Project Specification Document

Date of Lecture	
Link to Notes	

Project Specification Document

Collaborators

Kuhu Halder (kh761)

Katherine Lee (ksl103)

Goal

The goal of this project is to study Internet benchmarks of existing document sharing applications. Specifically, we are interested in the performance differences for the user who is generating the content (the user who is typing) and the user who is viewing the content.

In order to achieve this goal, we are asking the following questions:

- What is the network utilization of some popular document sharing applications?
- How do performance metrics (latency, etc.) vary with different link capacities?
- How do these applications respond to connectivity disruptions?
- How does competing traffic affect performance?
- Does different usage modality (multiple viewers, active vs passive viewership) have an effect on performance?

The applications for which we will be generating these benchmarks are:

Document Sharing:

- Google Docs
- Notion

- Microsoft Docs

Existing Work

There is some existing work in the area of measuring internet benchmarks for real-time activities.

- Michel, et al. Enabling Passive Measurement of Zoom Performance in Production Networks: https://www.cs.princeton.edu/~jrex/papers/zoom22.pdf
- MacMillan, et al. Measuring the Performance and Network Utilization of Popular Video Conferencing Applications: http://internetequity.uchicago.edu/wp-content/uploads/2021/05/Measuring-the-performance-and-network-utilization-VCA-1.pdf
- Zhang & Liu, On Crowdsourced Interactive Live Streaming: A Twitch.Tv - Based Measurement: https://arxiv.org/pdf/1502.04666.pdf

There are also some existing programs and tools that measure *some* of the benchmarks that we plan to study.

- TwitchTest: https://r1ch.net/projects/twitchtest
- Twitch Inspector: https://inspector.twitch.tv/#/
- Google Lighthouse: https://developer.chrome.com/docs/lighthouse/overview/
- Google Web Vitals: https://chrome.google.com/webstore/detail/web-vitals/ahfhijdlegdabablpippeagghigmibma?hl=en

Description

This is a study of internet measurements for document sharing applications. To our knowledge, there are no existing studies, that study these particular applications. There are, however, studies which measure and compare performance and network utilization of other types of applications, such as live streaming applications or video conferencing applications. We suspect this is because the benchmark differences for measurements like latency are much less pronounced than with other applications like video conferencing, however, the these performance pitfalls are still existent.

Approach

The approach for this project is to simulate different scenarios under which popular document sharing applications might be used and take measurements of various internet benchmarks.

The scenarios that we plan to study are:

- Where only a single user is actively participating in some activity, all other users are passively viewing
- Where multiple users are actively participating in some activity
- Where network connectivity is disrupted
- Where the application must compete with other traffic

In order to automate the environment, we plan to use the Python PyAutoGUI package which allows developers to automate mouse movement, keystrokes, etc. In order to generate measurements, we plan to use Clockwork and the WebRTC Statistics API. Clockwork is able to provide time stamping and clock syncs to make accurate latency measurements. The WebRTC Statistics API will allow us to generate measurements (such as uplink and downstream throughput) for third-party Internet applications. Google also provides the PageSpeed Insights API which provides various performance measures of web pages.

Tentatively, some of the measures we plan to study are:

- **Latency:** Latency will be measured through tracking timestamps for individual activities on both the user side and the viewer side.
- **Throughput:** WebRTC Statistics API
    - Uplink
    - Downstream throughput
- **Network Utilization:** WebRTC Statistics API
- **Time to Recovery (TTR):** Timestamps

Evaluation

We will be evaluating our study based on whether or not we are able to provide clear and concise measurements for each of the testing scenarios that we plan to observe. This is similar to how results are evaluated in MacMillan et al.'s “Measuring the Performance and Network Utilization of Video Conferencing Applications.” In their study of Zoom, Google Meet, and Microsoft Teams, the individual metrics and performance are measured relative to the specific scenarios posed. There are some cases in the MacMillan et al. study, particularly with Zoom, where they are not able to provide accurate measurements for all the applications they are studying because they are working with third-party applications. While successful project execution will ultimately compare performance between the two applications we plan to study, should there be measurements that we are not able to gather because we are working with third-party applications we will, to the best of our ability, provide reasoning and hypotheses for their performance.

Risks

Because we are looking to measure performance of third-party in-browser applications, we are unsure how many Internet performance benchmarks we will actually be able to analyze. The WebRTC Statistics API has been used in similar studies for video conferencing applications, however, generally is used exclusively for audio and video content. However, we are hopeful that the API will also be able to provide meaningful statistics to our study.

We will need to set up an environment that is able to simulate multiple users in order to measure different situations, for example, situations where only one user is participating in an activity and where all other users are viewers, situations where all users are actively participating, and situations where connectivity is disrupted. Simulating the environment as well as the various scenarios pose potential risks to accurate results. Simulating the environment will require many physical or virtual machines.

There is a risk with looking at these internet measurements with regard to document sharing applications that the differences between users will be so minute so as to not be significant, and while those results would be significant on their own, it is not necessarily conducive to an interesting paper. Should we find the results unsatisfying early, we may (with instructor approval) switch over to taking these measurements with regard to live streaming applications (Twitch, Youtube Gaming, etc.).

Brainstorming (1)

To-Dos (1)

Timeline (1)

Report Rough Draft

Data (custom collaboration app in react)

Third party apps

Custom App Experiments Summary

Google Docs Experiments Summary

Notion Experiments Summary

Microsoft Word Experiments Summary