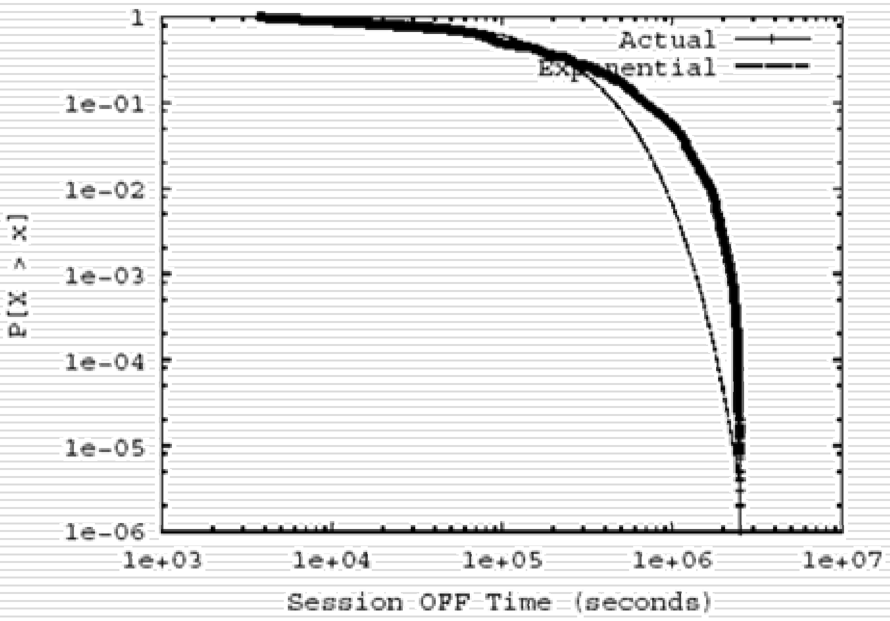
- Session ON Time is lognormal
- Relatively insensitive to diurnal patterns



Session Layer: OFF Time

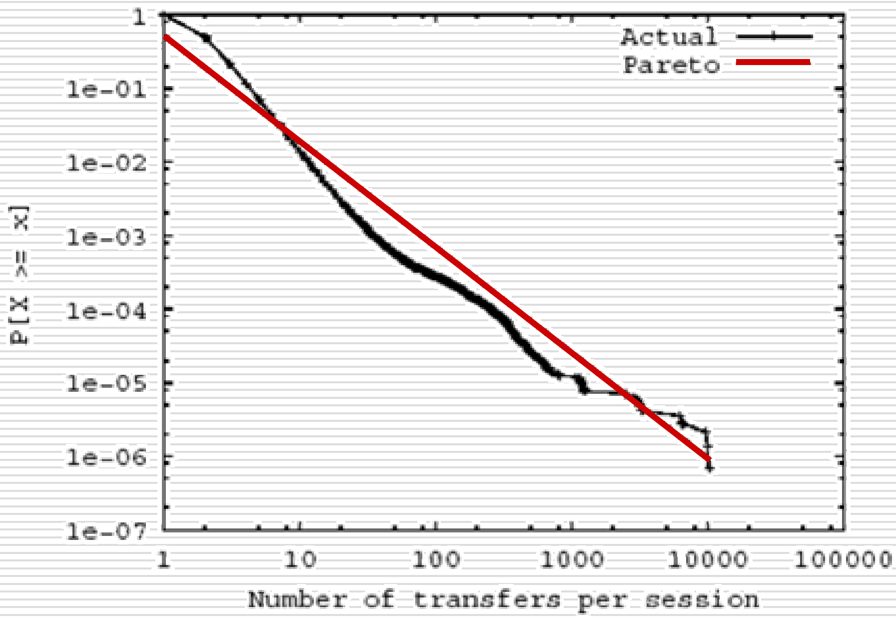
- Session OFF Time fits an exponential
- Ripples around 24-hour increments





