|  |
| --- |
| **Macintosh HD:Users:festor:Desktop:Logos Lorraine INP:LOGO LINP-  Couleur.gifuniv-lorraine** |
| Master’s Thesis |
| Development of a cloud-based Configuration Management Database |
|  |
| David Rubino |
|  |

Year 2014-2015

Final year internship conducted in Prime Resources

to obtain the engineering degree of TELECOM Nancy

Internship supervisor: Ritch Houdek  
University supervisor: Moufida Maimour

**Déclaration sur l’honneur de non-plagiat**

**Je soussigné(e),**

**Nom, prénom : Rubino, David**

**Élève-ingénieur(e) régulièrement inscrit(e) en 3e année à TELECOM Nancy**

**N° de carte d’étudiant(e) : 31010226**

**Année universitaire : 2014 - 2015**

**Auteur(e) du document, mémoire, rapport ou code informatique intitulé :**

**Development of a cloud-based Configuration Management Database**

Par la présente, je déclare m’être informé(e) sur les différentes formes de plagiat existantes et sur les techniques et normes de citation et référence.

Je déclare en outre que le travail rendu est un travail original, issu de ma réflexion personnelle, et qu’il a été rédigé entièrement par mes soins. J’affirme n’avoir ni contrefait, ni falsifié, ni copié tout ou partie de l’œuvre d’autrui, en particulier texte ou code informatique, dans le but de me l’accaparer.

Je certifie donc que toutes formulations, idées, recherches, raisonnements, analyses, programmes, schémas ou autre créations, figurant dans le document et empruntés à un tiers, sont clairement signalés comme tels, selon les usages en vigueur.

Je suis conscient(e) que le fait de ne pas citer une source ou de ne pas la citer clairement et complètement est constitutif de plagiat, que le plagiat est considéré comme une faute grave au sein de l’Université, et qu’en cas de manquement aux règles en la matière, j’encourrais des poursuites non seulement devant la commission de discipline de l’établissement mais également devant les tribunaux de la République Française.

**Fait à Milwaukee, le 08/ 31/ 2015**

**Signature:**

|  |
| --- |
| Master’s Thesis |
| Development of cloud-bases Configuration Management Database |
|  |
| **David Rubino** |
|  |

**Year 2014-2015**

Final year internship conducted in Prime Resources

to obtain the engineering degree of TELECOM Nancy

David Rubino  
W170 N11525 Armada Dr. Germanton, WI 53022   
262-302-9049  
[david.m.rubino@gmail.com](mailto:david.m.rubino@gmail.com)

TELECOM Nancy  
193 avenue Paul Muller,  
CS 90172, VILLERS-LES-NANCY  
+33 (0)3 83 68 26 00  
[contact@telecomnancy.eu](mailto:contact@telecomnancy.eu)



Prime Resources LLC  
N58 W24793 Quail Run Ln Sussex, WI 53089  
262-465-6750

Internship supervisor: Ritch Houdek  
University supervisor: Moufida Maimour

Acknowledgments

This project would not have been possible without the support of many people.

First and foremost, I wish to express my deepest gratitude to Ritch Houdek, my industrial supervisor, who followed me during the entire project. With his numerous suggestions and ideas, he was able to help me pursue the development of the application. Ritch was always there when I needed help, and I am deeply grateful to him.

Then, I would like to thank John Houdek, director of Prime Resources, for giving me the opportunity to do my internship in his company. I also thank him for his help with all the paperwork involved, especially for getting the J-1 visa in order to come to the United States.

I would like to show my gratitude to Kasandra Wysocki and Colin Brady, concierges at the Hudson Business Center, for the help they provided me as well as their everyday good mood who made each day of work a great experience.

I also would like to thank all the other Hudson members I was able to meet, for their advice, good mood, and uplifting spirit who helped me understand the spirit of American entrepreneurship.

I thank my university supervisor, Moufida Maimour, for following my progress during my internship and correcting this report.

Last but not least, I thank all members of the jury for reviewing my work and giving me the opportunity to present the results of this great professional experience that my internship was.

Summary

[Acknowledgments 3](#_Toc429663972)

[Summary 4](#_Toc429663973)

[1 Introduction 6](#_Toc429663974)

[2 Presentation of the company 8](#_Toc429663975)

[3 Description of the project 10](#_Toc429663976)

[4 State of the art 13](#_Toc429663977)

[4.1 ITIL 13](#_Toc429663978)

[4.2 Configuration Management Database (CMDB) 14](#_Toc429663979)

[5 Problem analysis and our solution: a cloud-based CMDB 17](#_Toc429663980)

[5.1 From a user point of view 17](#_Toc429663981)

[5.2 From a developer point of view 19](#_Toc429663982)

[5.3 Why a minimally viable product 19](#_Toc429663983)

[5.4 Answering the flaws of enterprise CMDBs 21](#_Toc429663984)

[5.5 The importance of cloud-based applications 22](#_Toc429663985)

[6 Technologies used 26](#_Toc429663986)

[6.1 Programming languages 26](#_Toc429663987)

[6.2 XAMPP 26](#_Toc429663988)

