**פרויקט גמר**

**מערכת ניהול וכריית מידע**

**Mining Data IIOT**

**SDD:**

**Software Design Document**

**שמות הסטודנטים:** אברהם שקד

כדורי שחר

**שם המנחה:** דרשר תמיר

**אישור מנחה למסמך:**

**תאריך:** 31.5.21

**Main Structure**

1. Introduction3

Purpose and Overview 3

Scope and Context 4-6

References and Related Documents/Links/Data-Links 7

Terms, Acronyms and Abbreviations8

1. Requirements Specification Essence9

Requirements Definitions Summary9

Use Cases Summary10-15

Traceability Matrix16

High-Level SW-Architecture Diagram17

1. Data Design18

ERD-Entity Relation Diagram18

1. Object Design19

Dynamic Object Modeling (Sequence Diagram)19

Static Object Modeling (Class Diagram) 20

1. User Interface Design21
2. Refined Project Plan22

**3. Data Design**

Formal/Informal Representation of:

* + ERD (if relevant)
  + DFD (if relevant)
  + Other DB Schemas (e.g. XML Schemas, etc.)
  + Complete Design of non-standard Communication-Protocols (if relevant):
    1. Messages Packet Structures
    2. Relevant Data Types (for each field)
    3. **Field 🡸🡺 Layer** association

**3. Data Design**

Formal/Informal Representation of:

* + ERD (if relevant)
  + DFD (if relevant)
  + Other DB Schemas (e.g. XML Schemas, etc.)
  + Complete Design of non-standard Communication-Protocols (if relevant):
    1. Messages Packet Structures
    2. Relevant Data Types (for each field)
    3. **Field 🡸🡺 Layer** association

**1. Introduction**

**1.1. Purpose and Overview**

אחת הבעיות במפעלי ייצור וציוד ובדיקה היא יצירת לוגים שונים בלתי אחידים כלומר בעלי מבנה לא ברור, וברוב המפעלים המידע לגבי התפקוד של המכונות יכול להתפספס ושונה בין מפעלים, חברות, ולפעמים מכונות. מטרת הפרויקט היא פיתוח מערכת ניהול וכריית מידע ממכונות ייצור וציודי בדיקה אלו, אשר מפיקים את הלוגים והנתונים שלהם בפורמטים בלתי אחידים. מטרות נוספות הן ביצוע כריה של המידע על פי הגדרה על מנת לכרות אתה המידע מהלוגים השונים בצורה דינאמית מאחר וכל לוג שונה במבנה, ביצוע פעולות עיבוד על המידע לצורך קבלת תצורה נוחה לשמירה במערכת. למעשה ,המערכת שלנו תשלב את כלל הדו"חות המתקבלים מן המקורות השונים ע"י איזשהו שרת ותבצע עליהם טרנספורמציה כך שיוכלו לעבור דרך המערכת וזו תהפוך את המידע לקריא ונגיש.

המערכת מהווה בסיס לניתוח הלוגים המתקבלים מהמכונות השונות. היא מבצעת ניתוח והצגת המידע באופן מסודר ומאפשרת למשתמש לקבל מידע בזמן אמת. המשתמש יוכל לבחור איזה חלקים בדוח הוא רוצה לראות.

החלק התכנותי של המערכת מורכב מ 3 חלקים עיקריים:

1. PARSER :

* יבצע התחברות למסד הנתונים.
* איסוף לוגים מבסיס הנתונים
* המרתן לפורמט הניתן לשימוש ע"י כלי הניתוח, פורמטים כגון: JSON / CSV
* העברת הקבצים המפורמטים אל ה מערכת ה ELK – elastic search and kibana

1. הelastic search - מבצע אנליזה על המידע שיתקבל מה PARSER.
2. ניתוח המידע ע"י elastic search והצגת גרפים יבוצע על גבי dashboards באמצעות מערכת הצגת הנתונים הפועלת יחד עם elasticSearch הנקראת Kibana .

**1.2. Scope and Context**

**Relevant Regulatory Requirements**

* The big goal of the system is to pull the logs of machines the database

and manipulate the data taken from it for the benefits of the users.

* Database should be SQL, other databases will require changing in the software.
* Software can be change only by the developers of the product (Aka IIOT System).
* The software will not be used for purposes other than those for which it is intended.

**Software Safety Class**

This software requires the software safety class: A.

**Context and External Interfaces Overview**

Neighbor systems:

* SQL Server Management software system like SSMS, MYSQL, and etc.