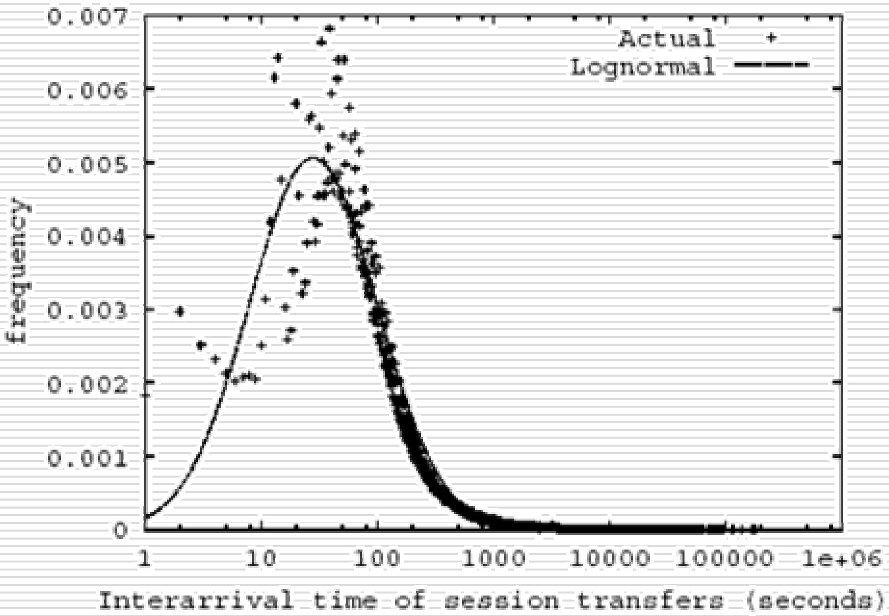
Session Layer: # of Requests

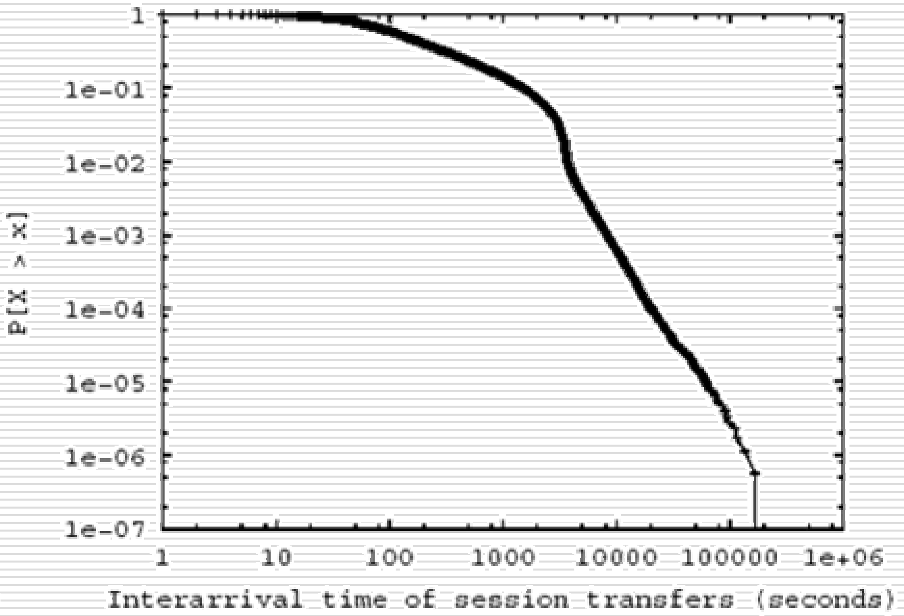
- # of requests (transfers) per session is a measure of client interactivity
- # of requests per session is Pareto



Session Layer: Request IAT

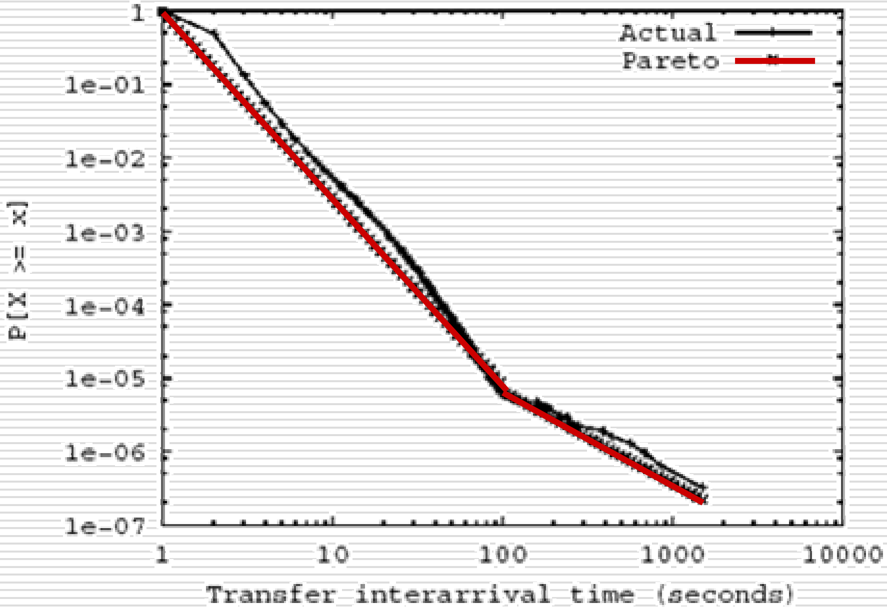
- Inter-arrival time of requests within a session is best fitted to a lognormal





