

Software Engineering 2

A.Y. 2019/2020

**SafeStreets**

DD – Design Document

Version 1.0

**Authors:**

Daniele Comi 10528029 - 944534  
Anton Ghobryal 10501942 - 945577

**Professor:**

Elisabetta Di Nitto

Sommario

[1. Introduction 3](#_Toc25681242)

[1.1 Purpose 3](#_Toc25681243)

[1.2 Scope 3](#_Toc25681244)

[1.2 Definitions, Acronyms, Abbreviations 5](#_Toc25681245)

[1.2.1 Definitions 5](#_Toc25681246)

[1.2.2 Acronyms 6](#_Toc25681247)

[1.2.2 Abbreviations 6](#_Toc25681248)

[1.4 Revision history 6](#_Toc25681249)

[1.5 Reference Documents 6](#_Toc25681250)

[1.6 Document Structure 7](#_Toc25681251)

[2. Architectural design 8](#_Toc25681252)

[2.1 Overview 8](#_Toc25681253)

[2.2 Component view 12](#_Toc25681254)

[2.3 Deployment view 14](#_Toc25681255)

[2.4 Runtime view 16](#_Toc25681256)

[2.5 Component interfaces 16](#_Toc25681257)

[2.6 Selected architectural styles and patterns 22](#_Toc25681258)

[2.7 Other design decisions 27](#_Toc25681259)

[3. User interface design 32](#_Toc25681260)

[4. Requirements traceability 35](#_Toc25681261)

[5. Implementation, integration and test plan 38](#_Toc25681262)

[6. Effort spent 38](#_Toc25681263)

[7. References 38](#_Toc25681264)

# Introduction

## Purpose

The purpose of this document is to give more technical and design details than the RASD about SafeStreets application system. RASD showed a general view of the system and described the functions which supposed to be executed, while DD aims to present the implementation of the system including design viewpoints; such as context viewpoint, composition viewpoint, logical viewpoint, dependency viewpoint, information viewpoint, patterns use viewpoint, interface viewpoint, structure viewpoint, interaction viewpoint, state dynamics viewpoint, and finally resource viewpoint. It also presents in more details the implementation and integration plan, as well as the testing plan. More precisely, the document presents:

* + Overview of the high-level architecture
  + The main components and their interfaces provided one for another
  + The Runtime behaviour
  + The design patterns
  + Additional details about UI
  + Mapping requirements on the defined components
  + Implementation plan
  + Integration plan
  + Testing plan

The purpose of this document is to provide an overall guidance to the architecture of the software product.

## 1.2 Scope

Here’s a review of the application’s scope that refers to what is already explained in the RASD.

SafeStreets allow users to report a violation to the authorities when they spot one. In order to obtain the ability of using SafeStreets the user will have to register himself into the application system. Users have two different modes to register themselves into the system: the first one is the proprietary authentication which also requires email validation and the second one consists of SPID9 authentication. Generally, they will have to subscribe with their full name and fiscal code since they’re mandatory to be able to fill certain reports. Registered Users obtain points that indicate their integrity through their continuous voluntary participation in order to provide the possibility of achieving the goal of making the streets safer. These points are called integrity points. Initially, users, who have registered with SPID9, have more integrity points than the proprietary authentication (according to demonstrating more integrity into the society verifying his own identity through a public system of digital authentication). Moreover, when a report is verified by the authorities, integrity points of the notifier increase. Users can see also, through a map, the security level of a zone. Allowing users to mine general information about notified violations doesn’t violate the privacy of the reporting user according to the Legislative Decree 196/031 and the regulation 2016/6793 given since they aren’t authorized to access other users’ private information such as fiscal code, name, surname etc. Security level is calculated being based on the statistics of the types of violations committed in the interested area. Of course, any user will have the possibility to change the password in case it is forgotten through the normal process of password change link sent to their email address.

As it is in the specification of the S2B, Reports are composed of date, time, position, a note (with a maximum fixed number of characters) and a clear picture of the committed violation in which the licence plate should be included, but it isn’t a restricted requirement because, in the worst case, there are two possible situations: in the first one the licence plate isn’t clear (e.g. poor quality or blurry image) the user is allowed to do one out of two possible actions that consist of re-take the picture of the violation or modify the licence plate number, and if the user chooses to do the second action, the system shall recognize the report as one, instead, with a modified licence plate number and this induces minor level of credibility; instead, in the second situation, if the system doesn’t recognize a vehicle in the taken picture it will take an immediate action to discard this picture and it will eventually ask the user to take a new clearer picture to be able to proceed, and that precludes the fact that user might send pictures that are not in accordance with the domain of the application (e.g. photos that don’t contain a vehicle such as selfies).

Since the violation must be notified in real-time domain, the user is not allowed to upload a picture at all. So that, situations as creating a false violation or manipulating data of a certain violation. For the same reason the user is not allowed to modify a photo. If the user notices something that should be mentioned, there’s a note that he can fill in briefly with possible observations. Also, the user must have a stable active connection to be able to submit the violation.

A report should satisfy the application domain before it becomes in hands of authorities and in order to realize this fact a report should include the preconditions described earlier. When a report is filled in completely the authorities must be able to receive it through the application. Within this context, the authorities are defined as Italy’s law enforcement agencies. The authorities, interested in the application willing to use it for increasing local security, must have a valid digital certificate provided by the police forces through the Ministry of the Interior5 and the Ministry of the Defense6. An authority must register making a formal and direct request to the Certified Email7,8 address of SafeStreets through his Certified Email7,8 which will give him in a secure way the credentials generated to be able to use the application or by just using SPID with his Certified Email. The login process with the authority credentials requires also a valid digital certificate. Once an authority is registered and SafeStreets has added his credentials in the system, he will be able to receive notifications about the committed violations. Registered authorities have the maximum authorization to access all the data notified by users. They also have access to all normal user functionalities, thus the capability of reporting violations. The authorities can also verify and validate the visualized reports depending on the veracity of the notified violations. The authorities are also guaranteed a second access to SafeStreets through a web service which require the same login process.

Either the registration process or the reports made and of the user who carried it out are respects the terms established by the Legislative Decree 196/031, Legislative Decree 82/052 and the regulation 2016/6793.

SafeStreets offers also the possibility to be an important participant as an independent entity which can provide suggestions to the improvement of a certain area. In order to realize such a functionality, SafeStreets should have access to accident records of the applied areas. Interested municipalities, in order to let the authorities benefit this functionality, must guarantee access to those data records because it helps the application to cross the provided data about accidents with its own data to provide suitable suggestions depending on the identified situation. It will then notify the authorities regarding those suggestions.