Logical Business Data:

* Log Files

Diagram

Description automatically generated

**Dynamic UI Behavior**

In this section we will describe the behavior of our user interface space.

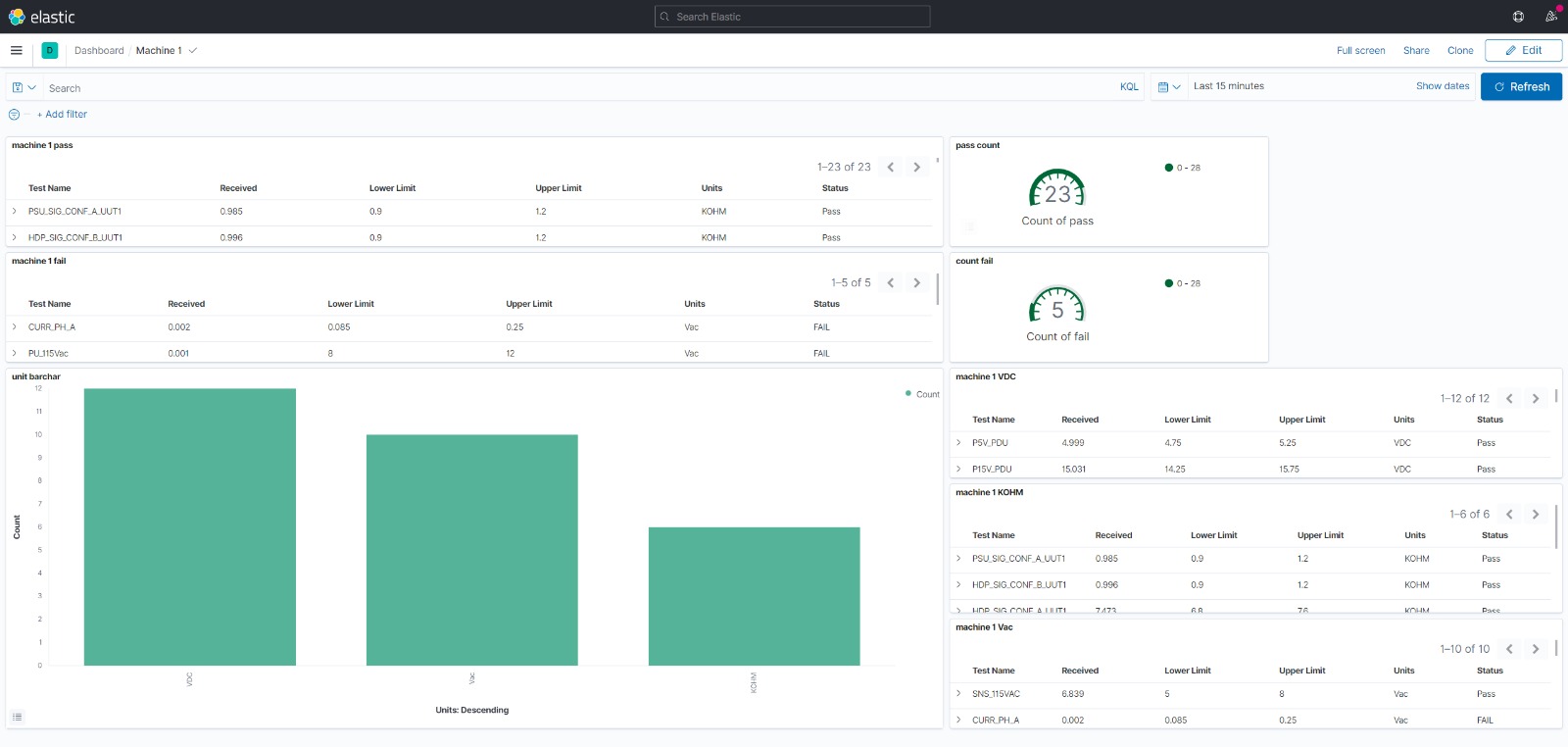
במילים אחרות אנחנו נתאר את ההתנהגות של המסך שיראה המשתמש.Diagram