[6.3 phpMyAdmin 27](#_Toc429663989)

[6.4 jsTree 27](#_Toc429663990)

[6.5 Vis.js 27](#_Toc429663991)

[6.6 Aptana Studio 28](#_Toc429663992)

[7 Implementation of the solution 29](#_Toc429663993)

[7.1 Data model 29](#_Toc429663994)

[7.2 User administration 33](#_Toc429663995)

[7.3 Interaction with the database 37](#_Toc429663996)

[7.3.1 Connection to the database 37](#_Toc429663997)

[7.3.2 User session 38](#_Toc429663998)

[7.3.3 Retrieving data with Ajax 39](#_Toc429663999)

[7.3.4 Updating data with Ajax 41](#_Toc429664000)

[8 Conclusion 43](#_Toc429664001)

[9 References 44](#_Toc429664002)

[List of figures 47](#_Toc429664003)

[Glossary 48](#_Toc429664004)

[Addenda 49](#_Toc429664005)

[Résumé 52](#_Toc429664006)

[Abstract 52](#_Toc429664007)

# **Introduction**

This CMDB project took place for my last year internship at TELECOM Nancy to get my Master’s Degree in Computer Science. It was directed by a small startup company from Wisconsin, Prime Resources.

The main goal of the internship was to apply on a specific project all the skills and knowledge I learned during my training as a graduate student. Though challenging, it helped me develop new skills as a software developer, but also helped me to improve my analysis and conception skills as well.

The task that was given to me was to develop a Configuration Management Database (more commonly referred to as CMDB for the rest of the report) application aimed at small and medium-sized companies. The objective was mainly to write something simple, user-friendly, and efficient in terms of rendering. The notion of Minimally Viable Product (MVP) was essential to keep in mind during the whole development of the application.

The goal was to create a CMDB that users would want to use. Without obviously competing with major CMDBs available on the market such as OneCMDB, our application was aimed for people who would like to access quickly to the different dependencies of their system, and who would feel the need to use it without spending too much time on it. Thus, the software was designed as a cloud application for an easy access, whether on a phone, tablet, or PC. The advantage of not having to download and install something on the computer or phone was essential for us, as it improved the attractiveness of the product and the time that a user would spend to configure the application.

A CMDB consists mainly in listing all the different resources that a company has (servers, databases…) in order for a specific user to know which of these resources are used by a specific department or network. It comes really handy when an administrator would like to upgrade one of his servers, but first would want to make sure that the different dependencies linked to the server would not create a version problem for a specific app linked to that server. Thus, it enables not only administrators, but every single user to know exactly how the company resources are implemented, and to have an idea of which parts would be impacted by a specific change.

The problem about CMDBs though, is that it can soon become too heavy to maintain up to date. For example, big retail companies likes Kohl’s or Macy’s require heavy CMDB software in order to support their vast network spreading across the country. It becomes soon a big charge to make sure the CMDB is up-to-date and running with the latest versions of the different components of the company. Administrators do not always have time to update the software, and users can get discouraged to use it since it does not really reflect the state of the company at present time. CMDBs can then become neglected, which is a financial loss since thousands of dollars are usually invested in the licensing.

Obviously, small and medium-sized companies would have no interest in investing such a big amount of money in products that are complicated to use and that would require too much time to keep up-to-date. Therefore, it is very unlikely that they would invest in such a product. Although CMDBs on the market are apparently complex, the idea of knowing exactly what is going on in the company with all the different resources is a really important, if not a necessity measure every company should strive to have. We believe that such companies are looking for a simple and easy way to manage their network, without having to spend too much time learning how to do it and actually doing it.

That is where the subject of my internship comes up. My objective was to write a CMDB specifically aimed at such companies. It should be easy to use by any employees, no matter what their function could be, and easily maintainable as well. To present that subject into more details, I am going to present the company in which I did my internship, Prime Resources. I will then describe the project into more details. Then, we will take a look at the different solutions available on the market. Next, I will analyze the problem and describe the solution that was implemented before exposing the different arguments of why we chose this solution and how we are sure that it is going to work. We will then study into more details the implementation of the solution before talking about the economic strategy behind it. Finally, after describing the future potential of the application, I will expose the different management methods used during the project.

# Presentation of the company

In this part, I will present the company in which I had the opportunity to do my internship.

Created in 2014, Prime Resources LLC is a startup company focusing on providing top-notch services to other companies. Its main goal is to recruit people having the skills necessary to answer one of the clients’ most challenging problems. The different domains of expertise of the company do not only include information technology but also manufacturing, executive search, and engineering. [1]

Although the company was created very recently, its story is interesting to tell as it illustrates the spirit of American entrepreneurship that I was able to experience a lot during my internship. In 2014, John and Karen Houdek, residents of Sussex, WI, created their own company, Prime Resources. Owned by Mrs. Houdek and led by her husband, it was their desire, after spending 25 years in the industry leadership, to run their own business. The company has grown since then and answers the challenges of many other businesses who come to Prime Resources in order to find the best professionals available to solve their problems.

