Project Proposal
CS 410 – Text Information Systems
N. Wood (nadiaw2)

Search engine for indoor environment data using ElasticSearch

NetID: nadiaw2. I will be working on this project individually.

Abstract:

In the age of sensors, devices and platforms collecting millions of datapoints every second, it comes necessary to be able to sift through all the data to develop insights efficiently Big data offers the solution for analyzing large amount of data and using the technique of Elasticsearch, access to data can be made faster.[1]

I will be creating a web application in C#.net to use ElasticSearch to search content from a database of environmental datapoints collected by sensors (indoor air, humidity, temp etc). Currently, it is difficult to search for data in a RDMS database and it takes significant time using traditional SQL queries. The project will take the data transfer it to ElasticSearch server. The front end written in C# will allow users to search for data. Measurable outcomes are going to be the amount of time it takes to run a query against a traditional RDMS database vs. using ElasticSearch. The planned architecture is shown in Figure1:

---

[1] Gujarat, India, Darshita Kalyani, and Dr. Devarshi Mehta, "Paper on Searching and Indexing Using Elasticsearch," *International Journal Of Engineering And Computer Science*, June 30, 2017, https://doi.org/10.18535/ijecs/v6i6.45.

Cloud (AWS or Google or Azure)

ElasticSearch Cluster

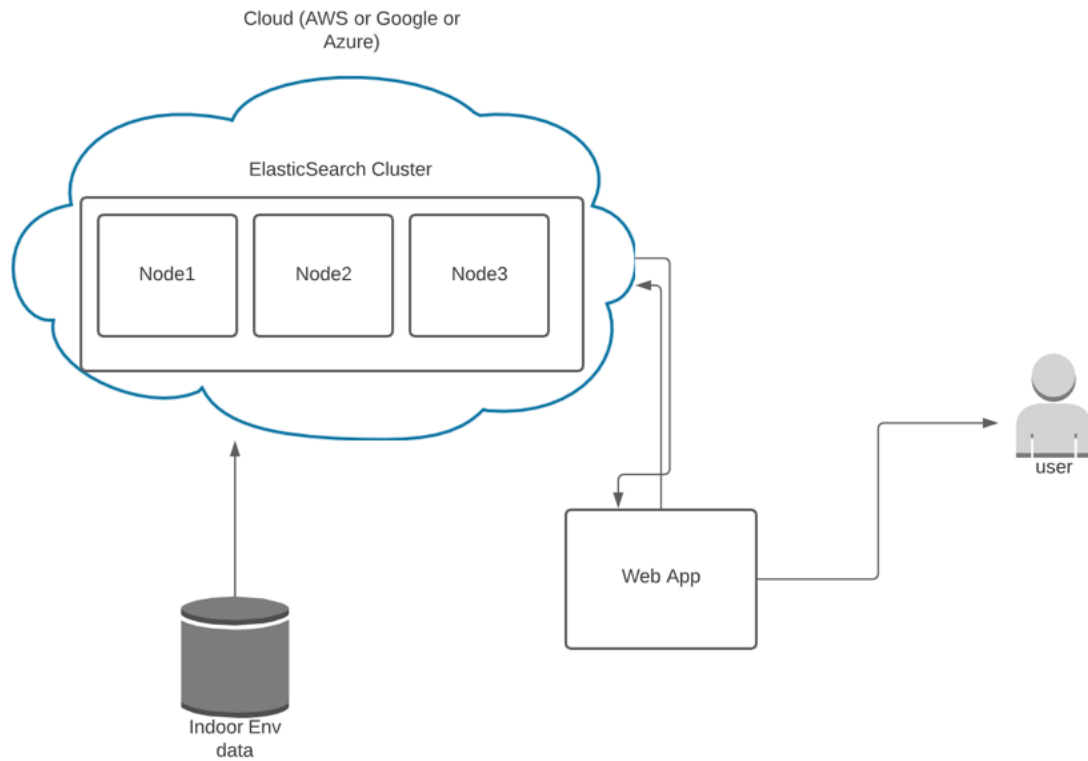Node1 Node2 Node3

Indoor Env data

Web App

user

Figure 1

The estimated time to complete this project is about 30-40 hours.

Tasks and time:

1. Determine how to set up, configure and deploy ElasticSearch on a cloud platform : 15 hours
   a. The cloud platform determination will base on the cost difference between Azure, Google and AWS.
2. Development and deployment of Web application: 10 hours
3. Migration of SQL database to Elastic search server : 3 hours
4. Test and measure outcomes: 4 hours