Description automatically generated

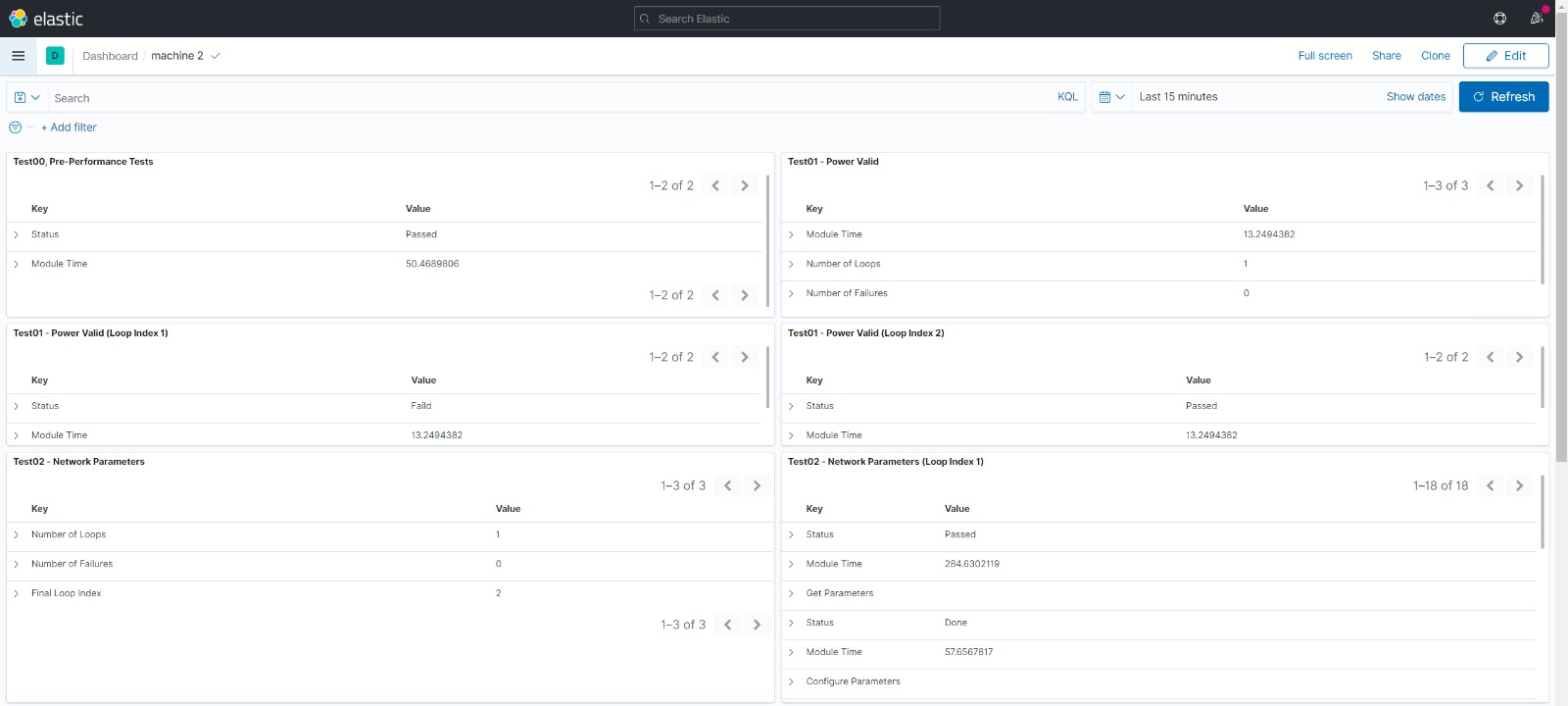
**Static UI**

**Kibana Dashboard**

**Panel 1**



**Panel 2**



**1.3 References and Related Documents/Links/Data-Links**

<https://www.learnentityframeworkcore.com/>

<https://www.mssqltips.com/sqlservertip/5771/querying-sql-server-tables-from-net/>

<https://docs.microsoft.com/en-us/dotnet/api/system.data.sqlclient.sqlconnection?view=netframework-4.7.2>

<https://docs.microsoft.com/en-us/aspnet/core/fundamentals/http-requests?view=aspnetcore-5.0>

<https://www.newtonsoft.com/json/help/html/SerializeObject.htm>

<https://html-agility-pack.net/documentation>

<https://www.guru99.com/c-sharp-access-database.html>

<https://github.com/Dynalon/JsonConfig>

<https://www.w3schools.com/js/js_json_intro.asp>

<https://www.json.org/json-en.html>

<https://www.youtube.com/channel/UC7z5VlhDHnorjUm6oW5dXcw/>

<https://www.elastic.co/guide/index.html>

<https://www.elastic.co/start>

**1.4 Terms, Acronyms and Abbreviations**

* CSV – Comma-Separated Values File.

Explanation - csv is used to store data in this particular way of comma separated values to export/import data from and for other programs and also can easily viewed by different editors

* HTML – Hyper Text Markup Language.

