CIS 655

HOMEWORK 5:

Name: Himani Patel

SU ID: 495936092

Email ID: [htpatel@syr.edu](mailto:htpatel@syr.edu)

**Reading and Research Assignment**

1. **Analyze and compare at least two cloud computing services for machine learning. And which one would you choose and why?**

Ans. Cloud computing can be described as the delivery of the computing resources that are in demand over the internet like database storage, computer power, applications and some other resources related to IT. The service made to provide these resources to the user on demand via the internet from a cloud computing provider’s server is called cloud computing service.

The cloud computing services for machine learning are provided by the following:

Google Cloud: Google provides a collection of public cloud computing services that is Google Cloud Platform which operates on the same framework as that of other end-user products that Google uses, like YouTube and Google Search. The services that are performed by Google Cloud Platform are computing hosted services, storage and application development running on Google hardware. The access to these services is given to the software developers, cloud administrators and other IT related professionals. This is done over the network via a dedicated network connection. Creating efficient machine learning models for all kind of information is easier with this google cloud computing for machine learning. It involves bolsters reprocessing and information recovery because of google cloud utilization. Instant predictions and programmed calculation proposals are provided to the clients by this cloud computing service.

Advantages:

* It helps to control tensor procedures by providing command line interface with Google utility.
* Versatile models are built which is beneficial in preparing them for deploying tensor stream.
* Mainly, users can build their own models by referring to the pre-prepared models which supports exploring pictures, video testing, discourse acknowledgement, etc.

Disadvantages:

* The pre-prepared models are not easily available.
* The coders can establish noticeable control whereas the clients that are not specialized can face some problems.

Amazon Web Services: Amazon provides a comprehensive and expanded cloud computing platform. A mixture of Infrastructure as a Service (IaaS), Platform as a Service (PaaS) and Software as a Service (SaaS) is delivered by Amazon Web Services (AWS). Amazon cloud computing service for Machine Learning is one of the automated solutions in today’s market.

Advantages:

* Data is accepted rom multiple data sources.
* The processing steps on the data are automated.
* Also, the noticeable benefit is that the computing service also differentiates between categorical and numerical fields.

Disadvantages:

* It supports only three predictive capabilities: regression, binary classification and multiclass classification.
* There are no autonomous learning methods i.e. the methods that are not supervised and that is the reason why it is a fully automated but limited solution.

Microsoft Azure: The Microsoft Azure cloud computing service is provided by the Microsoft and it is a public cloud computing platform. The services are provided over the network by the data centers managed by Microsoft. Azure provides intuitive interface which helps the users to build their applications according to their thoughts. The operations like data exploration, preprocessing, method selection and model validation are performed using the graphical drag and drop interface provided by Azure. Data Science field is also benefited by Azure cloud computing service as it consists of algorithms for classification, regression, anomaly detection, recommendation and text analysis. Moreover, K-means clustering algorithm is also supported by Azure.

Advantages:

* Activities like information investigation and information approval can be performed physically.
* It strengthens parallel, relapse and identification calculations.
* It supports more than 100 different techniques.

Disadvantages:

* It is not useful as a learning process to the users who need a concise usage. Also, preparing part is very moderate.

I would choose Amazon cloud computing services for Machine Learning over all of the above because it is easy to implement and its functionalities and other important procedures are well documented which is very helpful to the users.

