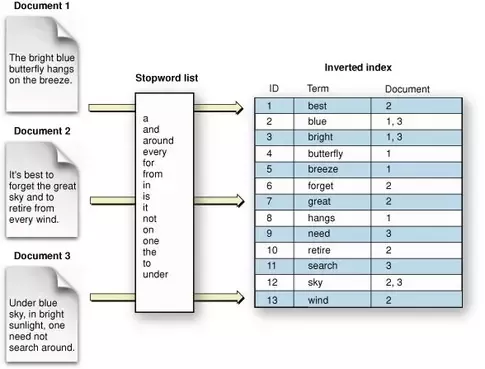
**Key Terms: -**

1. Elastic search is based on Apache Lucene’s search engine library written in Java. Reliable for full text search. Works on inverted index principle.



1. Cluster – One or more nodes instances (servers) connected. It helps in distribution of tasks, searching and indexing, across all nodes in cluster.
2. Nodes – Single server in a cluster which stores data and participates in cluster’s indexing and search capabilities.
3. Master Node – Tracking which nodes are part of cluster. Adding and removing nodes.
4. Indexes – It is Backbone. Elasticsearch is able to achieve fast search responses because, instead of searching the text directly, it searches an index instead.

You can (very roughly) think of an index like a database.

**MySQL => Databases => Tables => Columns/Rows**

**Elasticsearch => Indices => Types => Documents with Properties**

An Elasticsearch cluster can contain multiple Indices (databases), which in turn contain multiple Types (tables). These types hold multiple Documents (rows), and each document has Properties(columns).

**Example 1** - In your car manufacturing scenario, you may have a MarutiCarFactory index. Within this index, you have three different types:

People

Cars

Spare\_Parts

Each type then contains documents that correspond to that type (e.g. a Maruti Wagonr doc lives inside of the Cars type. This doc contains all the details about that particular car).

Searching and querying takes the format of: **http://localhost:9200/[index]/[type]/[operation]**

So to retrieve the Maruti document, I may do this:

<strong>$ curl -XGET localhost:9200/MarutiFactory/Cars/MarutiWagonr

**Example 2 - Indices for Logging**

Now, the reality is that Indices/Types are much more flexible than the Database/Table abstractions we are used to in RDBMs. They can be considered convenient data organization mechanisms, with added performance benefits depending on how you set up your data.

To demonstrate a radically different approach, a lot of people use Elasticsearch for logging. A standard format is to assign a new index for each day. Your list of indices may look like this:

logs-2013-02-22

logs-2013-02-21

logs-2013-02-20

Elasticsearch allows you to query multiple indices at the same time, so it isn’t a problem to do:

$ curl -XGET localhost:9200/logs-2013-02-22,logs-2013-02-21/Errors/\_search?query="q:Error Message"

Which searches the logs from the last two days at the same time. This format has advantages due to the nature of logs – most logs are never looked at and they are organized in a linear flow of time. Making an index per log is more logical and offers better performance for searching.

Indices are fairly lightweight data organization mechanisms, so Elasticsearch will happily let you create hundreds of indices.

**Example 3 - Indices for Users**

Another radically different approach is to create an index per user. Imagine you have some social networking site, and each users has a large amount of random data. You can create a single index for each user.

Your structure may look like:

Zach’s Index

Hobbies Type

Friends Type

Pictures Type

Fred’s Index

Hobbies Type

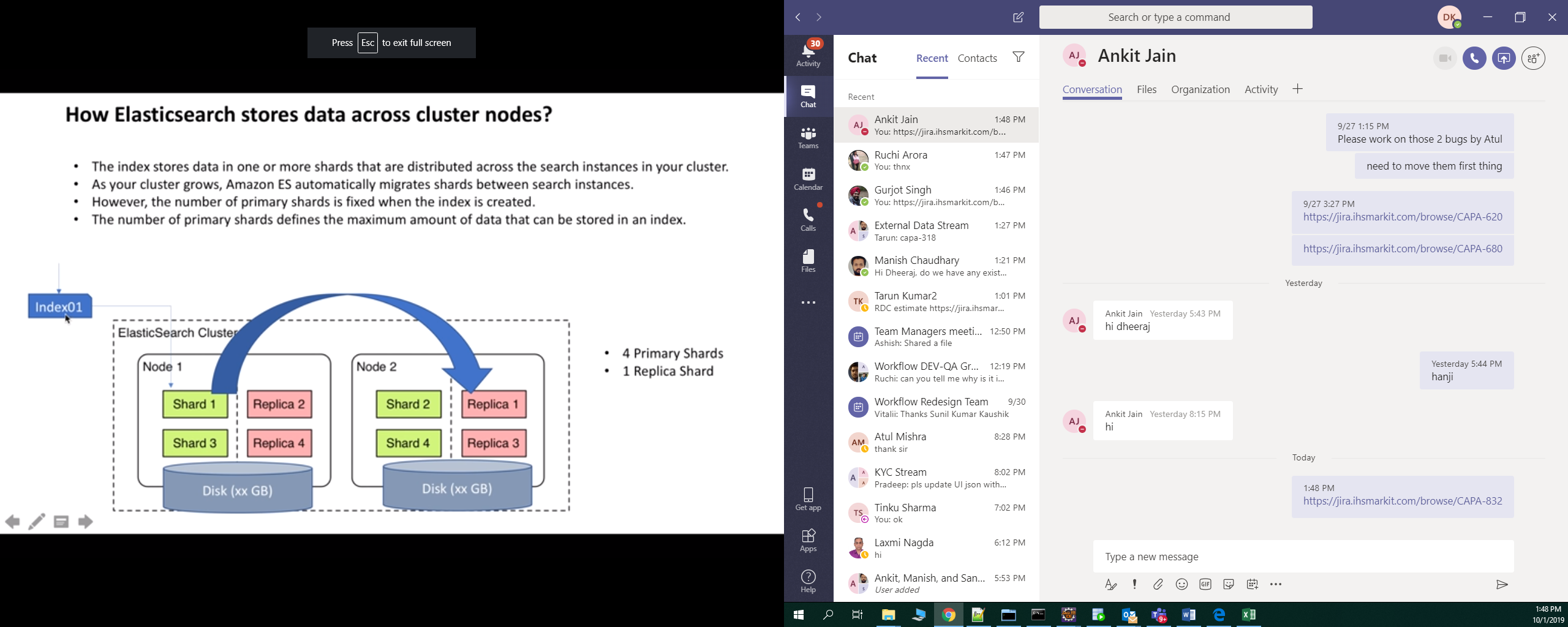
Friends Type

Pictures Type

1. Why amazon Elastic Search – To easily deploy and manage elastic search which becomes an admin task like checking nodes health, heartbeat etc.

Also provide services like Firehose (Analytic graphs on elastic search), Lambda and cloud Watch.

1. Shard and Replica – Indexes are stored on shard in one node and replica of same on another node. This duplication of data helps in handling scenario if one node is down.