Explanation - HTML is the standard markup language for documents designed to be displayed in a web browser and are the building block of HTML pages.

* SQL – Structured Query Language.

Explanation - SQL is a standard language for storing, manipulating and retrieving data in databases.

* ELK Stack – Elastic Search , Log Stash, Kibana

Explanation - the **ELK Stack** is a collection of three open-source products which created by [Elastic NV](https://www.google.com/search?sxsrf=ALeKk00M_jxlS4WWu63CPkWb8e-oDb0NtA:1627207980809&q=Elastic+NV&stick=H4sIAAAAAAAAAONgVuLVT9c3NMwwiDcrMbIoXsTK5ZqTWFySmazgFwYACVGfyh4AAAA&sa=X&ved=2ahUKEwjr-66N_v3xAhWLy6QKHT-rB0oQmxMoATAsegQINBAD) and stands for — Elasticsearch, Logstash, and Kibana. **ELK stack** provides centralized logging in order to identify problems with servers or applications. It allows you to search all the logs in a single place.

**E** stands for ElasticSearch: used for storing logs and analyze them.

**L** stands for LogStash : used for both shipping as well as processing and storing logs.

**K** stands for Kibana: is a [visualization tool](https://www.guru99.com/best-data-visualization-tools.html) (a web interface) which is hosted through Apache and etc, which is a http web server.

**2. Requirement Specification Essence**

**2.1 Requirements Definitions Summary**

**Functional requirements**

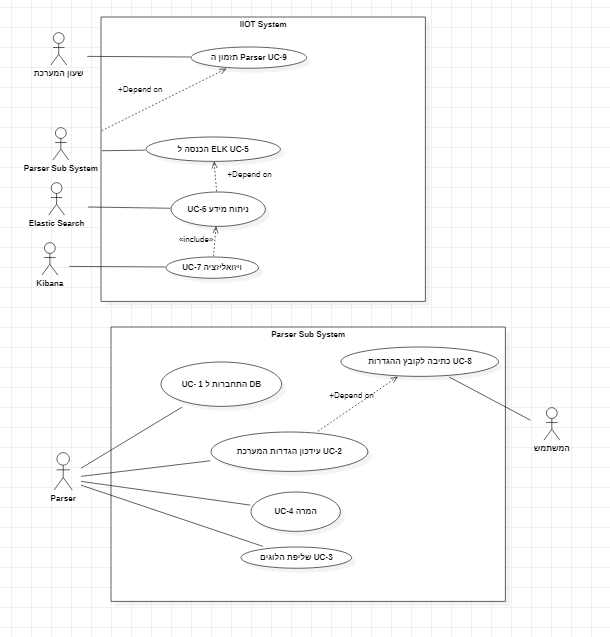
|  |  |  |
| --- | --- | --- |
| Requirement | Priority | Identifier |
| המכונות יעלו את הדו"חות באופן עצמאי על מנת שהמידע יהיה נגיש לשליפה מבסיס הנתונים הקיים. | 2 | REQ01 |
| המערכת צריכה לספק אפשרות למשתמש להגדיר אלמנטים במבני לוגים שונים ל parsing ומה לפרסר על מנת לבצע parsing בצורה דינאמית ע"י קובץ הגדרות. | 1 | REQ02 |
| המערכת צריכה לקרא את קובץ ההגדרות כל יום בהתאם לזמני ביצוע ה parsing. | 2 | REQ03 |
| המידע מהלוגים צריך להיות נגיש לקריאה עד יום מרגע הכנסתם לבסיס הנתונים. | 3 | REQ04 |
| המידע מהלוגים צריך להיות מ"פורסר" ומועבר ל ELK מידי יום באופן שוטף 24/7. | 2 | 5REQ0 |
| המערכת צריכה להציג את הנתונים על גבי ממשק נוח וקריא אשר יספק מידע עדכני ובזמן אמת 24/7 אשר יהיה זמין למשתמש. | 4 | 6REQ0 |
| המשתמש אחראי לשנות את תוכן קובץ ההגדרות על מנת להתחיל את תהליך הפעלת המערכת. | 5 | 07REQ |
| לוגים לא תקינים שיתקבלו על ידי המערכת לא יעברו "פרסור" והמידע שלהם לא יעבור הלאה ל Elastic Search. | 4 | REQ08 |
| המידע המפורסר שמגיע מהלוגים צריך להיות מנותח בצורה כלשהי לפרטים על מנת לספק מידע מדויק ואמין על גבי מצב המוצרים שהמכונות בודקות. | 1 | REQ09 |