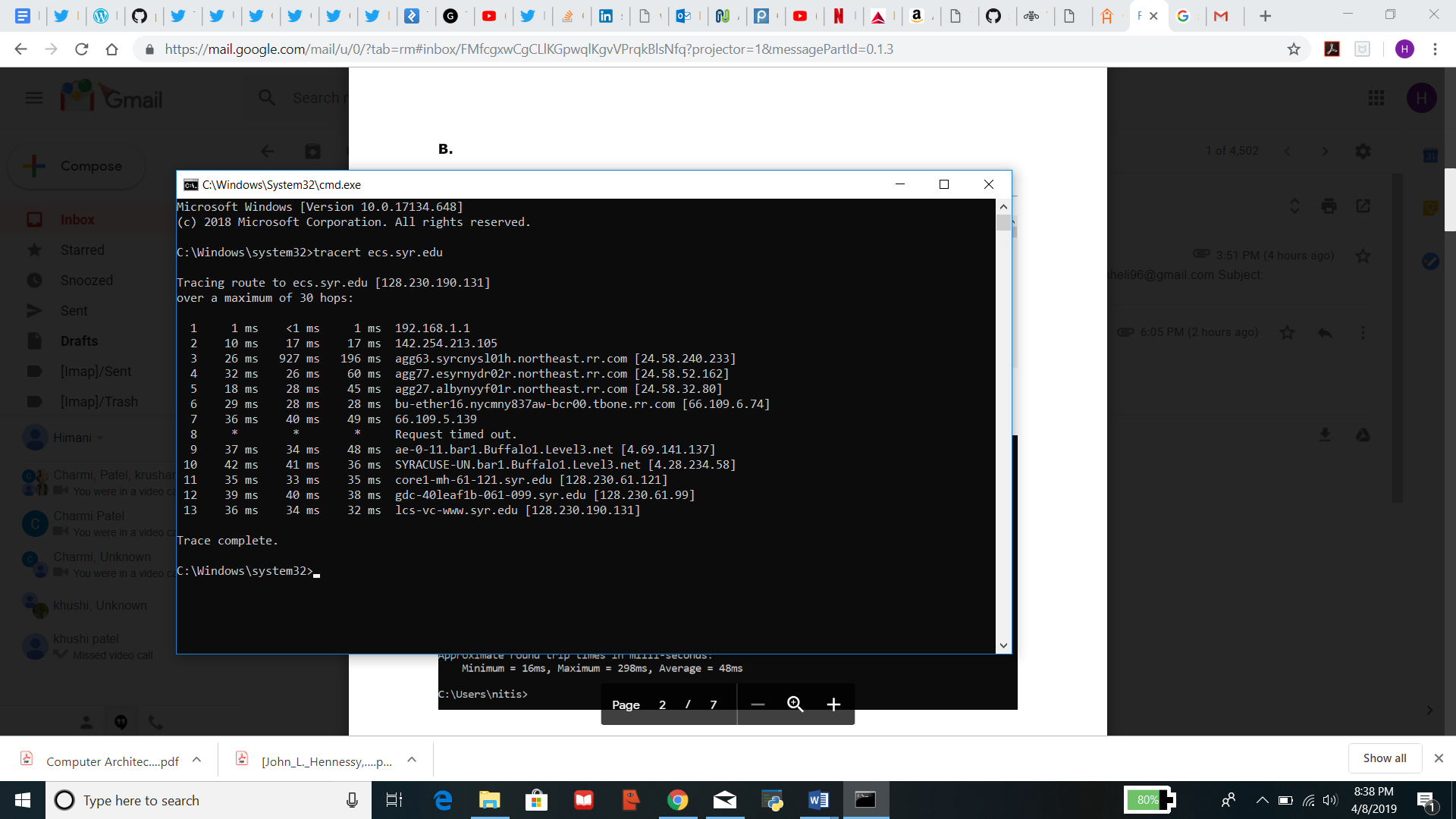
1. **The Windows ‘utility’ tracert, (or its Unix equivalent, traceroute) can be used to find the sequence of routers through which a message is routed. Use this to find the path from your site to some others. How well does the number of hops correlate with the RTT times from ping? How well does the number of hops correlate with geographical distance? Another Unix utility ‘ping’ can be used to find the RTT to various Internet hosts. Read the man page for ping, and use it to find the RTT to www.nyu.edu