## Definitions, Acronyms, Abbreviations

### Definitions

* **Violation**: a subset of anything that is classified as a traffic violation by the Traffic regulation and laws document. This subset is composed of:
* Double line parking
* Expiry of the parking time limit
* No parking area
* Parking in places reserved to people with disabilities
* Parking in the middle of bike lanes
* Parking near bus stops
* Parking on crosswalk
* Parking on residents reserved spots
* Parking ticket missing
* Possible vehicles damage by third parties (e.g. broken glass)
* **Vehicle**: any terrestrial identifiable vehicle subject to Traffic regulation and laws document, like cars, motorbikes, trucks, etc…
* **User**: any citizen registered in the system who is using any of SafeStreets functionalities.
* **Violation report, notification**: acknowledgment in SafeStreets system of a new violation occurred.
* **Authority**: any registered law enforcement using SafeStreets application alongside its authority-restricted functionalities
* **Municipality**: any central administration of a city or a town which may or may not give open access to its incidents data.
* **Reliability score**: score assigned to any user account which gives a sense of how much a user I reliable in giving information regarding violations.
* **Safe area**: a low radius geographical area where violations are lower than a certain threshold or lower than other areas.
* **Suggestion**: an automatically inferred hint given to the authorities by SafeStreets regarding how they could improve, with the help and permission of their municipality, area marked as high-risk area due to a high correlation of violations and incidents reported from the same municipality. Possible suggestions are:
  + Add a barrier between the bike lane and the part of the road for motorized vehicles
  + Install a towaway zone sign
  + Increase parking slots
  + Increase local police controls
* **Galileo**: Global localization system based on a network of 24 satellites commissioned by European Union and ESA (European Space Agency)
* **SPID**: is the unique system of access with digital identity to the online services of the Italian public administration and of private members: citizens and companies can access services with a unique digital identity in a secured way
* **Certified Email**: A certified email is an email that can only be sent using a special Certified Email Account provided by a registered provider. When a certified email is sent, the sender's provider will release a receipt of the successful (or failed) transaction. This receipt has legal value and it includes precise information about the time the certified email was sent.  A certified email account can only handle certified email and can't be used to send regular email.
* **Design viewpoint:** The specification of the elements and conventions available for constructing and using a design view.

### 1.2.2 Acronyms

* **API**: Application Programming Interface
* **AWS**: Amazon Web Services
* **D.L.**: Legislative Decree
* **DCPM**: Decree of the President of the Council of Ministers of the Italian Republic
* **DD**: Design Document
* **DDoS**: Distributed Denial of Service
* **DMV**: Department of Motor Vehicles
* **DMZ**: DeMilitarized Zone
* **DNS**: Domain Name System
* **EEA**: European Economic Area
* **EU**: European Union
* **GDPR**: General Data Protection Regulation
* **GPS**: Global Positioning System
* **IEEE**: Institute of Electrical and Electronics Engineers
* **QL**: Query Language
* **S2B**: Software to Be
* **SaaS:** Software as a Service
* **SPID**: Public Digital Identity System
* **UI**: User Interface

### Abbreviations

• Gn = nth goal

• Dn = nth domain assumption

• Rn = nth requirement

## 1.4 Revision history

* + Version 1.0: first release

## 1.5 Reference Documents

* + D.L. 196 of 2003 (196/03) <https://www.camera.it/parlam/leggi/deleghe/Testi/03196dl.htm>
  + D.L. 82 of 2005 (82/05) <https://docs.italia.it/italia/piano-triennale-ict/codice-amministrazione-digitale-docs/it/v2017-12-13/index.html>
  + General Data Protection Regulation (EU) 2016/679 <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:32016R0679>
  + IEEE 1016-2009 - IEEE Standard for Information Technology--Systems Design--Software Design Descriptions <https://standards.ieee.org/standard/1016-2009.html>
  + Specification document “Mandatory Project Assignment AY 2018-2019” <https://polimi365-my.sharepoint.com/:b:/g/personal/10528029_polimi_it/EXR1gN6gBoxJgMC86Ow45gMBFwZzkRSWuoaf5K7t1wZutA?e=SPnVkI>
  + Ministry of the Interior and digital certificates released <http://politichepersonale.interno.it/itaindex.php?IdMat=1&IdSot=35&IdNot=386>
  + Ministry of the Defence and digital certificates released <http://www.pkiff.difesa.it/#secEN>
  + Certified Email <https://www.agid.gov.it/it/piattaforme/posta-elettronica-certificata>
  + Certified Email RFC <https://tools.ietf.org/html/rfc6109>
  + SPID <https://www.agid.gov.it/it/piattaforme/spid>
  + Italian license plate verifier <http://www.targa.co.it/data/doc.aspx>
  + Police State license plate verifier <https://www.crimnet.dcpc.interno.gov.it/crimnet/ricerca-targhe-telai-rubati-smarriti/FAQ>
  + RASD
  + AWS <https://docs.aws.amazon.com/>
  + Cloudflare <https://developers.cloudflare.com/docs/>
  + Google Firebase <https://firebase.google.com/docs>
  + AES <https://nvlpubs.nist.gov/nistpubs/FIPS/NIST.FIPS.197.pdf>
  + RSA <https://community.rsa.com/docs/DOC-60094>
  + Argon2id <https://www.cryptolux.org/images/0/0d/Argon2.pdf>

## 1.6 Document Structure

* **Chapter 1** is an introduction to the design document. Its goal is to explain the purpose of the document and to highlight the differences with the RASD, whilst showing the link between them.
* **Chapter 2** aims to provide a description of the architecture design of the system, it is the core section of the document. More precisely, this section is divided in the following parts:
  + Overview
  + Component view
  + Deployment view
  + Runtime view
  + Component interfaces
  + Selected architectural styles and patterns
  + Other design decisions
* **Chapter 3** specifies the user interface design. This part is already contained in the RASD in the mockups’ section. However, we decided to insert some UX diagrams to better describe the interaction between the customer and the application.
* **Chapter 4** provides the requirements traceability, namely how the requirements identified in the RASD are linked to the design elements defined in this document.
* **Chapter 5** includes the description of the implementation plan, the integration plan and the testing plan, specifying how all these phases are thought to be executed.
* **Chapter 6** shows the effort which each member of the group spent working on the project

# Architectural design

## Overview

The SafeStreets system to be developed is a distributed application following the multitier architecture paradigm with a completely scalable multitier and data tier as shown in the Figure 1. The architecture is basically composed, without going in details yet, with three main layers which are the Presentation Tier, the Middle tier and the Data Tier. The Presentation Tier is the layer near the user where information is presented, and the user can start or receive interaction with SafeStreets. The middle tier, as will be better explained later, is the layer managing all the application and business logic behind SafeStreets coordinating all its functionalities. Finally, the Data Tier is the layer which its purpose is data storage of the SafeStreets system. So, all these different hardware layers that represent different computers and servers needed to do the respective tier work. This architecture will grant the system the characteristics of high scalability and high flexibility as it will be shown later.