**2.2 Use Cases Summary - Including Actors Table, UML Use Case Diagram and Glossary**

**Actor Table**

|  |  |  |
| --- | --- | --- |
| Use Case Name | Actor’s Goal (what the actor intends to accomplish) | Actor |
| התחברות ל DB (1-UC) | יצירת חיבור בין מסד הנתונים הנוכחי למערכת IIOT. | Parser |
| עדכון הגדרות המערכת (UC-2) | מעדכן את ההגדרות המערכת למה שרשום בקובץ ההגדרות בשביל להמיר בצורה נכונה את הלוגים על פי קובץ ההגדרות. | Parser |
| שליפת הלוג (UC-3) | שליפת הלוג הרלוונטי מתוך מסד הנתונים על פי ההגדרות. | Parser |
| המרה (UC-4) | המרת הלוג מדף HTML לפורמט CSV. | Parser |
| הכנסה לELK (UC-5) | ביצוע תקשורת עם Elastic Search לצורך העברת הלוג בפורמט CSV אליו. | Parser |
| ניתוח מידע (UC-6) | חיפוש ואנליזה בזמן אמת של המידע מהלוגים. | Elastic Search |
| ויזואליזציה ( UC-7) | מציגה את המידע על גבי dashboard מאפשרת visualization של המידע. | Kibana |
| כתיבה לקובץ הגדרות (UC-8) | כותב לתוך קובץ ההגדרות את ההגדרות לפיהם הוא מעוניין לפרסר את קבצי הלוג שלו. | המשתמש |
| תיזמון ה Parser. (UC-9) | מתזמנת את הParser לפי זמני המערכת. | שעון המערכת |

**Use Cases Diagram**



|  |  |
| --- | --- |
| **Use Case UC-1: התחברות למסד הנתונים** | |
| **Related Requirments** | REQ01, REQ04 |
| **Initiating Actor** | Parser |
| **Actor’s Goal** | התחברות למסד הנתונים |
| **Preconditions** | קבלת הפרמטרים של מסד הנתונים מקובץ ההגדרות. |
| **Flow of Events for Main Success Scenario** | |
| 1. The System established a connection with the DB. 2. The System ready to get the configuration data. | |
| **Flow of Events for Main Failure Scenario** | |
| 1. The System did not established a connection with the DB.  2. The System is not ready to get the configuration data. | |

|  |  |
| --- | --- |
| **Use Case UC-2: עדכון הגדרות המערכת** | |
| **Related Requirments** | REQ02, REQ03, REQ08 |
| **Initiating Actor** | Parser |
| **Actor’s Goal** | מעדכן את ההגדרות המערכת למה שרשום בקובץ ההגדרות בשביל להמיר בצורה נכונה את הלוגים על פי קובץ ההגדרות, זאת בשל מקרה שמישהו הכניס הגדרות אחרות מאלה שהכניסו בהתחלה לתוך קובץ הjson . |
| **Preconditions** | חיבור עם מסד הנתונים 1-UC, כתיבה לקובץ ההגדרות UC-8, תיזמון הParser UC-9. |
| **Flow of Events fo Main Success Scenario** | |
| 1. The Parser take the input (json config file) and stores\update the suitable json configuration file object. 2. The Parser read the propreties of the json to decide how to parse 3. The Parser start parsing according to the configurations | |
| **Flow of Events for Main Failure Scenario** | |
| 1. Configuration file is empty, thus the Parser take empty empty propreties. 2. The Parser behave by its default. | |

|  |  |
| --- | --- |
| **Use Case UC-3: שליפת הלוג** | |
| **Related Requirments** | REQ04, REQ08 |
| **Initiating Actor** | Parser |
| **Actor’s Goal** | שליפת הלוג הרלוונטי מתוך מסד הנתונים על פי ההגדרות. |
| **Preconditions** | התחברות למסד הנתונים UC-1, תזמון ה Parser UC-9. |
| **Flow of Events for Main Success Scenario** | |
| 1. The Parser take as input a valid log. 2. The Parser start the process in which the log is parsed. | |
| **Flow of Events for Main Failure Scenario** | |
| 1. The Parser take as input an invalid log.  2. The Parser continue to the next log. | |

