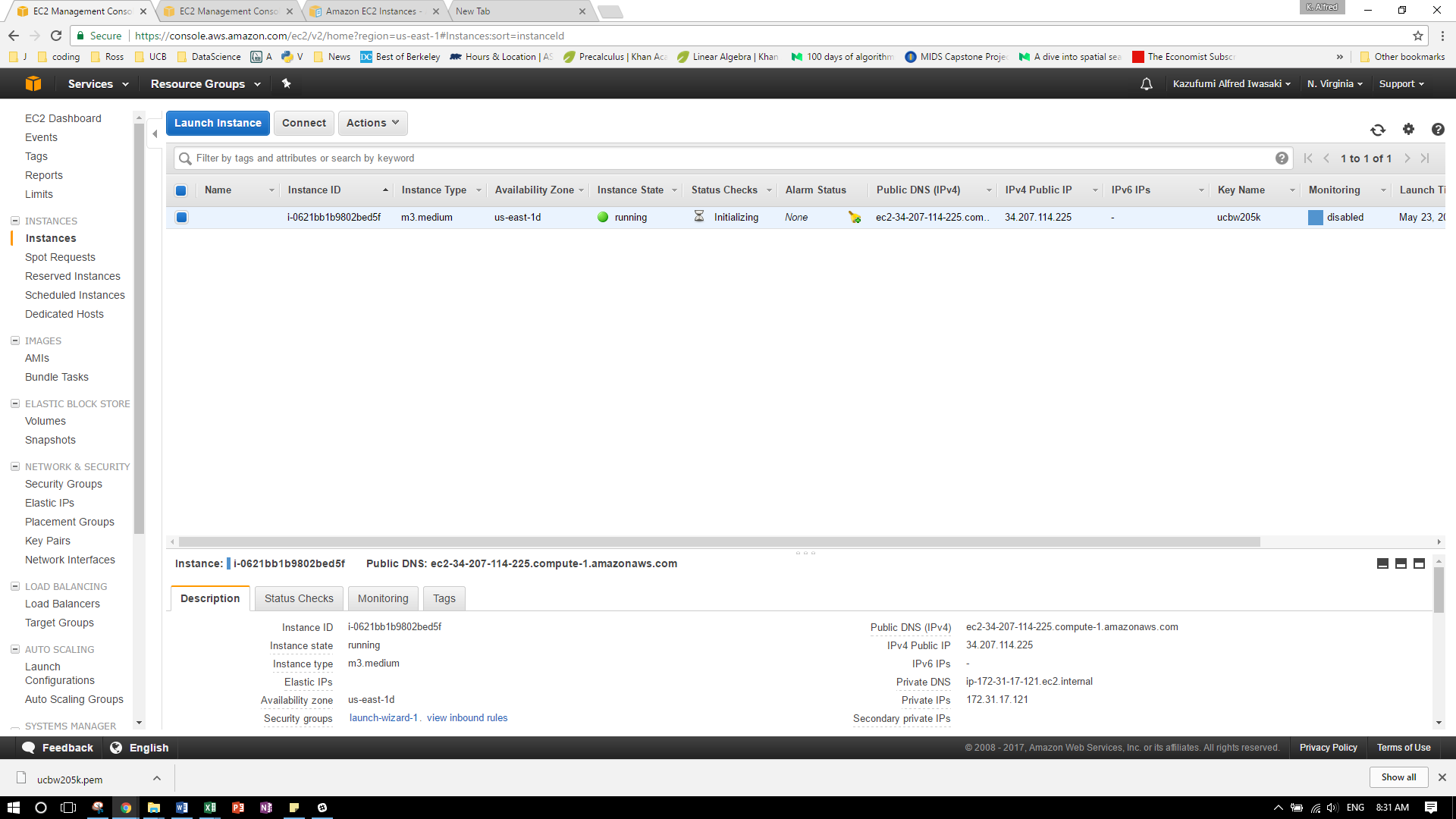
W205 summer 2017 section 2 – Lab01 - K Iwasaki

***1.Provide a screenshot of the running UCB AMI that shows you have successfully launched an EC2 instance using the UCB AMI.***



***2.Provide the answers to the following questions in a file called Lab1-answers.txt***

***Q1: What is the difference between EBS and S3?***

**EBS -** stores data as blocks of the same size and organizes them through the hierarchy similar to a traditional file system. This is *not* a standalone storage service so it can be used only in combination with Amazon EC2. Once you configure the volume in EBS, it’s can’t be easily scaled.

**S3 –** stores data as objects without hierarchy. Each object in the storage contains a header. Objects in Amazon S3 are associated with a unique identifier, so access to them can be obtained through web requests from anywhere. As compared to Amazon EBS, the process of requesting an object in S3 is slower, but S3 is highly scalable storage service with 99,99…% durability (very little chances to lose data or reach maximum storage capacity)

***Q2: When would you consider a cloud infrastructure for your data science tasks?***

**Look for scalability:** Rapidly increasing data can be a challenge for data scientist to analyze data. Data scientist waiting in front of the screen for data to be processed is a waste. A machine on cloud can scale on a click of the button.

**Look for lower cost:** If you have one-off project but it needs higher computational infrastructure. You can’t really buy a new machine for that. Rent out a higher configuration for a few hours or days a great solution to your headache.

**Look for larger ecosystem for ML system deployments:** A few cloud services like AWS, Azure provide complete ecosystem to collect data, run your models and then deploy them. This entire system must be built yourself in case of physical machine.

***Q3: What is the difference between spot instances and reserved instances?***

**Spot Instances** – are for applications that have flexible start and end times and for users with urgent computing needs. This is a bid for low price version of on-demand instance but could be shut down by AWS anytime the spot instance prices goes higher than bid price.

**Reserved instances** – are for applications that have steady state or predictable usage. Reserved instances provide user with a significant discount compared to using on-demand instances.

***Q4: List the names of four software packages installed on the UCB AMI***

PostgreSQL Database client, PostgreSQL Database server, MySQL Database client, My SQL Database server