*Figure 1 – High level multitier architecture*

Now, going in more details as shown in the Figure 2, a general diagram but not yet precise in every part of the architecture we can see starting from the Presentation Tier that is represented by the users using the application which can be Authorities or Normal users. As shown Normal Users will only be able to access SafeStreets through a mobile application where the Authorities will also have the possibility to use a SaaS through web access. Their requests towards SafeStreets will be asynchronously or synchronously depending on which action they will take in account. After sending a violation report they will be able of course to use any other kind of services offered by SafeSteets, as it will happen when for example using the Safeness are map functionality. Instead when composing a violation report in the process of taking the violation picture, they will have to wait, so a synchronous interaction, SafeStreets to control the validity of the image before proceeding in any further composing of the violation report. Also, interaction starting from SafeStreets to users are asynchronous, when for example sending to the selected Authorities notification on new reported violation by some users. This notification interaction will be better described later.

Moving on the middle tier as shown in the Figure 2 is highly scalable and flexible. Application servers and web servers are also decoupled to acquire even more stability and more security which will be later discussed. Even if the middle tier shows for clarity reasons only three distributed aggregated nodes in the diagram, it will be able to have many more of them. Each distributed node represents indeed a set of server geographically dislocated with different IP addresses, referencing the third level of the TCP/IP stack, but associated to the same DNS records. DNS, not part of the architecture, will be able to choose which node is better to initiate communication to. But this is not all, all main nodes containing various servers will have a load balancer to allow the workload to be correctly distributed without overwhelming any server during heavy workloads. Servers will run NGINX instances which will be discussed later, where different kind of services or even replicated services can work in parallel: a parallelized system will allow SafeStreets to get the requested Availability level for each functionality allowing it to have a high fault resistance characteristic.

The middle tier will have also access to external services provided by third parties to allow SafeStreets functionality to work correctly. As shown, we can spot the Google Firebase service, the SPID service, the License Plate API, the Police State services, the CA/TA Authorities for digital certificates, the Maps services and the Municipalities incident data access. All these various third parties will offer specific APIs used in the middle tier to access their services.

It is also present a Cloud architecture instantiated remotely on Amazon Web Services which will be used to work with heavier workloads regarding the image validity verification, license plate OCR system and the suggestions functionality of SafeStreets. As shown this deployed architecture to be defined during its instantiation will have an elastic load balancer which will grant to deploy any new hardware needed to keep the workload under control granting SafeStreets to function normally even when there are lots of data or requests regarding these functionalities. From the diagram can be spotted various Amazon SageMaker deployed instances like for Convolutional Neural Networks, for Natural Language Processing, for OCR or Bayesian Networks: their need will be explained in the further sub chapters.

Security is a topic even in the architectural design, where later on will be discussed security in software and network terms, in the architecture to better secure SafeStreets WAN there are present various Hardware Firewalls between Internet accessed by clients and the first layer they will encounter which is the web servers which are responsible for exchange information even if not in a direct browsable format for every users as explained before. This second hardware firewall is between this previous sub layer of the middle tier and the distributed application server layer: this will create a DMZ, Demilitarized Zone, for the application servers so that the external network can access only to the resources exposed in the DMZ.

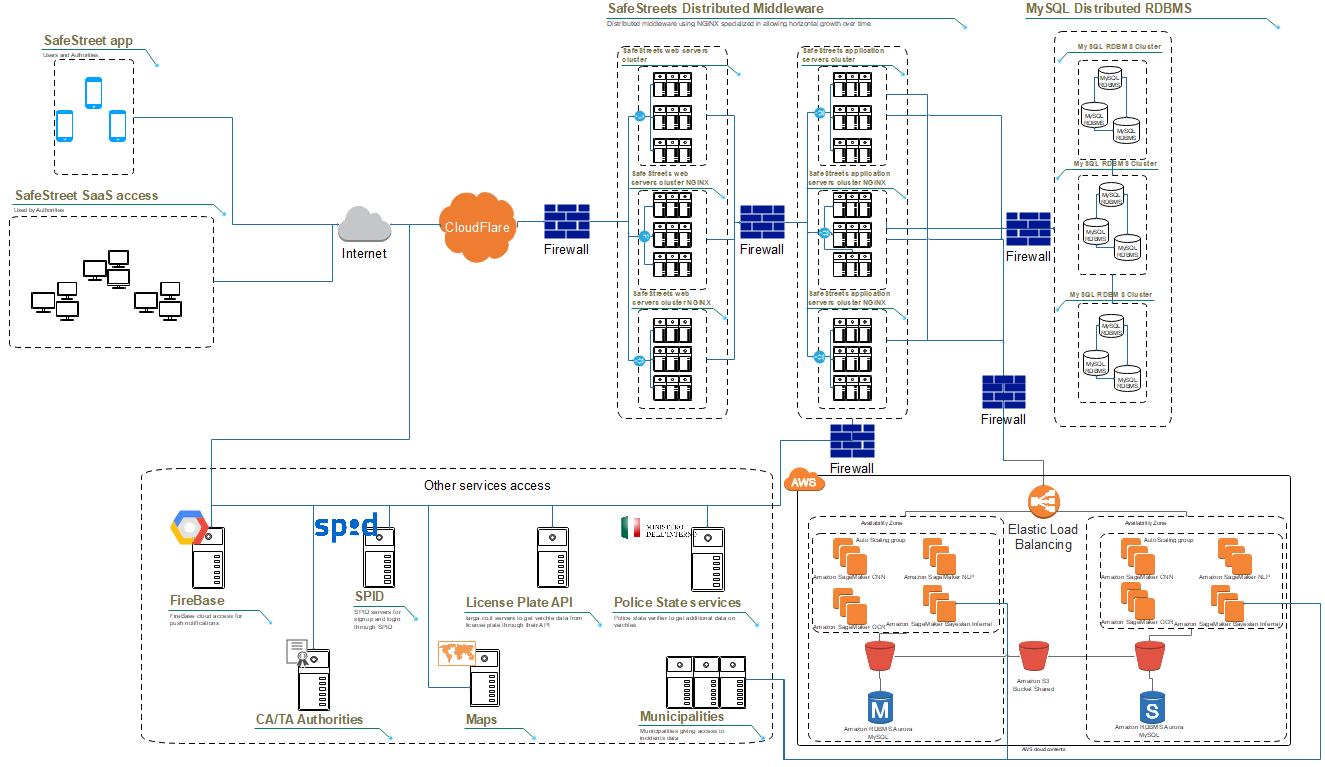
The web servers are not guaranteed the same level of security because their functionality is the management of information representation and their forwarding or receiving to/from users, and upon functionalities requests they will forward the requests to the application servers in the DMZ in a distributed way letting the DNS and the load balancer of the distributed application servers choose the most suitable node. This level of security is required, as mentioned, since the offered service deals with sensitive data of the users. The lower level of security of the more exposed layer of the distributed web servers is covered by the used service of Cloudflare.