in New York City, and www.intel.com in California. Measure the RTT values at different times of day, and compare the results. What do you think accounts for the difference?**

Ans. We know that the utility ‘tracert’ is used to find the sequence of routers through which the message is passed. The utility ‘ping’ is used to find RTT for various hosts over the internet.

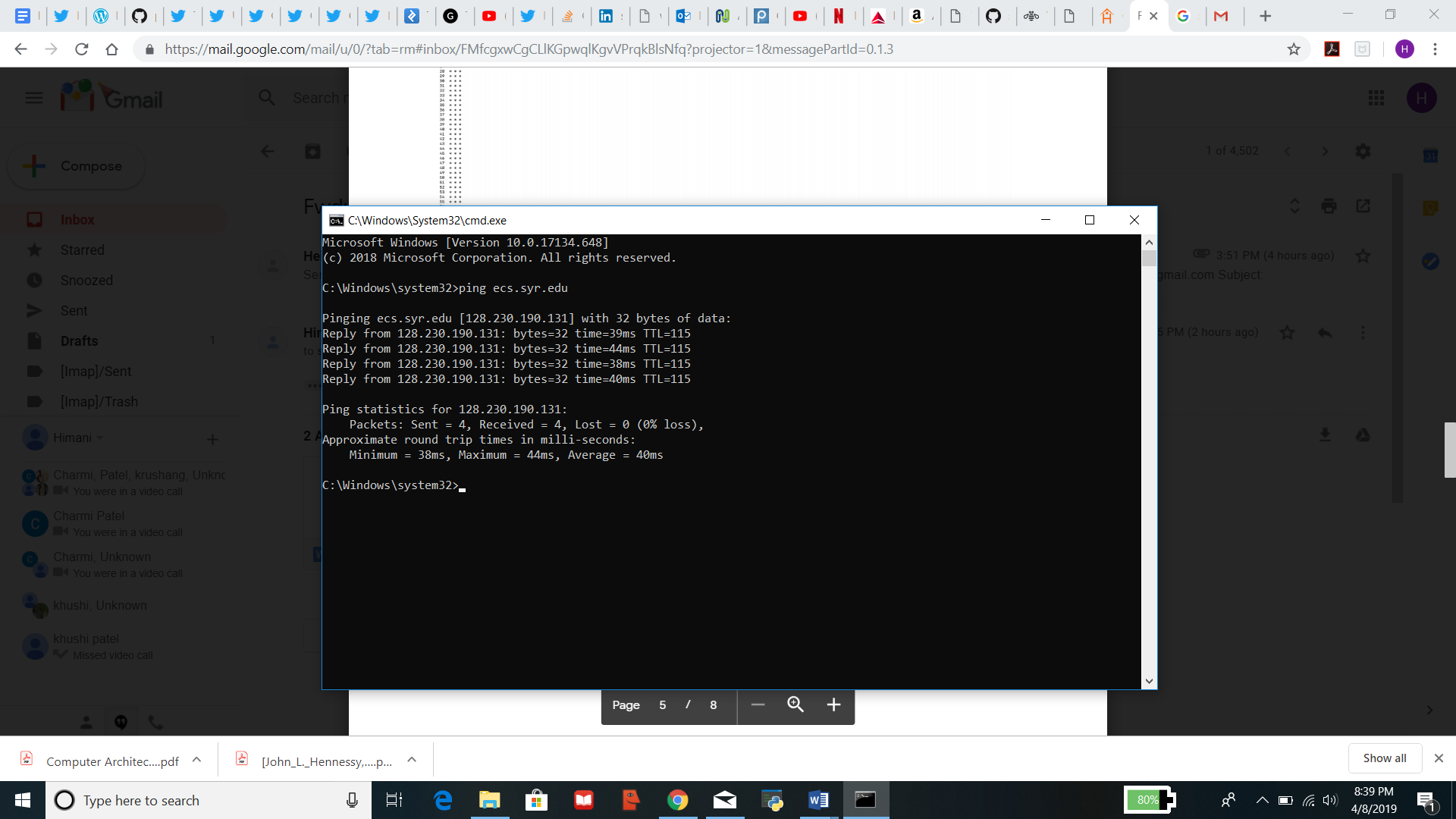
Hop count can be described as the number of routers through which the data packet passes starting from the source to the destination.