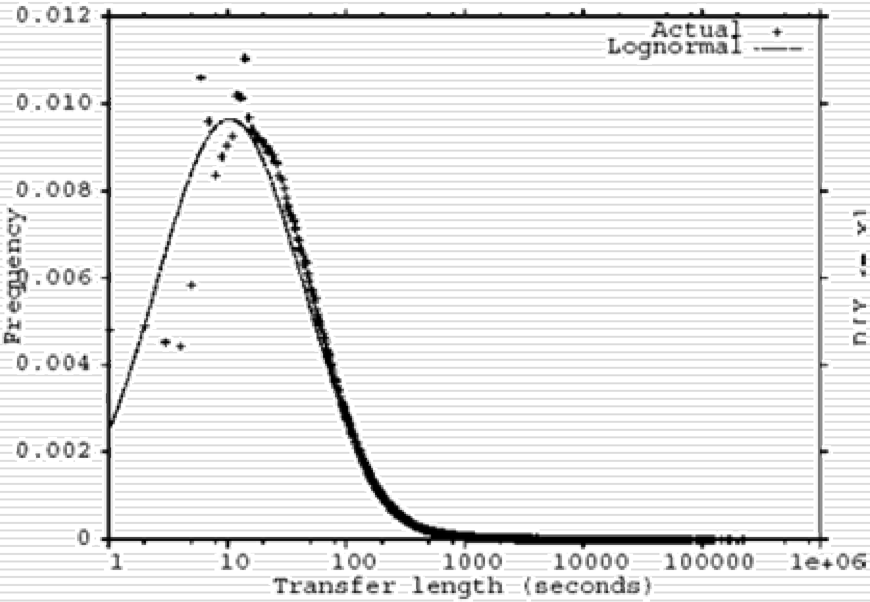
Transfer Layer: Transfer IAT

- Marginal distribution of transfer IAT is Pareto with two regimes:
  - A lighter tail (IAT < 100 sec) for popular periods
  - A heavier tail (IAT > 100 sec) for unpopular periods