Cloudflare can solve this issue and add also more services like DDoS protection, a Web application firewall, an Authoritative DNS and a Content delivery network. Their specific details will be discussed later.

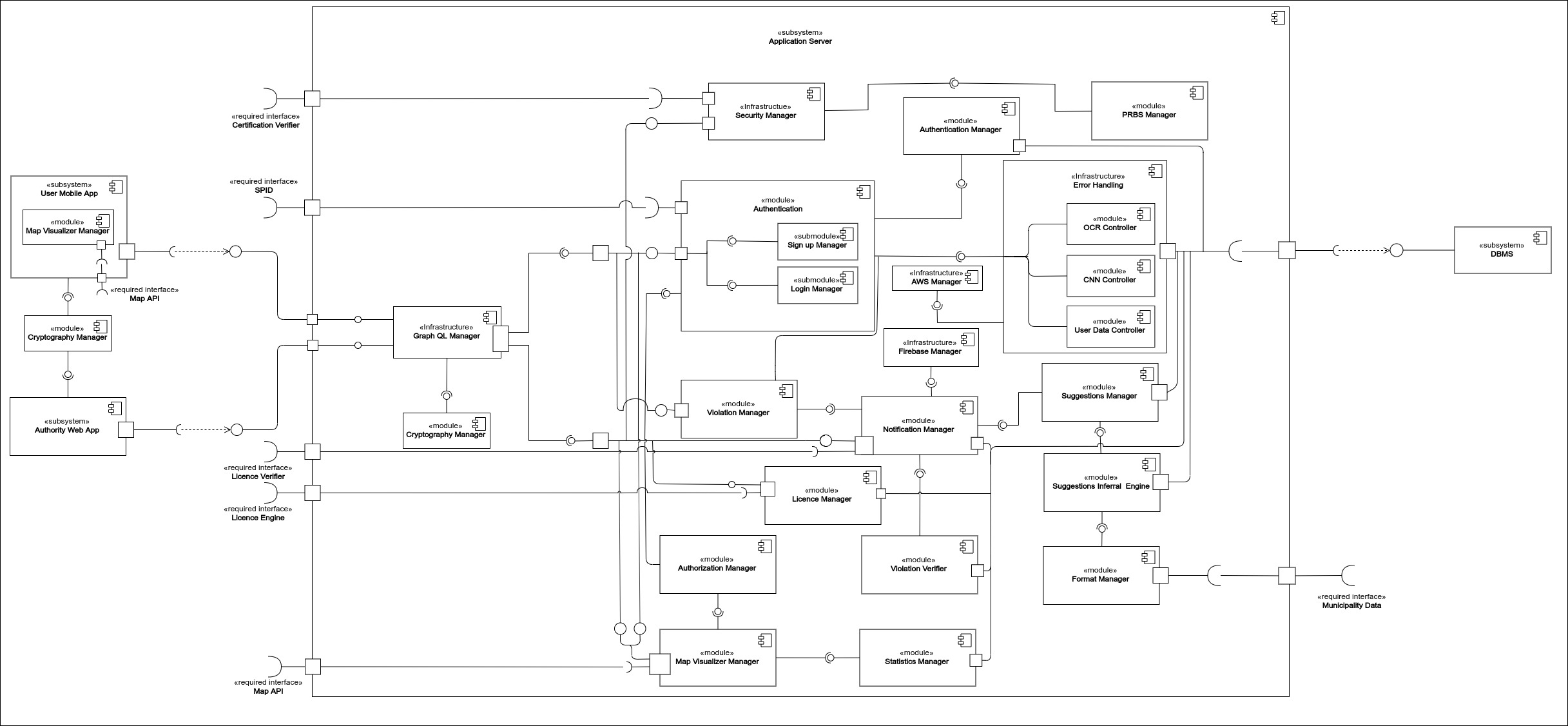
On the last layer called Data Tier we can find a distributed RDBMS, Relation Database Management System, based on MySQL which will grant a geographically advantageous access to data depending on user request and on the middle tier requests. Their specific details will be discussed later but it’s important to say that a hardware firewall between each layer of the multitier architecture is needed to get more security as possible. Indeed, even between the second sublayer of the middle tier and data tier there is a need of the firewall to better protect what’s, already secured, inside database: user and SafeStreets data.

Of course, to make the middle tier correctly communicate, there is the need to adopt a middleware. Where clients communicate through one of the distributed web servers, the application servers need to communicate with various heterogeneous distributed systems and remote third party services. So, a middleware is present, and it’ll be using some already existing frameworks in order to have a correct communication between the distributed web servers, the private cloud and the distributed RDBMS to the middle tier. To grant better performances the various network nodes associated with the distributed web servers but also with the distributed application servers will be better indexed through the best routing and switching algorithm available in terms of software defined networking of SafeStreets. To speed up the finding of the various nodes available it will be used, and configured, at a lower level in the ISO/OSI (TCP/IP) stack than the presentation level, an OpenFlow switch.

*Figure 2 – General system architecture of SafeStreets*



## Component view



* **User Mobile App**
* **Authority Web App**
* **Map Visualizer Manager**
* **Cryptography Manager**
* **Graph QL Manager**
* **Security Manager `**
* **PRBS Manager**
* **Authentication**
* **Authentication Manager**
* **Violation Manager**
* **Error Handling**
* **Authorization Manager**
* **Statistics Manager**
* **Violation Verifier**
* **Licence Manager**
* **Format Manager**
* **Suggestions Inferral Engine**
* **Suggestions Manager**
* **Notification Manager**
* **AWS Manager**
* **Firebase Manager**

## Deployment view



*Figure 3 – Deployment diagram*

The deployment diagram shows how is intended the physical deployment of artifacts on nodes.

In the Figure is shown the deployment diagram of SafeStreets, note that it is not displayed anything which is not under control at the design level so, for example, no deployment node of third-party services or of the private cloud.

Starting from the first tier where clients are located we can see two kind of deployment are available as already known.

The computer access will be deployed using an internet browser running on any hardware platform x86 or ARM available. It can be so generalized on different hardware deploying platforms because of its kind of interaction only through a web browser.

