Data Analysis and Solution Prediction using Elasticsearch in Healthcare System

Arpitha R, C K Vanamala[2]

[1]Student,MTech-Computer Network and Engineering, Department of IS&E, The National Institute of Engineering, Mysuru-08
[2]Associate Professor, Department of IS&E, The National Institute of Engineering, Mysuru-08

Abstract:

Health care systems aims towards the betterment of the people by providing efficient health services with the help of technically advanced and high quality medical devices. Any defects in these equipment leads to a severe problem.Remote monitoring is one of the best approach to mitigate the occurrence of the defects by constant monitoring of the log files generated from the medical devices .But maintenance and analysis of these log records in real time is more challenging. Odata(Open Data Protocol) service in combination with elastic search plays a vital role in maintaining these large log files and also in analysis of the same to identify the root cause for the problem and provides a prognosticate and zealous solutions to mitigate the chances of unpredictable failure of the system. Odata by allowing the creation and consumption of REST(Representational State Transfer)ful APIs through CRUD operations ,forms an entry point for the elastic search which helps in a fast and free text search of the indexed data and provides an output in the JSON format.Kibana is an open source data visualization plugin for the elastic search.

Keywords: **Odata (Open Data Protocol)service, Elastic Search , Kibana, JSON,RESTful API,CRUD operations**

1 INTRODUCTION

Health care system is the organization of resources, people, and institutions that aims at providing the health care services to the people in need and works for the betterment of the lives of millions of people around the world. The medical equipment are the key components which benefits the patients by helping the health care faculty to diagnose and treat the patients to improve their health condition.But these products can be defective in its design, manufacturing process which may impact the patients diagnosis process and leads to a disaster.Hence the constant and relevant monitoring actions are necessary to make sure the defects are avoided, or resolved and does not affect the patients.Remote monitoring is one of the most widely accepted approach which helps in constant monitoring and analysis of the devices based on the log files generated by them. But managing this big log data is a big challenge because traditional technology is not powerful enough to process huge data and also it is difficult to write applications which works with multiple data sources as each of them will expose the data in a slightly different way.The combination of Odata services and Elastic search technologies eases the task of maintenance of large data, utilizing generic tools and can consolidate data from different information sources. and depicting it in a way easier to analyze it.

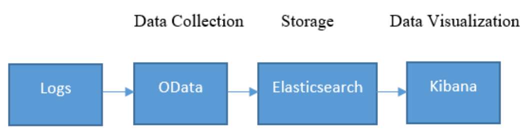

Fig 1.Data collection and search workflow

1.1 Open Data Protocol (OData)

[2]Odata is an open data access protocol from Microsoft that allows the uniform way of creation and consumption of query-able and interoperable RESTful APIs in a simple and standard

way.It is used to exchange the data over internet. The main advantages of Odata are :

- This service can be utilized by both the consumer and producer to devour and control the information The protocol is based on HTTP and hence any programming language with HTTP stack can consume the data.
- Resources are exposed [4]JSON format which is easy to understand.

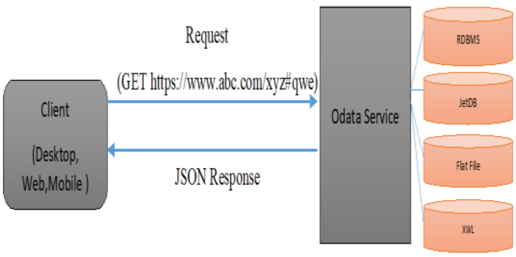

Fig 2.Odata Service

The request/response format in Fig 2. is as follows:

Method/ServiceURL/ResourcePath /Quey

a.**Method:**A methos is to specify the type of service that is being requested.It can be

i)GET:To retrieve the information from the server

ii)POST:To send the data to the server

b.**Service URL**: To specify the server system address to fetch the data from

c.**Resource Path**:To specify the location where the resource is located .

d.**Query** :To specify the actual resource required based on the conditon

As shown in Fig 1. Odata serves as an interface through which the data access is possible irrespective of the Schema in different databases. The above methodology is used to group each of the.

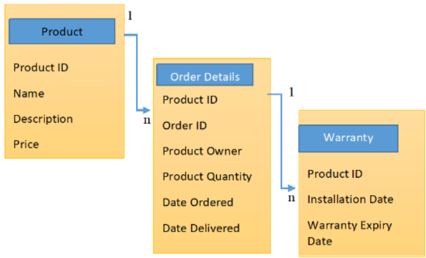

Fig 3.Relational Data model

The Relational data model structure is depicted in fig.3. The medical device related information which are present in different tables are retrieved through 1:n relation and all these information is consolidated into one JSON object which are the Key:Value pair sand returned as a response.Similarly the 'n' number of device details could be retrieved .As Shown, the Product ID is the key through which the data is retrieved and each product details is considered as one [6]index and the group of such indices forms a cluster which are the entry point for elastic search.