RTT i.e. Round-Trip Time can be described as the total time taken by a signal to travel from the sending point to the destination in addition with the time taken to receive the acknowledgement of the signal received.

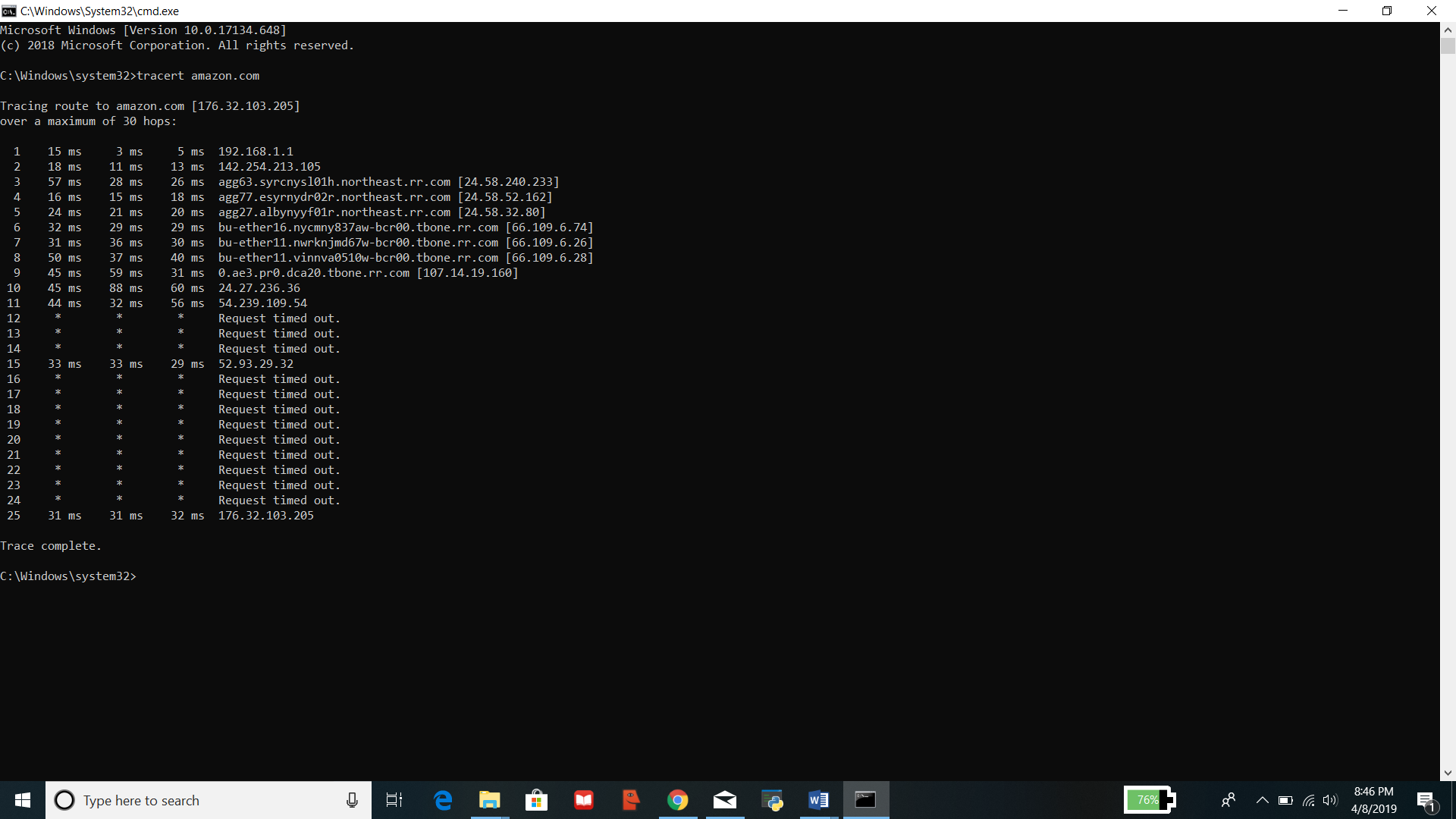
* tracert ecs.syr.edu