The users in general, comprehending the authorities, instead will have to access SafeStreets through a smartphone application. SafeStreets app will be deployed as executable using as operating system any Android ARM x64, ARM x32, x86 or x64 compatible hardware running Android 6.0 Marshmallow API 23 or higher. The Android application will be better explained in its details in the Chapter 5 of Implementation, integration and test plan. What can be already said is that it will use the Android SDK available with last APIs. It will have an Android activity for the main screen and others sections will be under another activity or an Android fragment. Any other functionalities needed to run into background or upon notifications will use Android services and Android broadcast receivers.

Moving to the middle tier the first device indicated as node in the deployment diagram it’s the external firewall. It’ll be deployed using a Cisco Firepower 8360 running on its IOS allowing a bandwidth of 60 Gbps. This will grant the best security in terms of unauthorized access to the network of the various servers. This very deployment will be the same in the various firewalls.

Both web servers and application servers will be running their various executable services of SafeStreets on Linux Red Hat Enterprise x64 operating system in an HPE ProLiant DL580 Gen10 running multiple scalable Intel Xeon 9200 series with up to 12 TB of DDR4 RAM and 512 GB of static RAM. The web server will be running NGINX natively or on VMware server instances to guarantee the required performance levels thanks to, first the various virtualized instances which reduces drastically the cost of failures and increment drastically the availability of the services and secondly thanks to NGINX internal structures using various kind of workers. The communication between web servers to application servers and from application to the, till to be analyzed, data layer deployment diagram will be done using already existing frameworks such as Remote Procedure Calls protocols using Java EE RMI, the Java EE RDMI and the Java EE Persistence APIs to communicate with an object/relation mapping fashioned way to the RDBMS. To gain ulterior enhance on performances of the middle tier, an OpenFlow controller switch to better manage internal cluster packets exchange through various switches which are controlled by the SDN defined in the controller which will use a Cisco Catalyst 8350.

The third layer corresponding to the data tier has various kind of nodes corresponding to the various distributed RDBMS. As it will be stated they’ll be running Oracle DB as RDBMS to guarantee very high performances. Indeed, the RDBMS has been chosen because of the high relational level of the tables for which data finds a secure storage. Data such as every proprietary authentication encrypted data, or all the violations and so on like it will be displayed, in the further pages, by the ER model. A RDBMS is also needed due to its really powerful skills of being very good in optimizing queries, especially the ones handling tons of data. It will use their algebraic representation together with internal statistics distributions of the data to make them faster and simpler to be executed, on an already fast DBMS. SQL indexes will also be set using specific main and secondary memory structures to guarantee the best performance available. The RDBMS, which is coherent with respect to the ACID properties, will also have various triggers to guarantee the perfect balance of data consistency and business logic implemented in the RDBMS to lighten the work on the middle tier when possible.

The various hardware available in the second and third tier for the web and application server and for the distributed RDBMS servers will have enough computational power, main and mass memory to handle everything. As computational power will be used HPE Primera 670 Storage based on Intel Xeon 9200 CPUs with about 6 PB of storage divided up between very high quality SSDs and HDDs, with also 1 TB of dedicated cache and 32 Gbps maximum throughput. Every Database Server will have a level 6 RAID system.

What was not already explained was the load balancer. Its instance will be deployed on a third party system because of the adoption of a hardware load balancer solution like for the various firewalls. It will be used the Barracuda Load Balancer 842 ADC which is the best suitable for large systems which can have very high traffic requests to handle in the fastest way possible having the maximum theoretical throughput at a rate of 20 Gbps.

## Runtime view

## Component interfaces

Here the SafeStreets class diagram with almost every details is shown, almost every details because further implementation analysis is required. The class diagram is shown firstly in full form to give a sense of the general architecture, although due to being big it was better to shown even some cut and more zoomed parts.



*Figure 4 – Full Class Diagram*

Now, for a better understanding zoomed parts are shown. Starting from the lower level corresponding to the Data tier, in the Figure 5 is shown how the interaction with the distributed ORDBMS, Oracle DB, is managed within a cluster. The available API for a better entity-object management are also used which are part of the Enterprise Framework available.



*Figure 5 – Class Diagram of Data tier*

The next part is now focused in the middle tier which due to being big was split in various parts.

In the Figure 6 it is shown the communication and security package. In the communication package are present the various classes for communicating according to the architecture through HTTPS, TLS and other types of secure communication which are using a standard dynamic scheduler for the various connections. The logical representation of the communication between users and the system is represented using GraphQL which will be better presented in the following paragraphs.

The visitor pattern in a little more abstract and general way is shown and how indeed communication needs to pass through the implemented DynamicRouter using the MessageDispatcher class.

In the security package the various symmetric and asymmetric encryption algorithm used are present as classes and also other security functionality functions regarding digital certificates X509 and hashing are present.