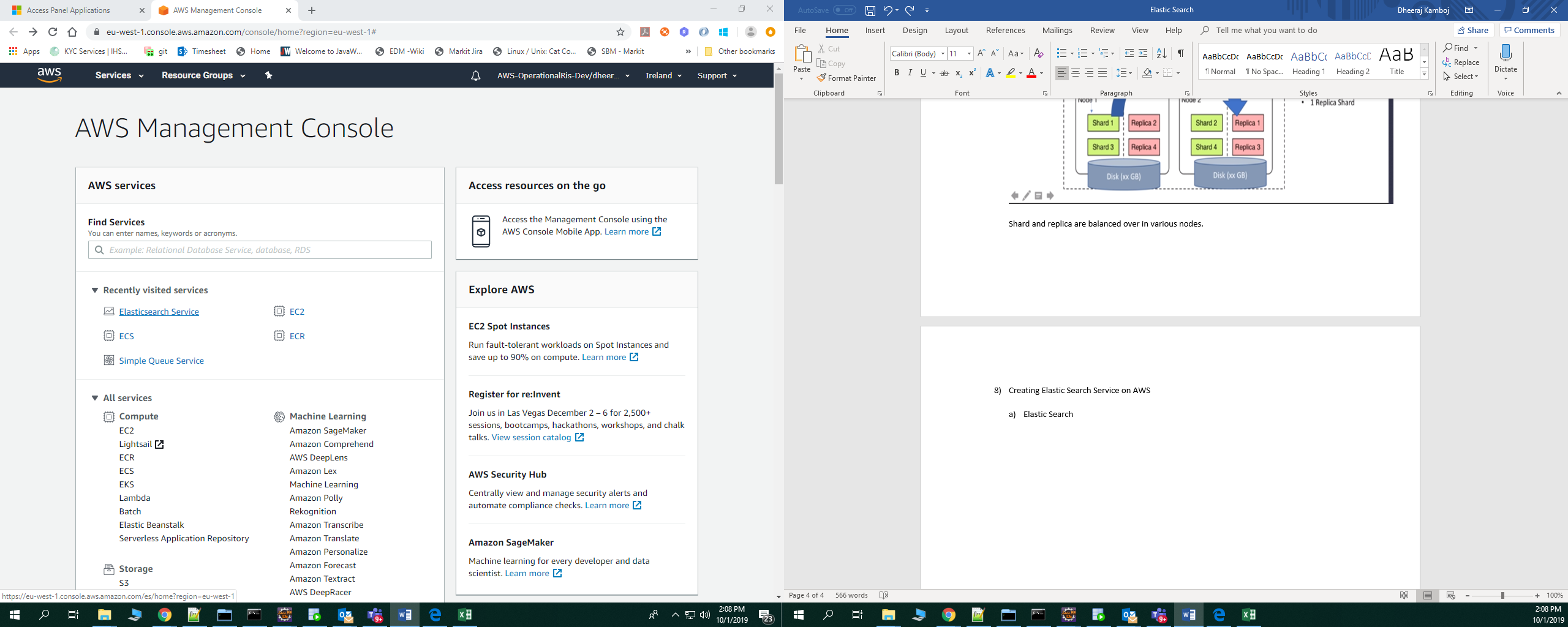
Shard and replica are balanced over in various nodes.

1. EBS - Amazon Elastic Block Store is a cloud-based block storage system provided by Amazon Web Services (AWS) that is best used for storing persistent data.