|  |  |
| --- | --- |
| **Use Case UC-4: המרה** | |
| **Related Requirments** | REQ05 |
| **Initiating Actor** | Parser |
| **Actor’s Goal** | המרת הלוג מדף HTML לפורמט CSV. |
| **Preconditions** | שליפת לוג UC-3, עדכון הגדרות 2-UC, תזמון הParser UC-9. |
| **Flow of Events fo Main Success Scenario** | |
| 1. The Parser parse the input log into suitable object to store, and from there to a CSV format. | |

|  |  |
| --- | --- |
| **Use Case UC-5: ELK הכנסה ל** | |
| **Related Requirments** | REQ05 |
| **Initiating Actor** | Parser |
| **Actor’s Goal** | ביצוע תקשורת עם Elastic Search לצורך העברת הלוג בפורמט CSV אליו. |
| **Preconditions** | המרת הלוג UC-4, תזמון ה Parser UC-9 |
| **Flow of Events fo Main Success Scenario** | |
| 1. The Parser succefully inserts the CSV files to ELK 2. The ELK get the data process it and analyze with the help of the tool called elastic search. | |

|  |  |
| --- | --- |
| **Use Case UC-6: ניתוח מידע** | |
| **Related Requirments** | REQ09 |
| **Initiating Actor** | ElasticSearch |
| **Actor’s Goal** | עיבוד הנתונים כלומר מבצעת חיפוש וניתוח (אנליזה) של הנתונים שהתקבלו כקבצי CSV |
| **Preconditions** | העברת קבצי CSV ל elastic search UC-5. |
| **Flow of Events fo Main Success Scenario** | |
| 1. The Parser succefully inserts the CSV files to ELK 2. The ELK get the data process it and analyze. | |

|  |  |
| --- | --- |
| **Use Case UC-7: Visualizing** | |
| **Related Requirments** | REQ06 |
| **Initiating Actor** | KIBANA |
| **Actor’s Goal** | מציגה את המידע על גבי dashboard מאפשרת ויזואליזציה של המידע. |
| **Preconditions** | ניתוח המידע UC-4 |
| **Flow of Events fo Main Success Scenario** | |
| 1. The Kibana get the analyzed data from ElasticSearch. 2. Kibana visualize the analyzed data in different ways. | |

|  |  |
| --- | --- |
| **Use Case UC-8: כתיבה לקובץ ההגדרות** | |
| **Related Requirments** | REQ02, REQ07 |
| **Initiating Actor** | משתמש המערכת |
| **Actor’s Goal** | כתיבה לתוך קובץ ההגדרות את ההגדרות לפיהם המשתמש מעוניין לפרסר את קבצי הלוג שלו. |
| **Preconditions** | None |
| **Flow of Events fo Main Success Scenario** | |
| 1. The system user write the configurations he wants in the json configuration file that comes with the system. | |

|  |  |
| --- | --- |
| **Use Case UC-9: Parser תזמון ה** | |
| **Related Requirments** | REQ03, REQ05, REQ06, REQ04 |
| **Initiating Actor** | שעון המערכת |
| **Actor’s Goal** | מתזמנת את ה Parser לפי זמני המערכת שהוגדרו בדרישות שמצריכות תזמון. |
| **Preconditions** | None |
| **Flow of Events fo Main Success Scenario** | |
| 1. The system clock scheduling the Parser. | |

**2.3 Traceability Matrix - With respect to: Requirements 🡸🡺 Use Cases**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **REQS** | **PW** | | UC01 | | UC02 | | UC03 | | UC04 | | UC05 | | UC06 | | UC07 | | UC08 | | UC09 | |
| REQ01 | 2 | | X | |  | |  | |  | |  | |  | |  | |  | |  | |
| REQ02 | 2 | |  | | X | |  | |  | |  | |  | |  | | X | |  | |
| REQ03 | 1 | |  | |  | |  | |  | |  | |  | |  | |  | | X | |
| REQ04 | 3 | | X | | X | | X | |  | |  | |  | |  | |  | | X | |
| REQ05 | 2 | |  | |  | |  | | X | | X | |  | |  | |  | | X | |
| REQ06 | 2 | |  | |  | |  | |  | |  | |  | | X | |  | | X | |
| REQ07 | 3 | |  | |  | |  | |  | |  | |  | |  | | X | |  | |
| REQ08 | 4 | |  | | X | | X | |  | |  | |  | |  | |  | |  | |
| REQ09 | 1 | |  | |  | |  | |  | |  | | X | |  | |  | |  | |
| **Max PW** | | 3 | | 4 | | 4 | | 2 | | 2 | | 1 | | 2 | | 3 | | 3 | |
| **Total PW** | | 5 | | 9 | | 7 | | 2 | | 2 | | 1 | | 2 | | 5 | | 8 | |