Transfer Layer: Transfer Length

- Transfer length is lognormal

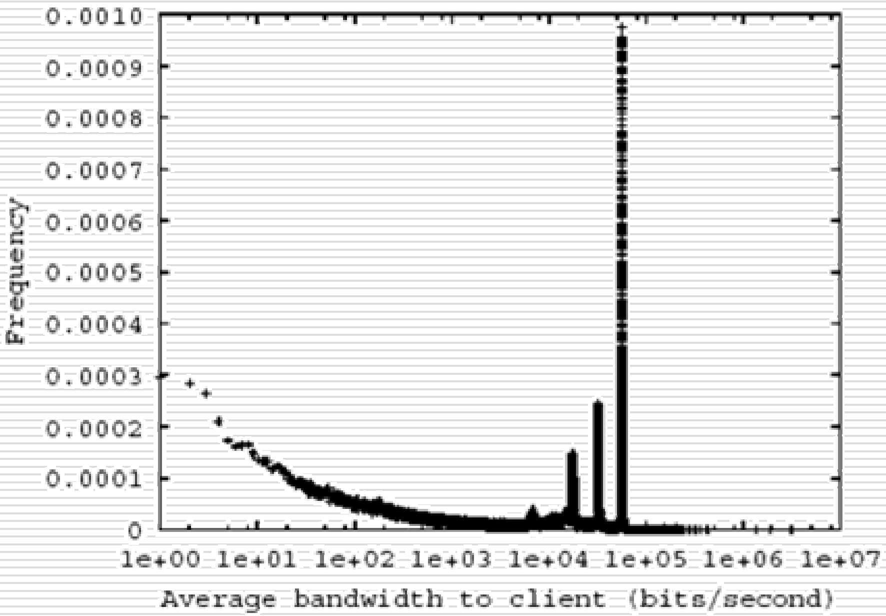


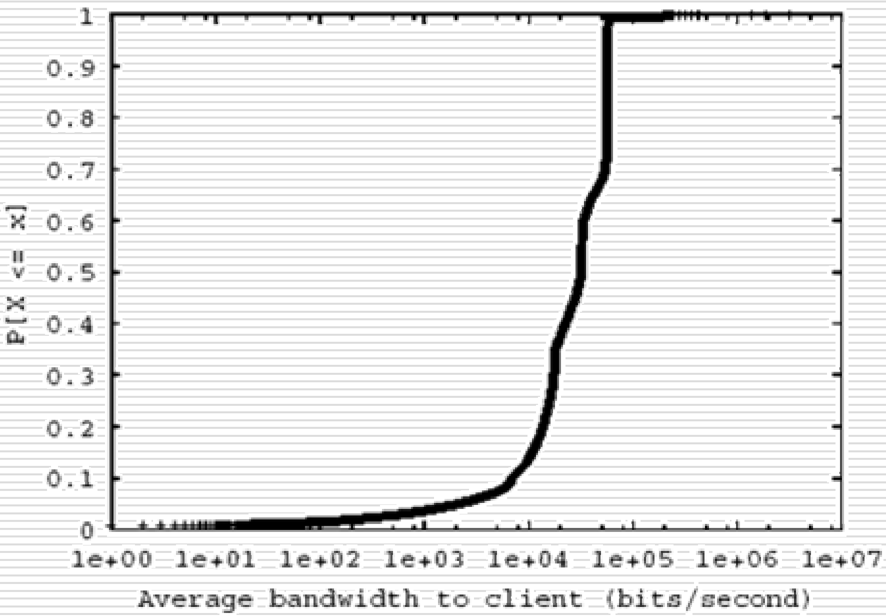
Transfer Layer: Transfer Length

- Heavy tail is not due to “file size” but rather to client “stickiness”
  - For stored content, transfer length is due to both object size and client behavior
  - For live content, transfer length is due *only* to client behavior

Transfer Layer: Transfer B/W

- Transfer bandwidth bound by:
  - Client connection speeds (> 90%)
  - Network congestion (< 10%)





Are these “typical” characteristics?

- Characteristics are likely to depend on nature of live content
- Need to characterize many live content workloads to answer this question
- We performed the same hierarchical characterization on a second “Live News and Sports” streaming workload
  - 30K requests to 12K clients over two-week period from same content provider
  - Live content is very different in nature from our main “Reality Show” workload

Are these "typical" characteristics?

- Surprisingly, workloads have similar characteristics (modulo parameters)!

	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	$\lambda = 5.025e{-}06$	Exponential	$\lambda = 6.008e{-}06$
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