2 ELASTIC SEARCH

[1] Elasticsearch is an extremely powerful search and analysis engine build on top of Lucene search engine part of this power lies in the ability to scale it for better performance and stability. It exposes a[3] REST API which can be accessed using HTTP GET, DELETE, POST and PUT methods. Its API allows one not only to query or delete documents, but also to create indices, manage them, control analysis and get all the metrics describing current state and configuration of ElasticSearch

2.1 Elastic search key components

1.Fields

Fields are the smallest individual unit of data in Elasticsearch. Each field has a defined type and contains a single piece of data that can be, for example, a boolean, string or array expression. A collection of fields are together a single Elasticsearch document.

2.Document

Documents are JSON objects that are stored within an Elasticsearch index and are considered the base unit of storage.In relational database it can be compared to a row in the table.Document is consist of:

- index :The index where the document resides
- _type : The type that the document represents
- _id :The unique identifier for the document

Ex:

```
{ "_id" :1,
  "_type": "product",
    "name": "magnetic resonance"
}
```

3.Types

Elasticsearch types are used within documents to subdivide similar types of data wherein each type represents a unique class of documents. Types consist of a name and a mapping (see below) and are used by adding the *_type* field. This field can then be used for filtering when querying a specific type.

4.Mapping

mapping defines the different types that reside within an index. It defines the fields for documents of a specific type — the data type (such as string and integer) and how the fields should be indexed and stored in Elasticsearch.

5.Index

Indices, the largest unit of data in Elasticsearch, are logical partitions of documents and can be compared to a database in the world of relational databases

6 Analyzers

Analyzers are used during indexing to break down phrases or expressions into terms. Defined within an index, an analyzer consists of a single tokenizer and any number of token filters.

7.Cluster

An Elasticsearch cluster is comprised of one or more Elasticsearch nodes

2.2 Example query in elastic search

Consider an example of searching the defective areas based on the indexed data.Initially the query to create an index is as follows:

i).Index Creation

```
PUT /product_details

  { "settings": { "number_of_shards": 1 }}
```

```
POST /product_details  /product/bulk

  { "index": { "_id": 1 }}

  { "product id": "123","product name" : "product1", "malfunction_area" : "motor is not working }

  { "index": { "_id": 2 }}

  { "product id": "456","product name" : "product2", "malfunction_area" : "motor is not working and display problem" }

  { "index": { "_id": 3 }}
{ "product id": "789","product name" : "product3", "malfunction_area" : "motor is not working,display error,keeps shut down" }

  { "index": { "_id": 4 }}
{ "product id": "012","product name" : "product4", "malfunction_area" : "fan is not working,monitor problem " }
```

ii) **a. Basic Match Query**

```
GET /product_details /product/bulk/_search?q=keeps shut down
```

b. Field-wise Query

```
GET /product_details /product/bulk/_search?q=product id:123
```

c. Multi match Query with logical operators

```
GET /product_details /product/bulk/_ search?q=error AND monitor
```

```
GET /product_details / product /bulk/_ search?q=error OR monitor
```

```
GET /product_details / product /bulk/_ search?q=error NOT monitor
```

2.3 Advantages of elastic search over Lucene search engine

Even though Elastic search is built on top of [5]lucene, it extends the functionalities to provide a better performance and stability to the search.Some of the advantages are as follows:

- **Full-text search**
- **Fuzzy Searching- Finding a search result even there are spelling errors**
- **Autocompletion & Instant Search**
- **Document- oriented**
- **Query Fine Tuning**
- **Distributed approach-**Routing and rebalancing operations are done automatically when new documents are added.
- **Multi-Tenancy**

3.KIBANA

Kibana is an open source analytics and visualization platform designed to work with Elasticsearch. It used to search, view, and interact with data stored in Elasticsearch indices. It easily performs advanced data analysis and visualize the data in a variety of charts, tables, and maps.Some highlights of kibana are :

- Able to Put Geo Data on Any Map
- Time Series can also be on the Menu
- Analyze Relationships with Graph
- Ability to Explore Anomalies with Machine Learning

4. CONCLUSIONS

This paper mainly focuses on the importance of the management and analysis of the large volume of log records generated from different health care systems for the remote monitoring of the medical devices to identify the problems that may occur and to suggest a suitable solution for the same.The Odata services and Elasticsearch are the key approaches to achieve the goal in a much better and an efficient way.The combination of advanced technologies come up with the new solution which may improve the lives of billions of people in need

REFERENCES

[1] Feasibility Analysis of Big Log Data Real Time Search Based on Hbase and ElasticSearch

[2] OData for Service-Oriented Business Applications

[3] REST-based sensor networks with ODatas

[4] Real-Time or Near Real-Time Persisting Daily Healthcare Data into HDFS and ElasticSearch Index inside a Big Data Platform

[5] Design and Implementation of an Indexing Method Based on Fields for Elasticsearch

[6]https://logz.io/blog/10-elasticsearch-concepts/