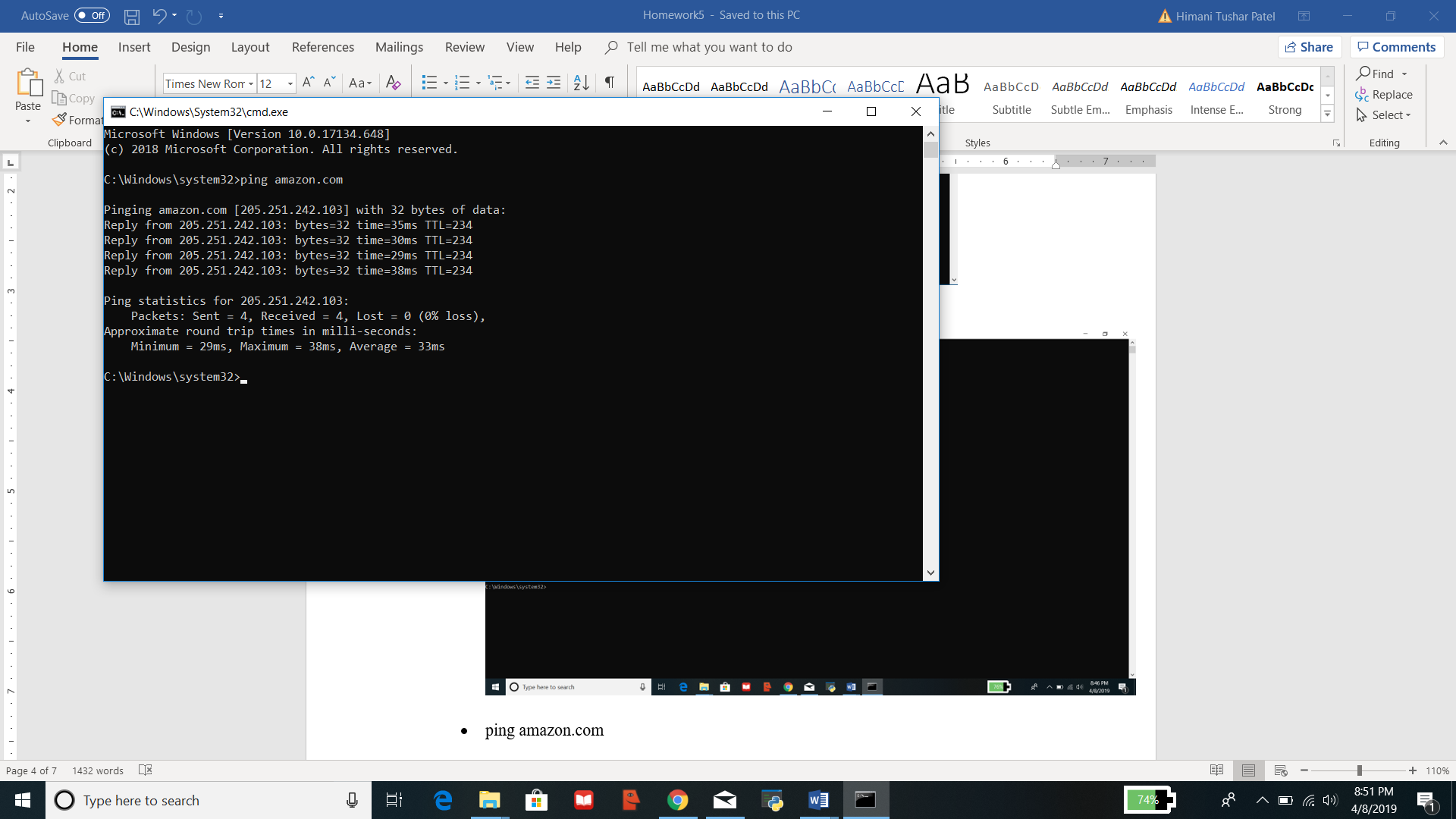
* ping ecs.syr.edu



* tracert amazon.com



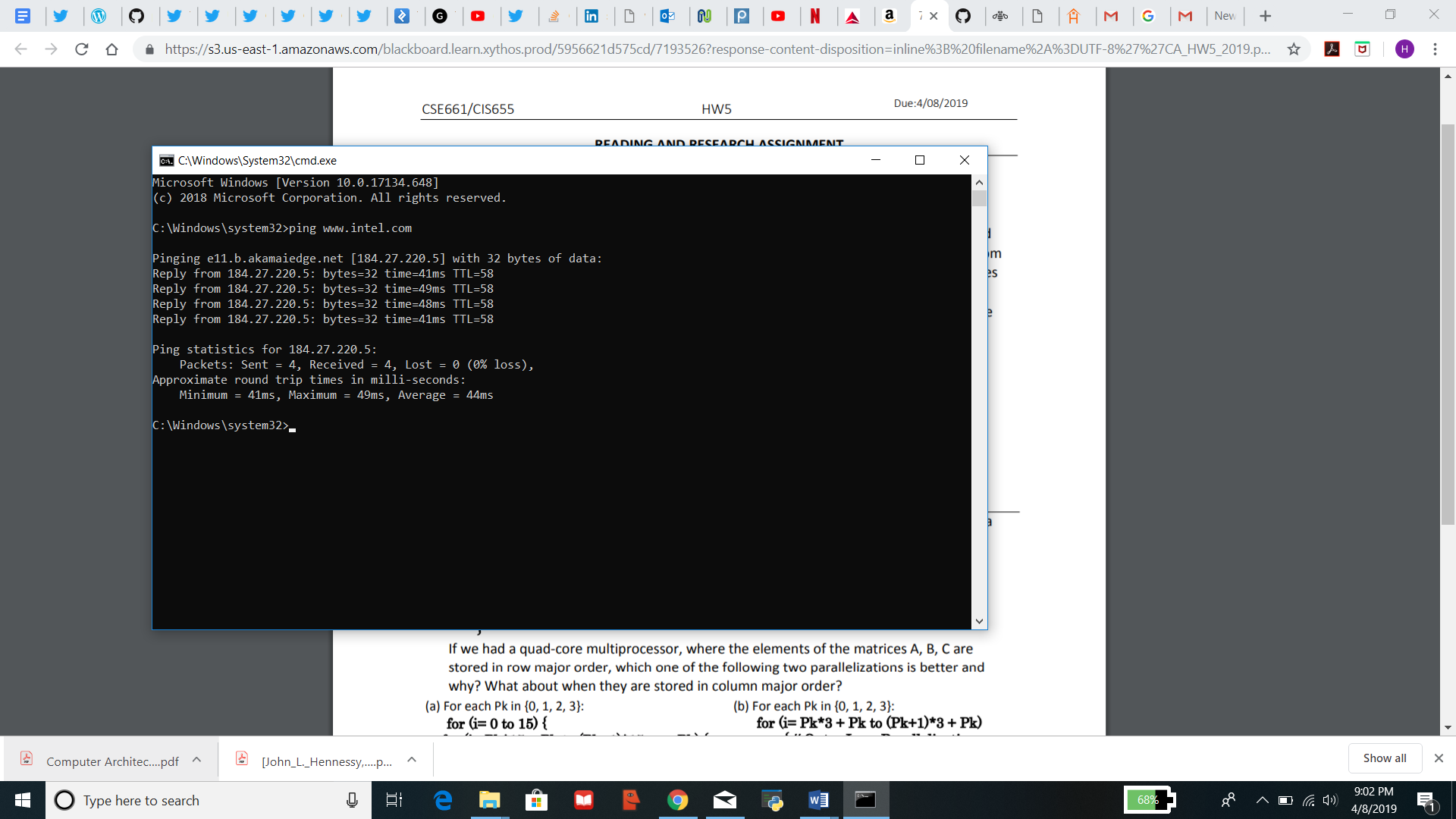
* ping amazon.com



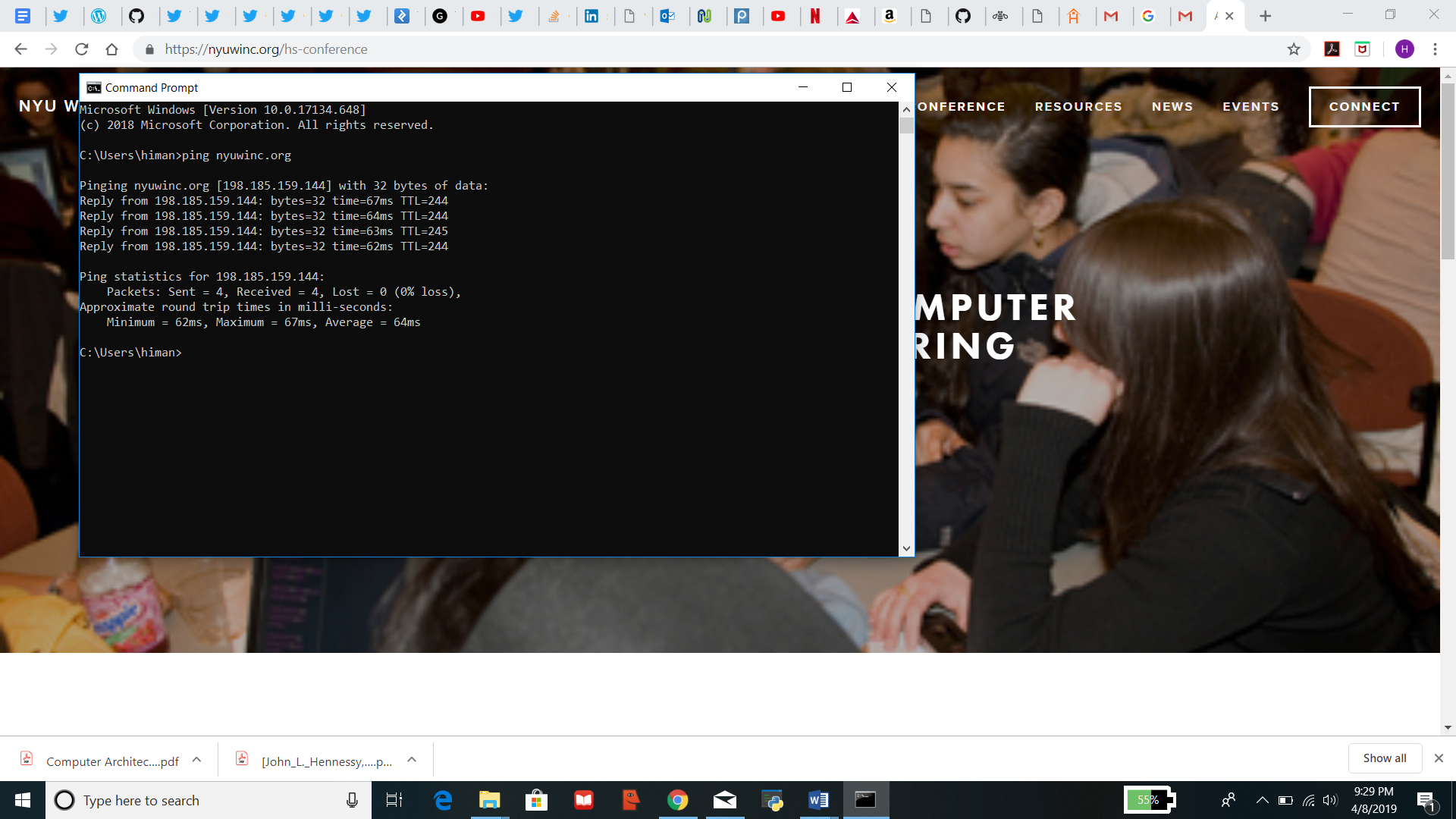
|  |  |  |
| --- | --- | --- |
| **Web server** | **Hop count** | **Average RTT** |
| Ecs.syr.edu | 13 | 40 ms |
| Amazon.com | 25 | 33 ms |

As we can observe that as the number of hops increases, i.e. as the hop count increases, the round-trip times (RTTs) also increases. But it is different for different routes, thus the correlation be different for the different routes.