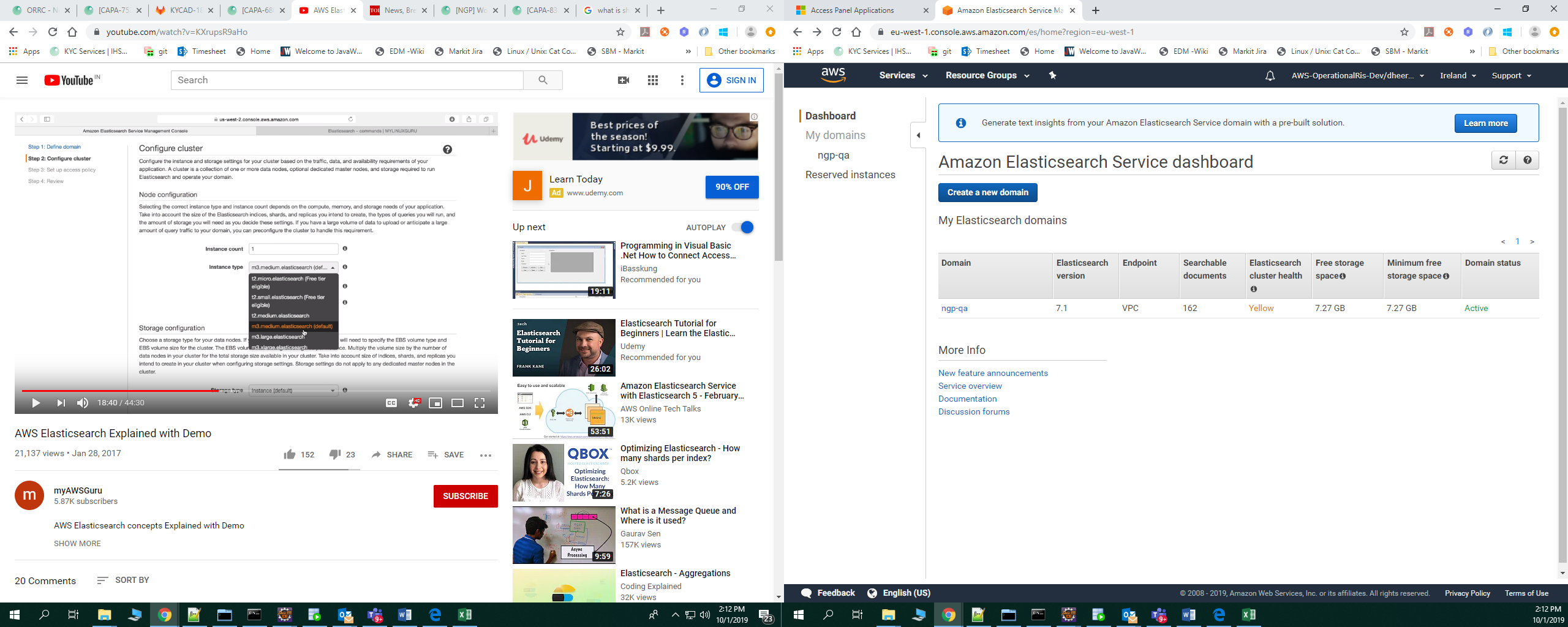
The service provides highly available block-level storage volumes for use with Amazon Elastic Compute Cloud (EC2) instances. Amazon EBS enables you to keep data persistently on a file system, even after you shut down your EC2 instance.