Diagram

Description automatically generated**2.4 High-Level SW–Architecture Diagram**

**3. Data Design**

**3.1 ERD – Entity Relation Diagram**

ההתחברות נעשית למסד נתונים קיים של המפעל\החברה המשתמשים בשירות

**4. Object Design**

**4.1 Dynamic Object Modeling**

Table

Description automatically generated

**4.2 Static Object Modeling**

**UML Class Diagram**

כפי שצויין קבצי לוג במערכת שלנו מוגדרים להיות קבצים מסוג html. מהותם במערכת זאת, היא להגיע לאותו אוביקט מסוג log parser, אוביקט זה יפרסר את הלוגים לפי ההגדרות שהוגדרו לו ע"י אוביקט מסוג configuration json שבו יתבצע ניתוח ההגדרות מקובץ ההגדרות של המערכת הנתון בפורמט json. שאר המחלקות מספקות איחסון נוח באוביקטים מתאימים על מנת להשתמש במידע לצורכי הפרסינג. מידע כמו מאיזה מכונה הגיע הלוג, מה שם הלוג, מה התוכן שלו, וכדומה..

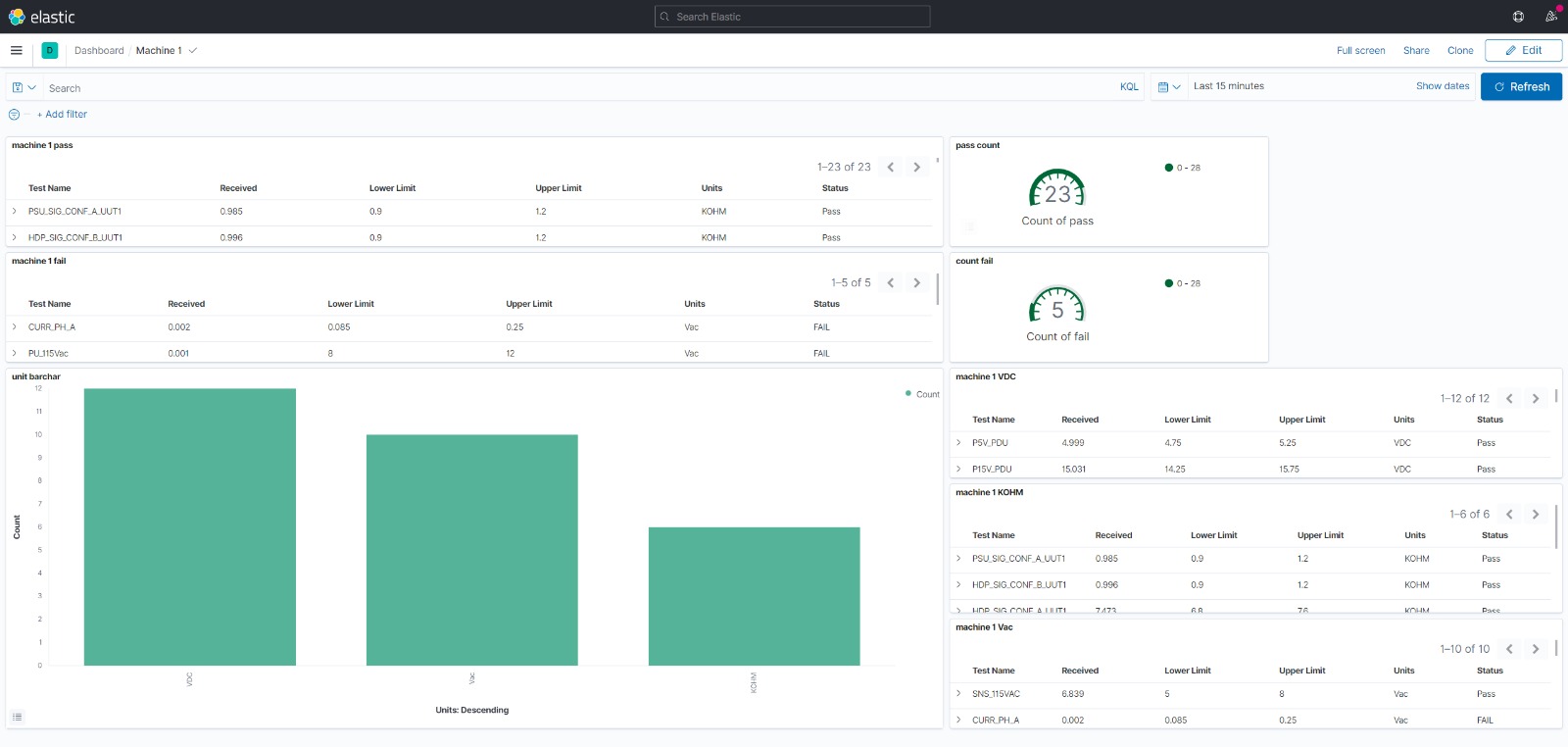
Diagram

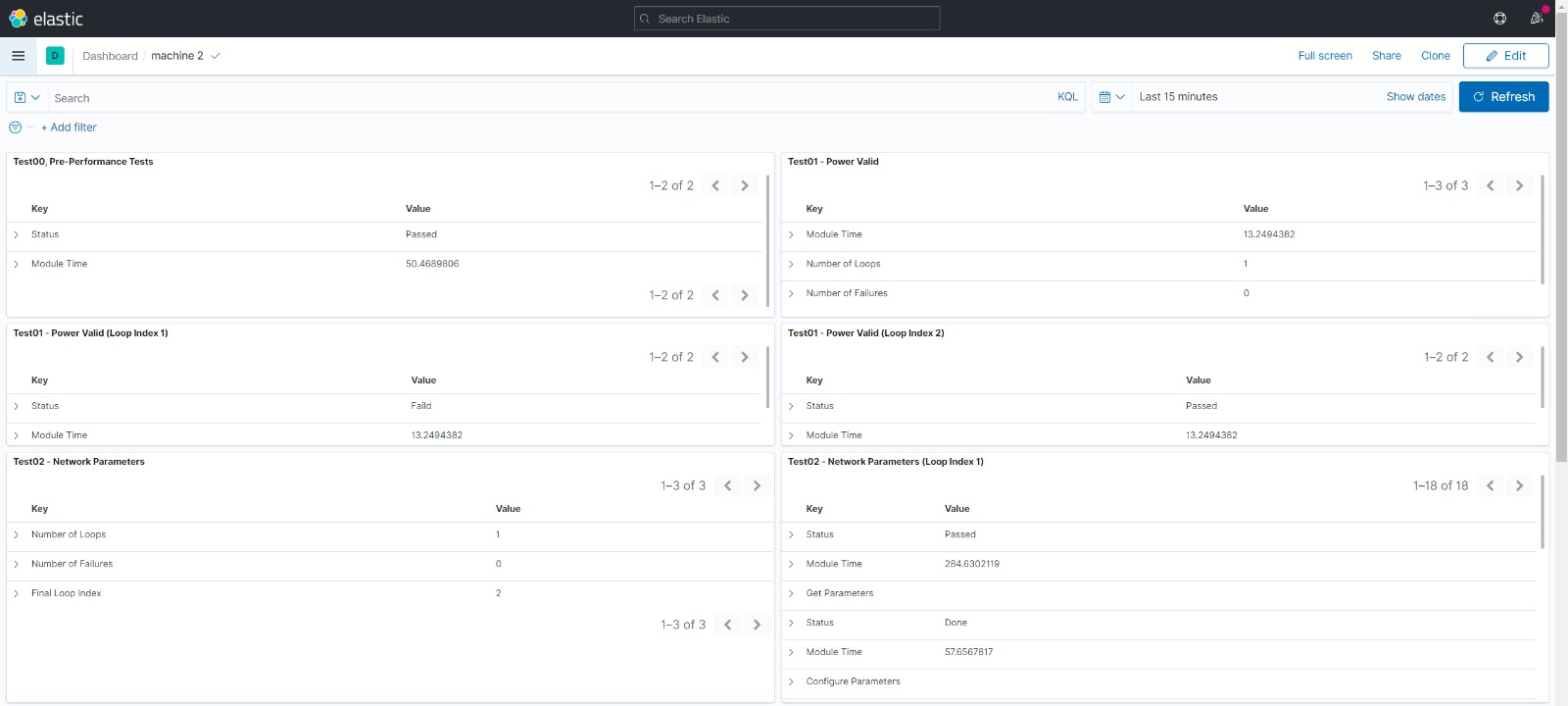
Description automatically generated

**5. User Interface Design**

ממשק המשתמש יהיה ממשק הגרפי של שרת הKibana. 

הממשק יציג את ניתוח הנתונים בdashboards כאשר בכל אחד ניתן יהיה לראות מידע המתקבל מהמכונות השונות.





**6. Refined Project Plan**