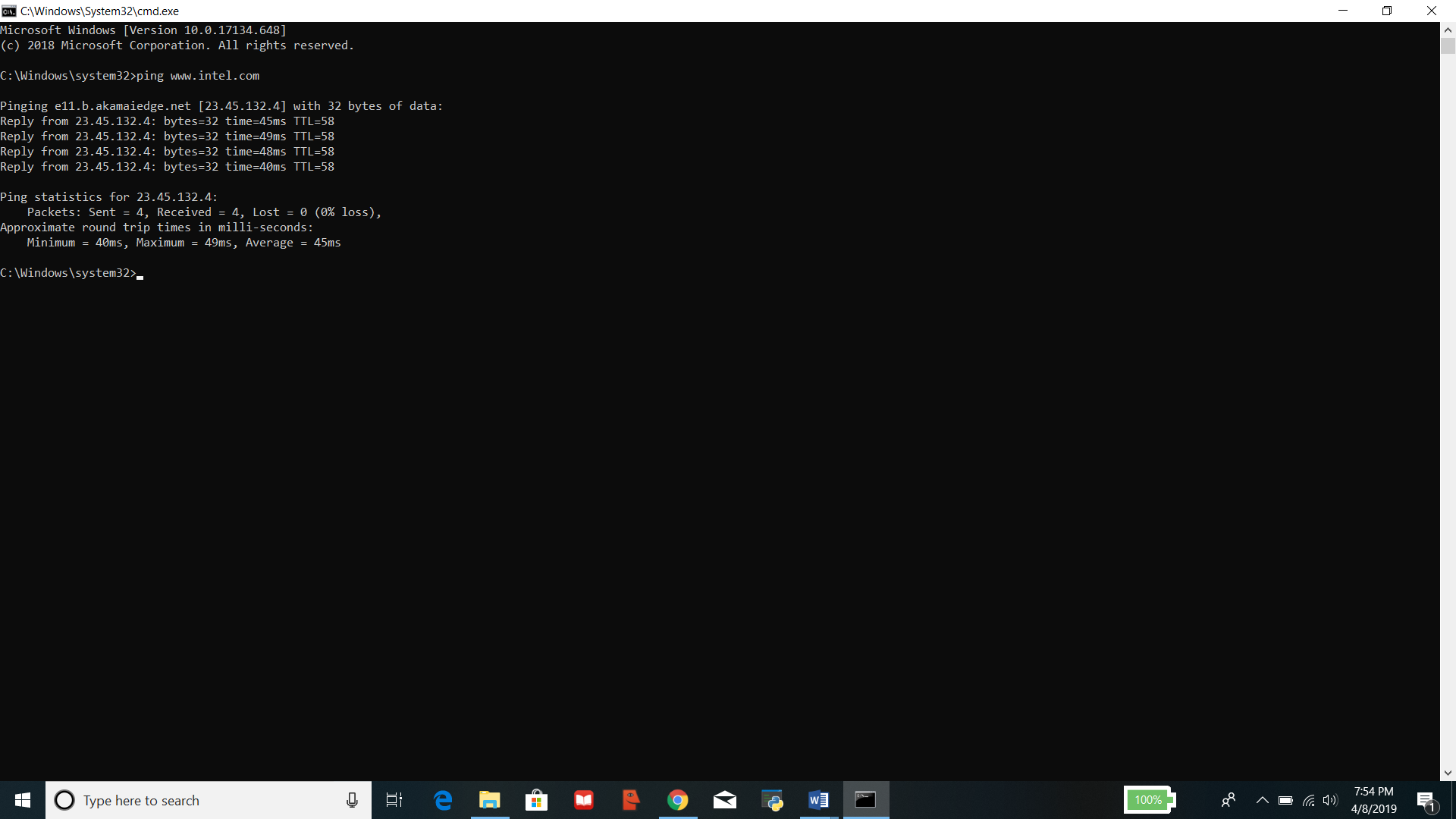
* ping intel.com



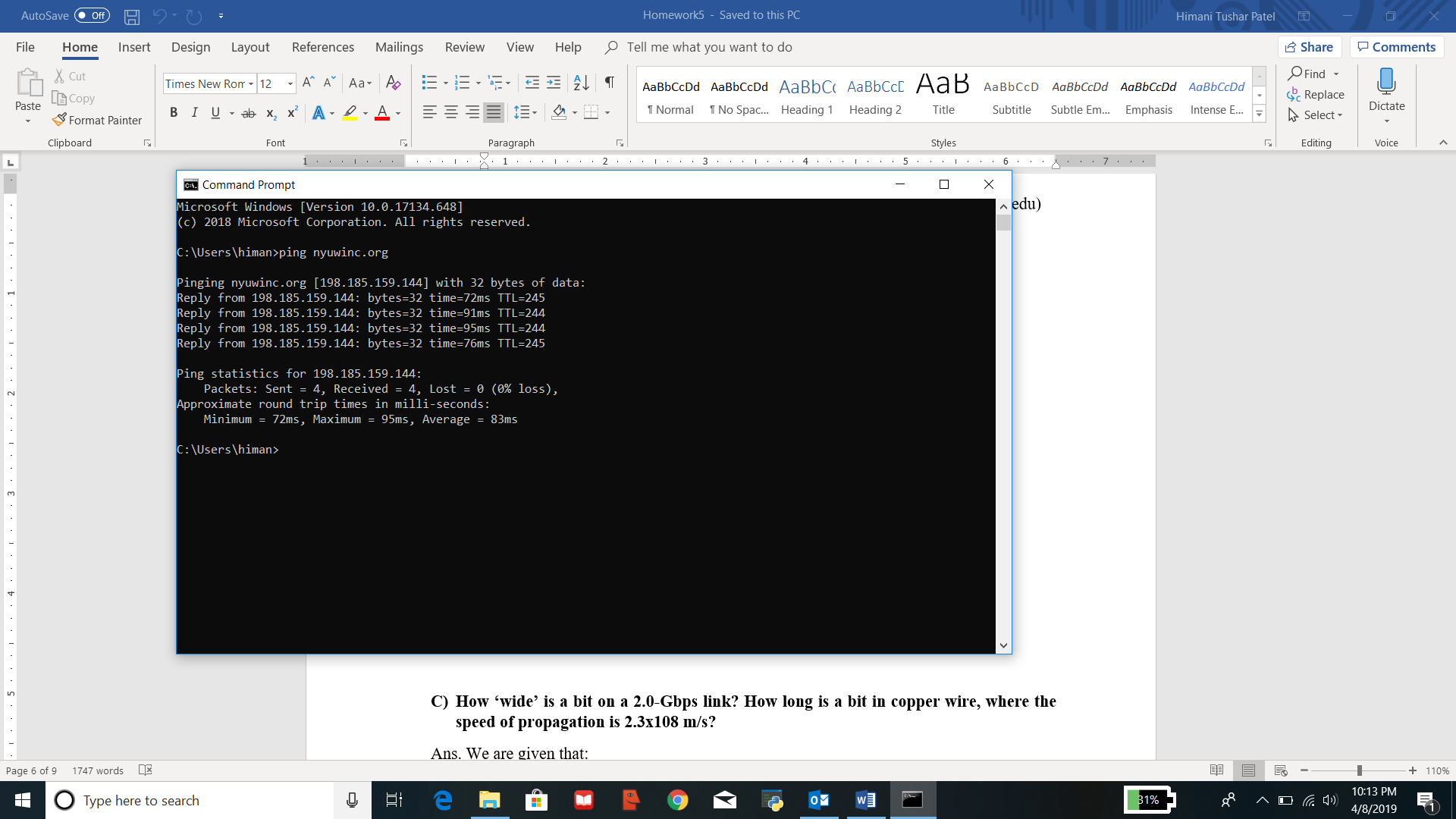
* ping nyuwinc.org (Using this site because it gave “Request timed out” for nyu.edu)



* ping intel.com



* ping nyuwinc.org



|  |  |  |
| --- | --- | --- |
| **Ping** | **RTT for** [**www.intel.com**](http://www.intel.com) | **RTT for www.nyuwinc.org** |
| 1 | 44 ms | 64 ms |
| 2 | 45 ms | 83 ms |

According to the above information we can observe that there is no correlation between the number of hops i.e. hop count and geographical distance. RTT might be dependent on traffic over the internet and the bandwidth allotment which can be different for both. Also, there might be other network reasons.