Amazon also provides local storage for EC2 instances that you can use while you run the instance, but you lose the data when you shut down the instance. This means, if you want to save that data, you need to store it on Amazon EBS.

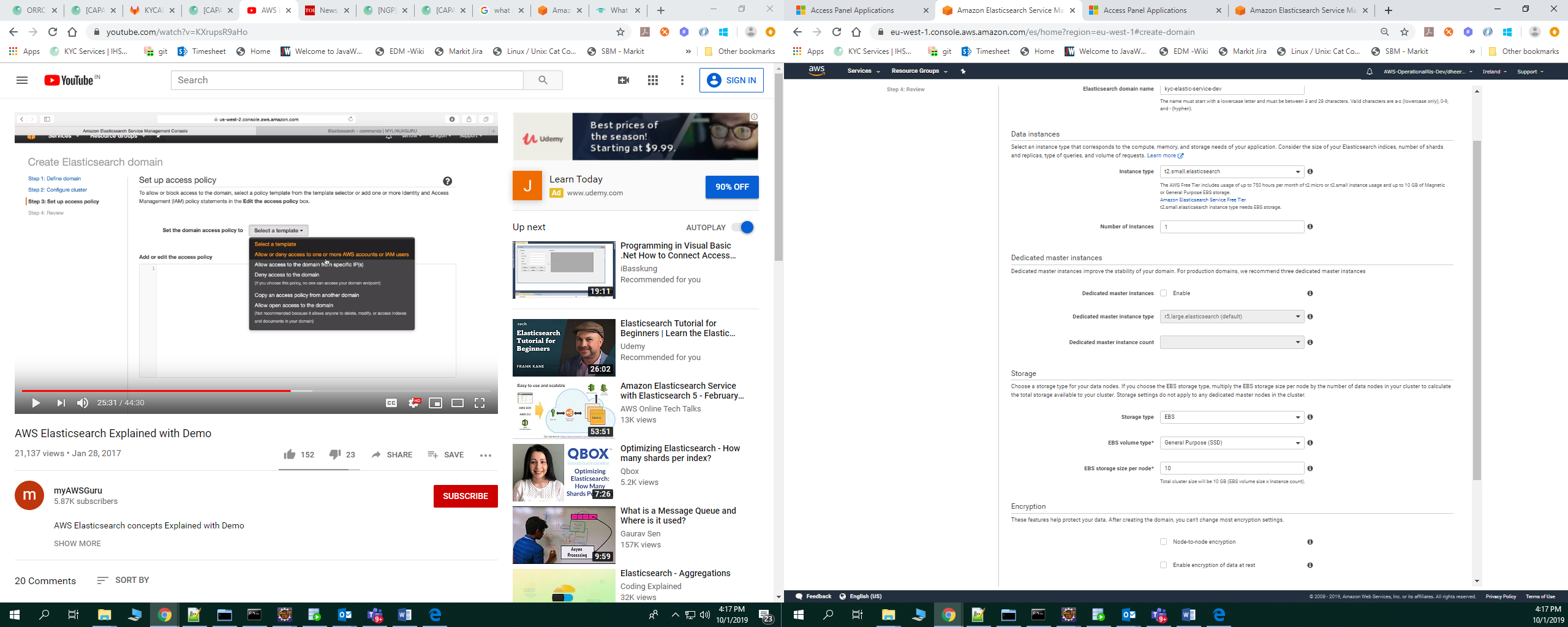
1. Creating Elastic Search Service on AWS
2. Elastic Search