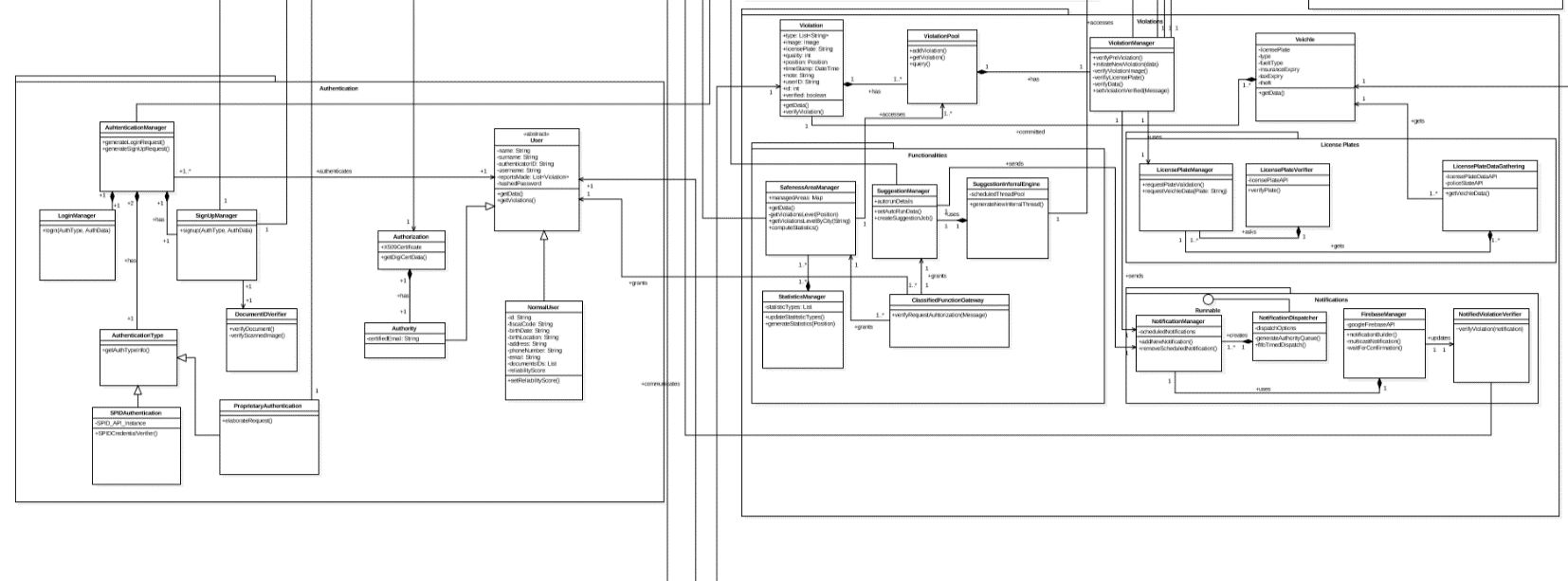


*Figure 6 – Class Diagram of middle tier – part 1*

In the following 7th figure the authentication part of the system and also the one regarding the various functionalities are shown. As shown in the RASD the authentication is composed of two types, one proprietary and one through SPID. The user class and its children are present as a form of entity-representing object through the previous shown Enterprise framework API.

The functionalities are accessed through a constantly present layer of security to also grant or not what a user can actually get in response and also which authorities receives a notification automatically regarding some particular event through the Google Firebase push notifications API.

Both authentication and functionalities access is ruled under the main dynamic visitor pattern.



*Figure 7 – Class Diagram of middle tier – part 2*

In the last part of the middle tier shown in the figure 8 it is shown the interaction with the private cloud through AWS and with Municipality access.

AWS SageMaker grants SafeStreets access through the APIs which are then used in the main AWSManager class which is the node through which any kind of AWS requests and responses are managed. Requests can also be automonous inferral on new data, a better fine tuning of hyperparameters, an improving of the already presents model and so on. As shown, in case it is asked to verify some data like a picture taken that will be asked to the CNNController which will then start a request to AWS through the AWSManager. The same is shown for the OCRController.

For the municipalities instead it is shown as an architecture a pool of data access for the various municipality which contains various metadata on the municipality itself and then the incidents data queried from their server, which specifications is of course present.



*Figure 8 – Class Diagram of middle tier – part 3*

The last part of the system is leaved to the presentation tier shown in the figure 9 where there is the standard client side network communication software architecture and then then a package for security and for actual presentation is present. The security package is needed also here to grant an actual end-to-end communication. In the GUI package are present the various Activities, Fragments, services and APIs used in the client side with a basic inner representation which is left for the further implementing phase part.

*Figure 9 – Class Diagram of presentation tier*



## 2.6 Selected architectural styles and patterns

Here are presented the architectural styles and patterns used in designing SafeStreets, their presentation will start from the architectural patterns, then to the architecture styles. That is because there is not a unique pattern and style but a variety of these to allow the design and future development of a better application both on end user level for satisfaction and on the developer level regarding the implementation, the power of the architectures combined and the accordingly easily maintainable and testable project.

As for Architectural Patterns for SafeStreets have been chosen these two main patterns:

* Multitier Architecture
* Dynamic Visitor Messagging Pattern

#### Multitier Architecture

The Multitier Architecture is intended to allow any of the n-tiers to be upgraded or replaced independently in response to changes in requirements or technology. For example, a change of operating system in the presentation tier would only affect the user interface code because they are physically separated.

The main presentation tier is the topmost level of the application. It is intended to display information related to users request upon various SafeStreets functionalities like Safeness area map, violations report, and statistics on violation in a certain area and so on. It communicates with other tiers and it is a layer which users can access directly using the SafeStreets app, or in case of Authorities using also the web access.

The application tier which corresponds to the distributed middleware running on different clusters in SafeStreets, is the tier that controls an application’s functionality by performing detailed processing regarding various functionalities, in this case it will be able to obtain data and communicate with other services and infrastructure as discussed before.

The data tier includes the data persistence mechanisms and the data access layer that encapsulates the persistence mechanisms and exposes the data. In this case the data tier is also distributed to grant even more scalability, flexibility and general performance improvements.  
  
A graphical general and generic representation of the Multitier Architecture is here displayed in the Figure 10.



*Figure 10 – General and generic multitier architecture*

#### Visitor Pattern

The visitor pattern is a way of handling inner communication in the middletier from ingoing and outgoing messages and requestes. This pattern is very powerful indeed it will be able to developers to add new functionalities without changing the main classes of SafeStreets, thanks to its separation between objects and algorithms. This approach will make the project more and more:

* Maintainable
* Testable
* Easier approach to test driven development (see Implementation and testing in chapter 5)

An important advantage is also that the various visitors are easy to separate and assign to different developers during the actual implementation phase.

The visitor pattern can be also better understood by looking at the Figure 11.

**

*Figure 11 – General and generic Dynamic Visitor Pattern*

As for Architectural Styles for SafeStreets have been chosen these three main styles:

* Web app - service oriented architecture style
* Private cloud computing architecture style
* Representational state style using GraphQL

#### Web app - Service oriented

The Service oriented architecture style (SOA) is a style of software design where services are provided to the other components by application components, through a communication protocol over a network. The basic principles of service-oriented architecture are independent of vendors, products and technologies. A service is a discrete unit of functionality that can be accessed remotely and acted upon and updated independently.

A service has four properties according to one of many definitions of SOA:

* It logically represents a business activity with a specified outcome.
* It is self-contained.
* It is a black box for its consumers.
* It may consist of other underlying services.

Service-orientation promotes loose coupling between services which is what we want in general for SafeStreets and what will be having especially in the SafeStreets web access for Authorities. SOA will separate various SafeStreets functionalities for Authorities into distinct services for example the service for checking the various kind of violations in multiple areas or the service for the suggestion inferred through municipalities data which can be dynamic over the time. This choice of using a web app for the authorities accessing SafeStreets it is a better approach because it will grant to deploy to different kinds of authorities the same access and will help do decouple the deployment of the various services for the authorities. The service-oriented architecture is well suitable in the SafeStreets architecture thanks to its large scalability, flexibility and performance characteristics.

#### Private Cloud computing

The second style adopted is the private cloud computing architectural style. Private because it will have a direct link with SafeStreets given the fact that is used to check for image validity, license plate reading and suggestions computation with municipality data. It will be based on AWS on which there will deployed various SageMaker instances:

* Convolutional Neural Networks: it will be used to recognize if an image is valid, so if it contains a vehicle and its identifiable license plate using a pretrained model with millions of tested images and having a 99% success ratio. Each instance will be improved upon any new image upload. The chosen instance would be the ml.p3.16xlarge.
* Natural Language Processing: it will be used to add more context and to aggregate in a better way incident data with their description which municipalities give SafeStreets access to. It will create so a model tree of an incident when possible to improve the suggestions capabilities and to do so it will be used the BlazingText Algorithm in AWS.
* OCR: it will be used to read autonomously the license plate using text recognizing algorithms based on continuous learning and improvement.
* Bayesian Networks: it will be used to create suggestions, it’s based on a conditional probabilistic model represented as a probability graph where it is possible to get the more probable event following a certain event conditioning it. It will be allocated in the Machine Learning section of AWS SageMaker.

