Measuring online live video streaming

Assigned: Jan 12, 2009. Due: Feb 2, 2009

Overview: The goal of this first assignment is to get a hands-on experience with understanding and measuring a real online live video streaming event. We are going to use the US presidential inauguration on Jan 20, 2009 for this project.

This event is going to be streamed live on multiple websites (youtube, cnn, cbs, nbc) using several different video streaming technologies (e.g., different video formats: flash, wmv, silverlight; different networking technology: unicast, p2p multicast, CDN). Your objective in this assignment is to pick one (or more) delivery modes and do a quantitative measurement study to analyze this event. (Yes, we are asking you to watch online video for class credit :-)).

Experiment Design:

- Pick one (or more) delivery platforms (e.g., youtube on a wireless network, nbc on a home dsl connection)

- Install packet collection software (e.g., tcpdump, ethereal, wireshark, your favorite thing) on the machine you will be running the video on.

- Collect the packet trace during the live video event

- Analyze the packet trace and report different diagnostic metrics of interest. For example, what is the average packet size, what is the stream download rate, what is the average inter-packet delay, jitter between packets, "video quality", how much total bytes are being downloaded/uplodaded, how many active connections are being maintained, where is the source of the video (local/remote/cdn).

Extra credits:

- pick interesting delivery technologies (p2p? iphone?, wireless 3g?). (the wackier, the better :-))

- contrast 2 or more delivery technologies

Report:

- Summarize your findings in a brief 2-3 page report

- Describe your setup (what is your machine configuration, what video delivery method did you study, what is your network configuration, briefly describe if there were any non-trivial steps in instrumenting the collection infrastructure).

- Read the assigned measurement studies for Jan 16 to get a sense of what are the typical metrics people are interested in, what sort of analysis methodology they use, how do they summarize and present the results etc. Present quantitative results using the metrics defined above as rough guidelines.

- Some sample questions to look beyond just the measurement: how do you think the delivery can be improved. are there obvious sources of inefficiency? is there too much overhead? how much of a cpu/memory resource hog is the technology? are there some biases in your measurement methodology?

Some rough guidelines:

- start early!

- test out your collection infrastructure using some test streams/test sites etc. e.g., make sure the stream actually works, make sure you can actually collect/analyze packets

- can your machine run tcpdump with a high speed stream without dropping packets?

- run tcpdump with -nn to avoid dns lookups etc while collecting data

- keep the collected packet traces so you can run more interesting analysis/metrics on them later.

- some useful tools to understand traces: nslookup, whois, traceroute