Rise of Web -based Computing Trends on the Client Side
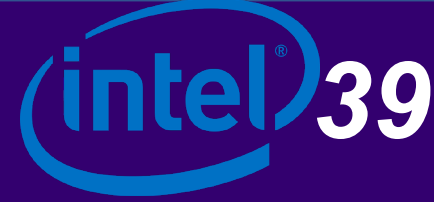
STUDENTS: Andrew Hadimaja, Ayush Singh, Furion Liu, Ishmeet Singh, Peter Lam

intel 39

INTRODUCTION / OBJECTIVES

Our team needed to analyze Intel's data based on four objectives:
- 1. Study native application usage vs. web category usage
- 2. Web persona classification and study how user behavior is different on desktop
- 3. A day in a life with desktop/web usage
- 4. Resource utilization for different categories

Progressive Web App (PWA) Analysis

- By creating snapshots of data on PWA on a predefined list, we can identify trends over time using Google Lighthouse website analyzer (350k sites analyzed)
- PWA score is based on Google Lighthouse checklist
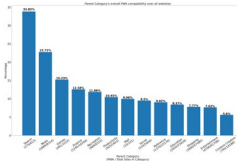
Category	Avg. PWA Score	Pass / Total	Full Pass
News	0.69	1889 / 8314	79 / 0.04
Content Creation	0.69	795 / 14188	25 / 0.03
Education	0.68	2004 / 23929	39 / 0.01
Shopping	0.72	3690 / 47480	193 / 0.05
Entertainment	0.69	4706 / 61708	142 / 0.03
Social	0.69	514 / 5409	24 / 0.04
Games	0.66	491 / 3222	23 / 0.04
Finance	0.69	1295 / 10294	22 / 0.01
Reference	0.68	1370 / 15176	22 / 0.01
Productivity	0.67	565 / 5403	12 / 0.02
Mail	0.68	33 / 331	1 / 0.03
Recreation	0.69	989 / 8333	31 / 0.03
Search	0.59	275 / 813	4 / 0.01

## Internet Protocol Usage Analysis
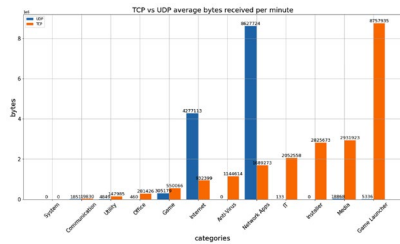
- The TCP is connection-oriented. It guarantees data is transferred intact and in order while sacrificing speed.
- The UDP is connectionless, and it is faster in rate while it doesn't guarantee reliability.

- Most categories favor TCP as it ensures the completeness of the data.
- Internet & Network Apps favors UDP as live streaming & real-time video conferencing apps are time-sensitive.

Page Fault Analysis

- Page fault(PF) is an interruption when a software program attempts to access a memory block not currently stored in the system's RAM.
- If the desired page is not present in the main memory and has to be fetched or paged in from the virtual memory, the fault is classified as a hard page fault(HPF). Space has to be allocated in the RAM in this case.
- Internet and Game Launcher processes are the main contributors to both PF per sec and HPF per sec.
- High rate of faults amongst Internet processes could be due to document object modelling and chrome.exe is the main source of HPF/sec and PF/sec amongst Internet processes.
- HPF's are highly correlated with disk read iobytes.
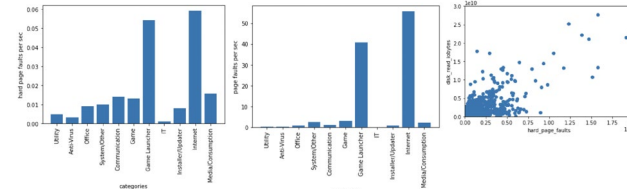
TCP / UDP Connection Usage by Number of Runs Analysis

- TCP or Transmission Control Protocol. It has a slower but more secure connection. (ex. Chat)
- UDP or User Datagram Protocol. It has a faster but less secure connection. (ex. Video chat)
- Ex. Microsoft Teams.exe is a collaborative app mainly used for video streaming and even chatting platforms.
- The most popular time for people using Microsoft teams is between 8PM-9PM (about 13k runs)
- For UDP, teams.exe have very high UDP connection usage between 8 PM-9 PM, which aligns with the time when the most number of runs happened.This means more video streaming during those times (video meetings)
- For TCP, teams.exe have a high TCP bytes meaning they use a lot of TCP connections between 8AM-9AM. Means more chat usage on those times (chat message, announcements).
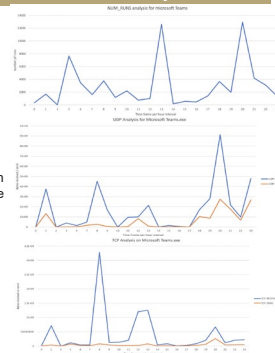
System behavior in different browsers

- Entertainment (specifically video streaming) and "other" contribute to most visits of a category for all the 3 browsers.
- Content creation and recreation had the least percentage contribution. Content Creation for Edge was 0%. Games was also a lot less than expected.
- Microsoft Edge:
  - Used primarily for social and mailing purposes.
  - Minimal usage for productivity - programming, crm, project management, web development, spreadsheets, etc.
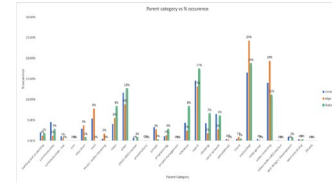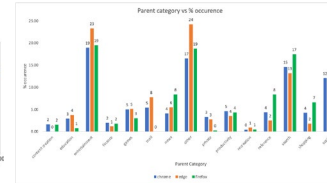
Implementation/ Methodology

- Statistical analysis:
  - Hypothesis and planning research design
  - Collecting, cleaning, processing data
  - Summarizing data with visualizations
  - Testing Hypothesis
  - Interpreting the results
- Tech stack used:
  - Python
  - Excel

Conclusion and Future work

- There are many different kinds of conclusions we can come to when looking at the data set. Depending on what we are looking at, Intel's team could use the information to optimize and address different issues that are analyzed.
- We hope that our work on the data sets provided to us could serve as the basis for the Intel company to dive deeper into the findings and conduct a similar analysis on their large-scale data sets.

ADVISORS: Bijan Arbab, Mohammad Haghighat, Sruti Sahani, Prof. Payman Arabshahi, Prof. Arindam Das

SPONSORS: INTEL

ELECTRICAL & COMPUTER ENGINEERING
UNIVERSITY of WASHINGTON