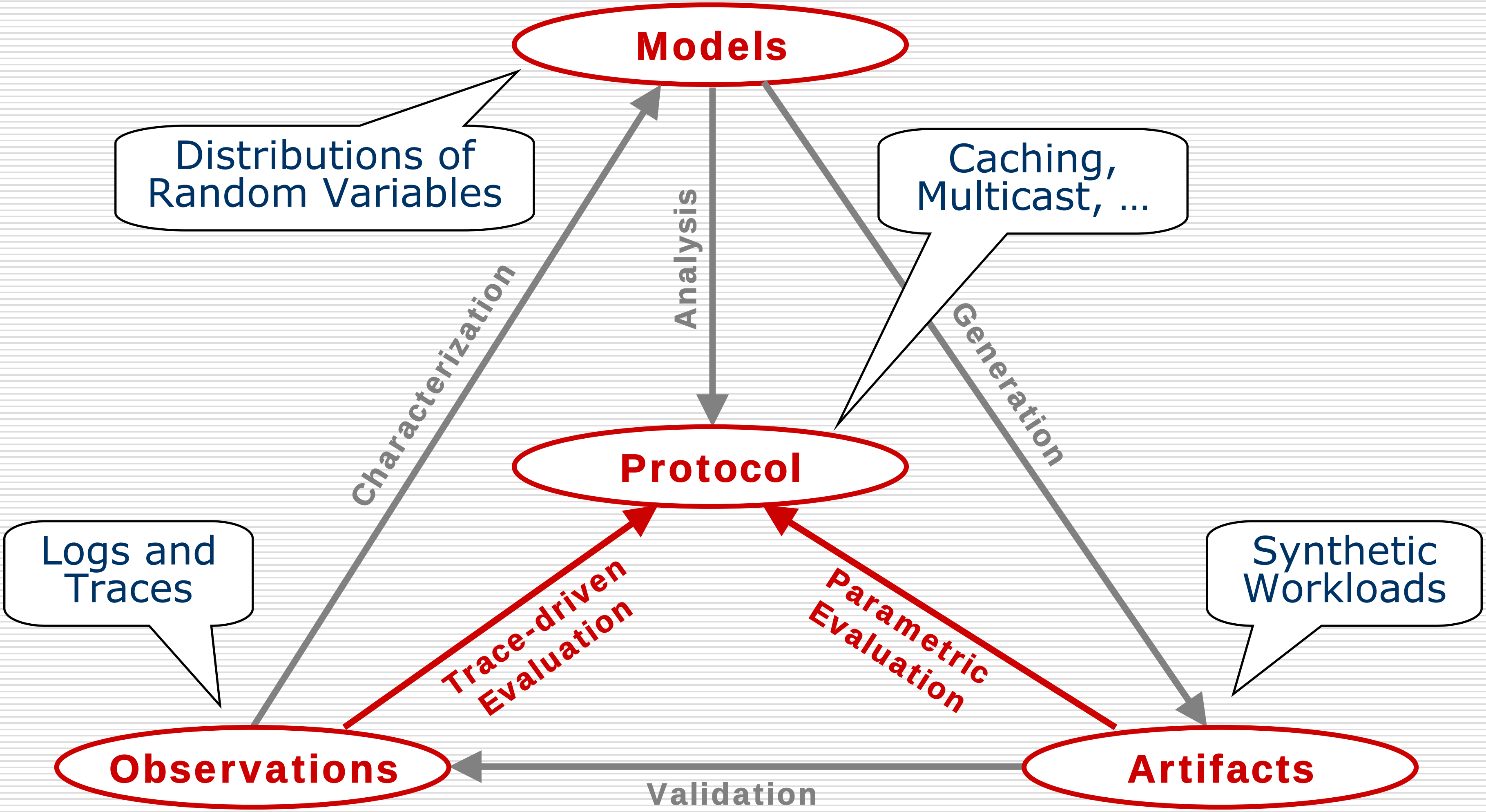
† The exponential distribution is of the form $\lambda e^{-\lambda x}$. The Zipf distribution is of the form $\frac{\beta}{x^\alpha}$. The Pareto distribution is of the form $\frac{ab^a}{x^{a+1}}$. The lognormal distribution is of the form $\frac{1}{\sqrt{2\pi}\sigma x} e^{-(\log(x)-\mu)^2/2\sigma^2}$.

What does the comparison suggest?

- Arrival processes for live content are dependent on nature of live objects:
  - Scheduled (e.g. game) vs continuous (e.g. reality TV)
  - Periodic (e.g. news) vs aperiodic (e.g. event)
  - Number of channels (e.g. one vs many feeds)

- Client “interest” and “stickiness” seem to be less dependent on nature of live content