AWS private cloud is composed of an elastic load balancing which will automatically instances new deployable instances for SafeStreets functionalities depending on the requests. SageMaker will be configured in using the best hardware available for Depp Learning which depending on the actual economical availability can be the NVIDIA DGX-2 or some NVIDIA Tesla V100 GPUs.

#### Representational state pattern using GraphQL

As for the third style, Representational state pattern using GraphQL it was chosen this composite style because in SafeStreets there will be a huge number of users that will access its functionalities and on which no server will have to save any kind of client state, which also does not have any control over. Although a REST standard query style has the disadvantages to be lower performant than GraphQL, to don’t have the possibility querying multiple different resources in an efficient way and to be server dependent on the query result representation. This does not mean that the REST architectural style does not need to be used: it guarantees, through its constraints, various advantages. So, it was chosen to combine it with GraphQL which compensate all those various disadvantages said above like performance, multiple different resource queries, client defined which led to a series of advantages which are useful to the actual SafeStreets system. Given the fact that Authorities and normal users will have different level of authorized access it is effective to be capable to make requests in a specific form defined from the caller of the query and receiving the result in that precise format. But this is not the end: it is also very suitable for relational data because GraphQL will by itself get the right data put together without the need of the developers to do various REST calls to different endpoints. Having also different kind of access methods for Authorities and normal users, a GraphQL approach will allow to simply adapt the specific request for visualizing data on the Service oriented style for Authorities using the SaaS and on the mobile app for both users and authorities and it will be able to better manage of push notification through Firebase. Also, adopting GraphQL will led to better maintainability because there is no need of a common interface for requests.

Having an underlying REST approach will create constraints which are already satisfied:

* Uniform interface: the goal is to have a common approach to access the resources, so that being familiar with one API means being familiar with all the other APIs.
* Client - Server: client and server are two different entities, which evolve separately without any dependency.
* Stateless: the client is responsible for managing the state of the application and this entails a simpler server design.
* Cacheable: this allows to avoid some interactions between the client and the server, speeding up the communication.
* Layered system: A client cannot ordinarily tell whether it is connected directly to the end server, or to an intermediary along the way. This means that the client doesn't know if it's talking with an intermediate or the actual server. So, if a proxy or load balancer is placed between the client and server, it wouldn't affect their communications and there wouldn't be necessities to update the client or server code.

All the communication in question exploit the HTTPS protocol through POST requests, HTTP over SSL, in order to guarantee the security and the reliability of the connection in any kind of situation. Regarding the format in which the data are transmitted, JSON is used because it is suitable for the data interchange between client and server and GraphQL services typically respond using JSON. JSON may seem like an odd choice for an API layer promising better network performance, however because it is mostly text, it compresses exceptionally well with GZIP. A graphical simple explanation of its working process is shown in the below Figure 6.

The utilization of this approach will ensure high future improvements, maintainability, and flexibility and will ensure a lower testing effort required to reach the desired quality. This is possible thanks to the improved of the already high decouple between server and client implementation and architecture. Any modeling change on the server side will not effect in any way the various client requests nor their security.



*Figure 12 – GraphQL, one of the advantages displayed*

## 2.7 Other design decisions

There are also other design decisions which were briefly shown in the overview of this   
chapter:

* Distributed RDBMS using MySQL
* Thin client
* Cloudflare
* Firebase for push notifications
* End-to-end encrypted communication
* End-to-end encryption for sensitive data
* Network architecture overview

#### Distributed RDBMS

The distributed RDBS is actually a pretty good choice for SafeStreets. Having already a distributed middle tier it was quite natural to couple it with a distributed RDBMS because it will allow to gain performance on data retrieval and their actual availability. A relational database was chosen given the data structures presents in SafeStreets which needs to be related with one another, for example a violation will be obviously related to the user who made it, to a vehicle, to certain city and so on. For the RDBMS it was chosen Oracle DB EE that offers one of the most performant RDBMS available, supporting ACID properties and a variety of SQL functionalities like triggers and data integrity upon tables creation and indexes.

In the following Figure 6 is displayed the ER diagram for the Oracle DB EE Distributed RDBMS.



*Figure 13 – ER diagram*

As shown in the figure 7 there are multiple entities with various kind of relationship and respective cardinalities. There are some peculiar points which merit to be explained.  
The password field is limited to 32 characters because as will be stated below in the following sections, only its hash will be saved for security and privacy reasons, accordingly to the Legislative Decree 196/03, Legislative Decree 82/05 and the regulation 2016/679, and the hash function that will be used has as outputs hashes of length of 256 bits which is translated to 32 chars. The image data will be saved as LONGBLOB type (Long Binary Large Object) to ensure that any image with any kind of quality or compression will be accordingly saved. The note attribute is limited as discussed in the requirements for clarity reasons regarding user explanation to exactly 140 characters. Vehicle entity has lots of data thanks to the license plate APIs and the Police State APIs, those data will be made visible to authorities and in a limited part to normal users. The authority will be as stated in the previous document capable of reporting a violation as a normal user can but also to verify violations which were notified to him. The verification process will output an enumeration which is represented as an integer in the attribute so that’s possible to state if no verification has been done or if there was a successful verification or a failed verification which will then change the reliability score for the reporting user through the internal business logic partially implemented as MySQL triggers.

#### Thin client

The thin client decision was straightforward. In SafeStreets all main computation is made on the middle tier so however a user is accessing safestreets, mobile app or SaaS, they won’t need to do any kind of heavy workload but the actual presentation of information and its rendering. However, some needed tasks are required in order to guarantee the desired high degree of security, like for example encryption/decryption especially in the end-to-end encryption for sensitive data explained below.

The client application, differently from the web app the authorities can access, will take advantage of the framework available for the local environment in which is being executed and it will use it to guarantee it can correctly communicate with the various servers and to be sure it will be able to do the few important operations requested.

#### Cloudflare

Cloudflare was adopted to improve principally the web servers security but also to add more services like DDoS protection, a Web application firewall, an Authoritative DNS and a Content delivery network which are not so easy to add natively in a complex architecture.