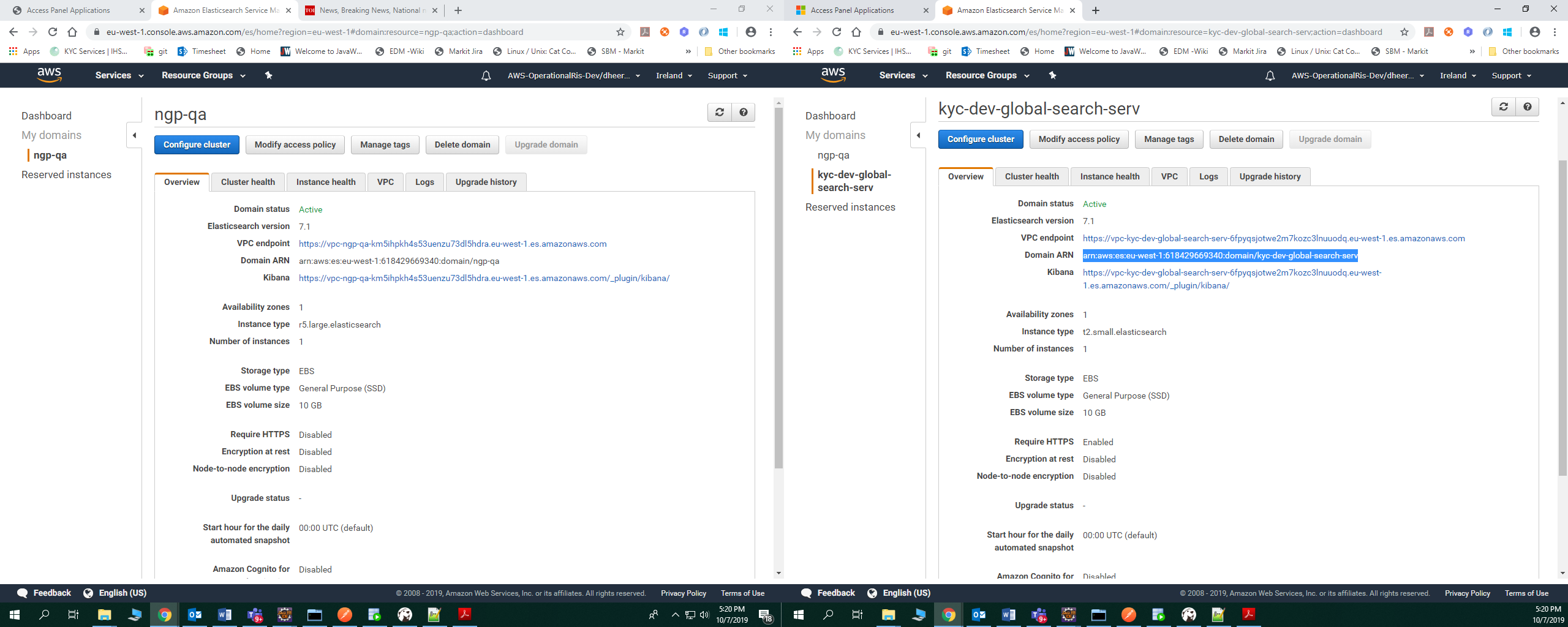
1. Create Domain

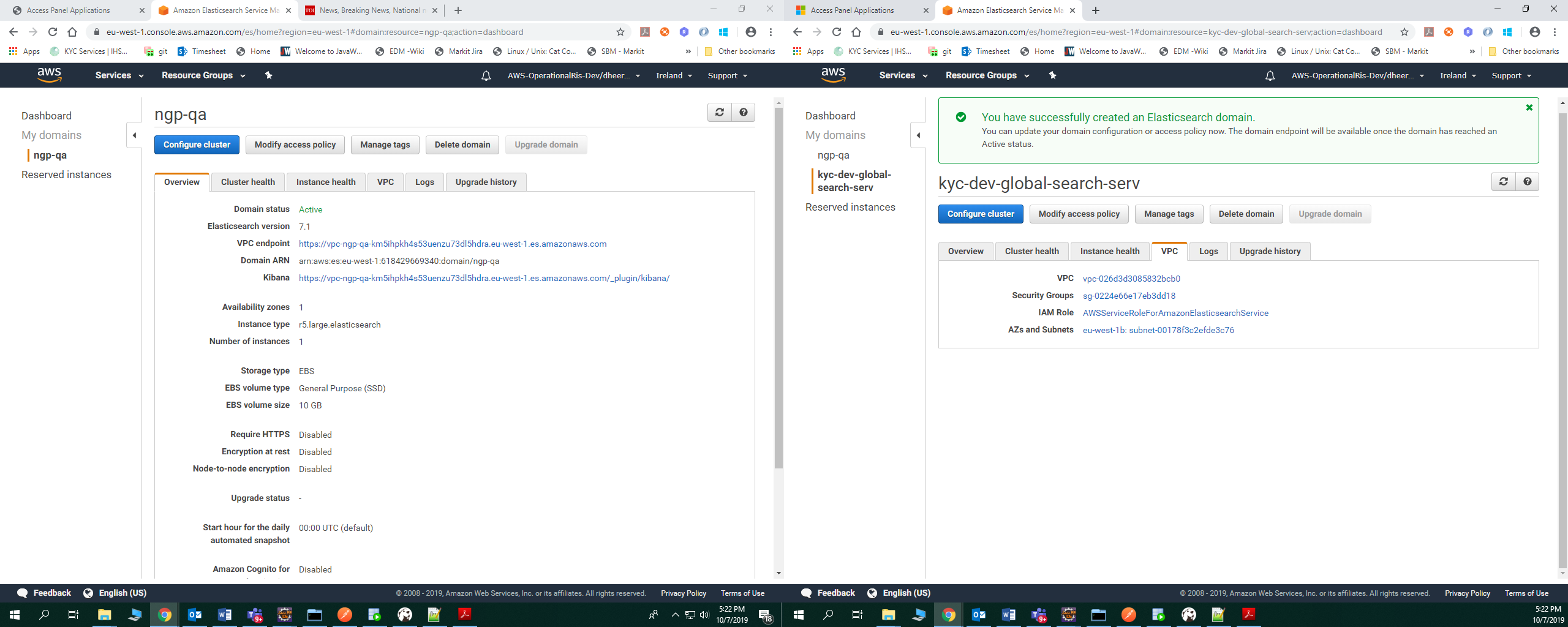


1. Config



1. Final





1. Cluster with single node will show yellow status.
2. You can create indices and types using curl command by connecting Linux aws machine.