1. **How ‘wide’ is a bit on a 2.0-Gbps link? How long is a bit in copper wire, where the speed of propagation is 2.3x108 m/s?**

Ans. We are given that:

Link = 2.0 Gbps = 2 \* 109 bps

Speed of propagation = 2.3 \* 108 m/s

Assuming, the delay bandwidth is 1 bit.

Thus, width of a bit = = 0.5 \* 10-9 s = 0.5 ns.

It means, a bit on 2.0 Gbps link is 0.5 ns wide.

Therefore, time = 0.5 ns.

Now, we know that: Distance = Time \* Speed

Thus, Distance = (0.5 \* 10-9) \* (2.3 \* 108) m/s

Distance = 1.15 \* 10-1 m = 0.115 m

Therefore, it means that a bit in copper wire is 0.115 m.

**Exercises**

1. **Consider the following code that adds two matrices A and B and stores the result in a matrix C:**

**for (i=0 to 15) {**

**for (j=0 to 63) {**

**} C[i][j] = A[i][j] +B[i][j];**

**}**

**If we had a quad-core multiprocessor, where the elements of the matrices A, B, C are stored in row major order, which one of the following two parallelization is better and why? What about when they are stored in column major order?**

1. **Loop A:**

**For each Pk in {0, 1, 2, 3}:**

**for (i=0 to 15) {**

**for (j = Pk\*15 + Pk to (Pk+1) \* 15 + Pk) {**

**//Inner Loop Parallelization**

**C[i][j] = A[i][j] + B[i][j];**

**}**

**}**

1. **Loop B:**

**For each Pk in {0, 1, 2, 3}:**

**for (i = Pk\*3 + Pk to (Pk+1) \* 3 + Pk)**

**{//Outer Loop Parallelization**

**for (j=0 to 63) {**

**} C[i][j] = A[i][j] + B[i][j];**

**}**

Ans. Considering the two loops:

1. When the elements of the matrices A, B, C are stored in row major order, the parallelization in Loop A is better because every time we update the value of variable ‘j’, the elements are accessed row by row which means the threads will utilize the caches’ spatial locality property.
2. When the elements of the matrices A, B and C are stored in column major order, Loop B will perform better parallelization because this time the elements will be accessed column by column. Due to this, it will experience many number of cache misses as the elements in the columns are separated by a given number of columns which is equal to the row size.
3. **Textbook (Sixth Edition) 5.2 (a) (b)**

Ans. For the given following sequence of operations, the total number of stall cycles generated by each implementation are:

1. C0: R, AC20; Read miss, Satisfied by Memory

C0: R, AC28; Read miss, Satisfied by C1’s cache

C0: R, AC30; Read miss, Satisfied by Memory, Writeback AC10

For Implementation 1: 100 + 40 + 10 + 100 + 10 = 260 stall cycles

For Implementation 2: 100 + 130 + 10 + 100 + 10 = 350 stall cycles

1. C0: R, AC00; Read miss, Satisfied by Memory

C0: W, AC08 <- 48; Write hit, Sends Invalidate

C0: W, AC30 <- 78; Write miss, Satisfied by Memory, Writeback AC10

For Implementation 1: 100 + 15 + 10 + 100 = 225 stall cycles

For Implementation 2: 100 + 15 + 10 + 100 = 225 stall cycles

1. **Textbook (Sixth Edition) 5.22 (a) (b)**

Ans. We know that it is a 64-process distributed-memory multiprocessor.

Clock Rate = 2.0 GHz

Base CPI = 0.75

Instructions involving a remote communication reference = 0.2%

Cost of a remote communication reference = (100 + 10h) ns, where ‘h’ is the number of communication network hops.

1. The worst-case remote communication cost for 64 processors arranged as:

A Ring:

Largest number of communication network hops = 32

Cost = (100 + 10(32)) ns = 100 + 320 = 420 ns

An 8\*8 processor grid:

Largest number of communication network hops = 14

Cost = (100 + 10(14)) ns = 100 + 140 = 240 ns

A hypercube:

Largest number of communication network hops = (log2 64) = 6 (because longest communication path on a 2n hypercube has ‘n’ links)

Cost = (100 + 10(6)) ns = 100 + 60 = 160 ns

1. We are given that the Base CPI = 0.75

Comparing the CPI with the 64 processors arranged as:

A Ring:

Worst-case CPI = 0.75 + (0.2% of communication cost) = 0.75 + ((0.2/100) \* 420)

= 0.75 + 0.84 = 1.59 CPI

An 8\*8 processor grid:

Worst-case CPI = 0.75 + (0.2% of communication cost) = 0.75 + ((0.2/100) \* 240)

= 0.75 + 0.48 = 1.23 CPI

A hypercube:

Worst-case CPI = 0.75 + (0.2% of communication cost) = 0.75 + ((0.2/100) \* 160)

= 0.75 + 0.32 = 1.07 CPI