Measure → Model → Synthesize

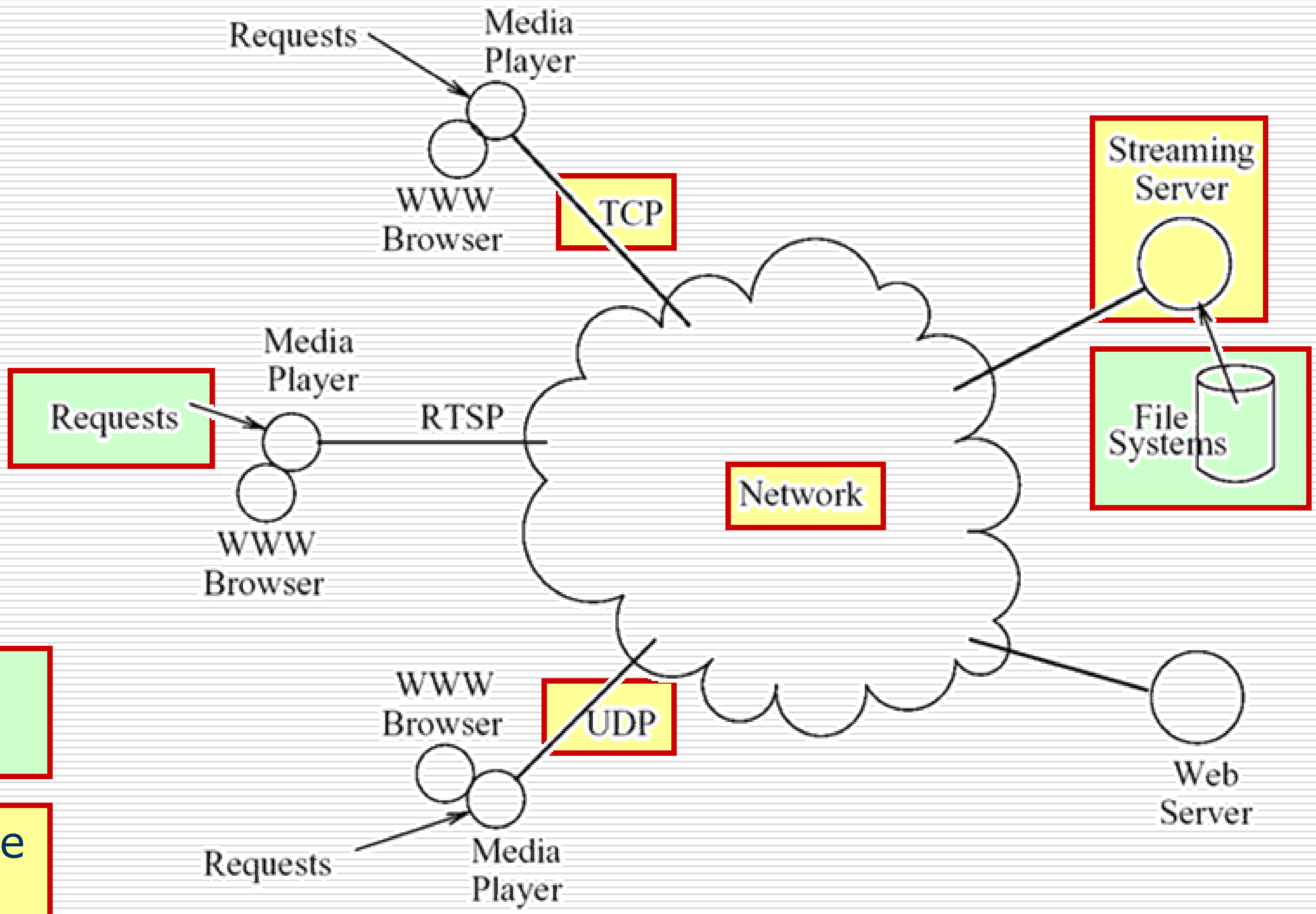


GISMO Workload Generator

- GISMO: A toolkit to generate synthetic streamed media workloads [JinBestavros:PER'02]
- GISMO generates
  - A set of "placeholder" streaming media objects, which can be installed on servers
  - Requests to these objects, initiated by clients subject to a prescribed access model

http://csr.bu.edu/gismo

GISMO: Components



Standard Elements

Programmable Agents

GISMO: Request Generation

1. Generate session arrivals using piecewise-stationary Poisson
2. Map clients to sessions to match client interest profile
3. Determine number of transfers per session
4. Generate transfer arrivals within a session based on distribution of IAT of intra-session transfers
5. Determine each transfer length

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	$\mu = 4.29$, $\sigma = 1.28$

Take-Home Messages

- With live content, there is a role reversal of clients and contents
  - Content (rather than clients) determine periodic (e.g. diurnal) behaviors
  - Clients (rather than content) determine length of transfers
- With live content, there are new dimensions to modeling user access
  - Client interest and stickiness

Hierarchical Characterization of
Live Streaming Media

Azer Bestavros

Joint work with

Eveline Veloso, Virgilio Almeida, Wagner Meira, and Shudong Jin

Computer Science Department
Federal University of Minas Gerais
Brazil

Computer Science Department
Boston University
USA

ACM Internet Measurement Workshop, Marseille, France
November 7, 2002