* DDoS protection: Cloudflare offers an "I'm Under Attack" mode for customers experiencing cyberattacks. It will be needed in case SafeStreets were under attack from malicious users making millions of requests very rapidly to try to disrupt SafeStreets services and other user experiencing their SafeStreets access.
* Web application firewall: Cloudflare will allow to utilize a web application firewall service. By default, the firewall has the OWASP ModSecurity Core Rule Set alongside Cloudflare's own ruleset and rulesets for popular web applications which will mitigate any non-compliant incoming packet.
* Authoritative DNS: Cloudflare will also offer a faster user accessing SafeStreets services by using their DNS which has one of the fastest DNS lookup speeds worldwide, with a reported lookup speed of 5.6ms.
* Content delivery network: Given that Cloudflare's network has the highest number of connections to Internet exchange points of any network worldwide it is possible to exploit Cloudflare cache content in its edge locations to use it as a content delivery network (CDN); all web requests from Authorities through the web access SaaS are then reverse proxied through Cloudflare with cached content served directly from Cloudflare when needed. So, in case of web server malfunction, which is very rare given their highly scalability characteristics and fault tolerance, no web access for the Authorities will be compromised.

#### Google Firebase

Google firebase is instead used for push notifications, push notifications are needed for the authorities to know whenever a new violation is reported, or a new suggestion is available. Push notifications are managed in the middleware but then it will be Firebase to send them to the right chosen users by the servers. A user will retrieve push notifications using even here, as described before, GraphQL. Developers will use Firebase available API, and this will also lower the testing and implementation cost together with simple yet powerful management.

The utilization of this cloud service will ensure high future improvements, maintainability and will ensure a lower testing effort required to reach the desired quality.

#### End-to-end encrypted communication

Every communication will be encrypted as said before, in order to do that it will be chosen an encryption based on asymmetric and symmetric keys. RSA with 4096 bits keys will be used alongside with AES 256, having in SafeStreets safely stored a KeyStore encrypted with a user dependent key, which contains the private key used together with the public key of SafeStreets to exchange a symmetric key for AES 256, gaining better performances and to allow secure communication on any network. This will make SafeStreets compliant with the Legislative Decree 196/03, Legislative Decree 82/05 and the regulation 2016/679.

#### End-to-end encryption for sensitive data

For sensitive data, so for the user personal data upon registration there will be the need of a further security layer. To do so it will be used end to end encryption, where not even SafeStreets will be able to read user personal and private data. To accomplish this will be used a symmetric encryption algorithm, which its key will be based on the actual password of the user salted with a function which has as domain a set of characters that are the unique identifier of a user. This end to end encryption works because the key is not known even to SafeStreets since in the database the password will be saved with its hash, using as hash function Argon2id which is the most powerful and resistant over attacks hash available. So, any user personal data is encrypted in the database using this procedure and sent over the network as it is, allowing over the actual owner to decrypt it as requested the Legislative Decree 196/03, Legislative Decree 82/05 and the regulation 2016/679.

#### Network architecture overview

In the following image is shown the already discussed architecture but shown in a more network fashion way so to be helpful in the future implementation at the network level for understating better the clustering and the configuration needed for the hardware network components such as the firewall, the load balancer and the OpenFlow switch defining an SDN.



*Figure 14 – Network architecture diagram overview*

# User interface design

The following design of the smartphone app of SafeStreets for both users and authorities and the design for the desktop application exclusively available for authority are here shown. They are just an intuitive way to lead the design towards this direction but any necessary changes or adjustments or new components can be added to ensure a better user experience and improve both usability and simplicity. UX diagrams will follow to understand user interaction with SafeStreets app and the authorities interaction even with the SaaS web app.

*Figure 15 – Login Figure 16 – Home screen*

*Figure 17 – Camera for violation picture Figure 18 – Violation report UI*

*Figure 19 – Statistics Figure 20 – Safeness Map*

*Figure 21 – User account details Figure 22– Inferred suggestions to authorities*

*Figure 23 – Violations reported to authorities Figure 24 – Authorities home screen*



*Figure 25 – exclusive authority SafeStreets desktop access through secure SaaS web-app*

# Requirements traceability

Design choices are not just for the sake of the various advantages and improvements already shown but also to grant the fulfilment of the project which was first described in the RASD, especially the goals and the various requirements.

Here is shown the complete list of the various requirements with the various associated components previously described which will grant their feasibility.

* [R1] The user must be registered to use the application.
* [R2] The user can register, and access securely, through two different authentication methods: SPID9 and proprietary authentication.
* [R3] The user registered with SPID9 has a higher initial integrity score than a registered user with proprietary authentication.
* [R4] Each user has an integrity score.
* [R5] Each user can access the details of his own and view his data, integrity score and reports made.
* [R6] Each registration made by a user follows the indications imposed by the Legislative Decree 196/031 and the Regulation 2016/6793, which are shown to the user.
* [R7] Each authority can securely access the application through his pre-given credentials and its digital certificate provided by the police forces through the Ministry of the Interior5 and the Ministry of the Defense6.
* [R8] Each authority can have access to the application features available for users without privileged access.
* [R9] Each authority has full access to the reports made.
* [R10] Each authority has full access on vehicle information pinpointed by the external Italian license plate verifier and by State Police portal.
* [R11] Each authority can access the details of the report made and of the user who carried it out according to the terms established by the Legislative Decree 196/031 and the regulation 2016/6793.
* [R12] Any authority shall receive notifications regarding the municipality in his current position.
* [R13] A report must consist of an image, date, time, location and metadata.
* [R14] The metadata of a report is the type of report, the quality of the report and the notes entered by the user.
* [R15] The notes entered by the user cannot be longer than 140 characters.
* [R16] Date, time and location must be added automatically via the Internet and GPS/Galileo satellites.
* [R17] The user can proceed with the violation signalling if the GPS location is precise and obtained securely.
* [R18] It is possible to report in the presence of an Internet connection only.
* [R19] User reporting image is recognized as valid for reporting only if it contains a vehicle that can be identified through the license plate.
* [R20] The system must be able to recognize the vehicle registration number.
* [R21] The user can decide to modify the result of the reading of the license plate made by the system.
* [R22] The system includes for each report a warning if the user has modified the vehicle license plate in which the notification will have reported a lower quality.
* [R23] The user can’t send multiple times the same violation report.
* [R24] Each user can access a map showing the security level in certain areas.
* [R25] Each user can have limited access to reports by viewing information that does not violate the privacy of the reporting user according to the Legislative Decree 196/031 and the regulation 2016/6793.
* [R26] Each user can view statistics based on reports made in certain areas.
* [R27] Authorities can indicate an alert as verified through the application.
* [R28] Each alert verified by an authority will give the user who has indicated it a higher reliability score.
* [R29] The system must be able to access the accident data present in a specific municipal area if present.
* [R30] The system must analyse accidents and violations data to produce a suggestion to be notified to the authority to improve road safety.
* [R31] Any authority shall receive suggestions regarding the municipality in his current position.

# Implementation, integration and test plan

# Effort spent

# References