The company is registered as an LLC (Limited Liability Company). According to the IRS (Internal Revenue Service), an LLC company is “a business structure allowed by state statute. Each state may use different regulations, and you should check with your state if you are interested in starting a Limited Liability Company.” [2]

This legal structure is the most convenient to adopt for an entrepreneur as the registration process at state level is pretty fast and do not include a lot of constraints. An owner of an LLC is called a member, and there is no limit on the number of members that can be included in an LLC. Actually, a single person can also be the sole member of his LLC, and different entities can be members as well, such as other companies for example.

The type of activities a company can do as an LLC usually depends on state law, but most of the time bank and insurance companies cannot claim this status. On the other hand, this business structure is very popular for IT services start-ups.

Regarding the different work methods, Prime Resources will usually place their employees at a client’s location. Most of time, people will actually not work from the headquarters but from each client’s site. In my case, it was a little different. Since I was working on a project that was not for a specific client but for a later release on the market as a software solution, I had the opportunity to work from different locations throughout my internship.

Although I worked from the headquarters, I also had the opportunity to work from a business lounge located in downtown Milwaukee: the Hudson Business Lounge. The Hudson is a meeting, work, and event center located in the heart of Milwaukee’s Third Ward, one of the most ancient part of the city. The purpose of this location is to provide start-ups a work environment as well as meet other like-minded professionals. Offering over 11,000 square feet of office space, it is a futurist concept that give people the opportunity to grow their network, either with meeting new clients or potential employers/employees. [3]

I was able to meet new people working on different IT projects. Some were just developers like me who worked on a specific project for their company, others were managers and recruiters who hold conferences and events in order to broaden people’s mind on the business world.

Working from this location influenced me deeply. It helped me understand better the world of start-up companies and the spirit of entrepreneurship that is so specific to the United States. Later on in this report, I will detail the influence I received from working at the Hudson; but before that I will present the different objectives of the problem I had to solve.

# Description of the project

The goal of my internship at Prime Resources was to develop a simple version of a Configuration Management Database, commonly referred to as CMDB. A CMDB is a software that allows businesses to keep track of their IT infrastructure to know exactly what resources are available to the company, whether it be hardware like servers, software like applications, or even human resources like the number of database administrators.

In today’s market, companies can use a CMDB in order to maintain up-to-date their different resources. While not all businesses can invest time and money in such a tool, most big companies like Kohl’s, General Motors, or Walmart will usually invest in a CMDB so they can have a better view of all resources available in each department of the company. Although this practice is well-defined in companies’ strategic development, the use of CMDBs can rapidly become overwhelming and too much time-demanding, thus making employees less inclined to update them, or even to use them at all.

Experience proved to my manager, currently a vice-president at Kohl’s Corporate, that people never use the CMDBs because of their complexity and their difficulty of access. A first problem is the time required to learn how to use a CMDB. Although different systems are available on the market, they all require a lot of training to know how to research or update a specific item in the database. Most of the time, employees only need to know about one specific function, but the complexity of the system makes it impossible to speed up the learning process or to specialize the training in a specific area of the software.

A second problem that comes to mind (and actually resulting from the first one) is the inaccuracies caused by the data in the CMDB. When the structure of the company changes (for example when a new device is added to the network or when a specific department is reworked), it becomes necessary to take into account these changes into the CMDB. They need to be registered in the software by either an administrator or a qualified employee who owns the right to modify the data according to his own department.

Those changes being numerous, it can be pretty hard to keep track of them, which ultimately leads to some inaccuracies in the CMDB. Employees will then not feel they can rely on a database that could potentially be out-of-date. Thus CMDBs become seldom updated and used.

Although bigger companies can find time and money to invest in getting a CMDB and training some employees to keep it updated, small- and medium-sized companies do not have this luxury. They usually have to use simpler methods to access the different resources of their company. The principles of CMDBs could however be applied to any company, the size not being a decisive factor for a CMDB to be efficient. It is clear, though, that with the current products available on the market, those companies would not be likely to invest time and money in purchasing a license for one of these applications.

This is where the project takes its place. What if there was a simple CMDB on the market that would meet all requirements regarding what a small- or medium-sized company could expect from a CMDB? Such an application should, of course, be simple to use and very intuitive, while enjoying the perks of having a top-notch interface while using some of the most recent technologies available on the market. Such an application could be easily implemented if the conception and development always keep in mind those three aspects: Minimally Viable Product.

My task was then to develop a software that could answer most flaws described earlier. The application would be divided into four parts:

* Configuration items listing
* Applications listing
* Data center mapping
* Human Resources inventory

The first part would rend the list of all configuration items from the company. A configuration item, as we will see later in the state of the art, is either a virtual element in the company like a server, a database, or a virtual storage service. The application should give the user the possibility to view all configuration items of the company, as well as their properties. An administrator should also be able to add new items or properties to existing configuration items. Moreover, the interface should display these items in a hierarchical tree.

The second listing should present the different applications used in the company. Each application can reference to one or multiple configuration items. Thus, in this listing, there will be two parts: first, a file explorer tree that shows the hierarchy of the applications grouped into folders; second, a graph that will list all configuration items used by a specific application. As with the previous part, there will be two views: user and administrator. The user should be able to access all applications and visualize the tree. The administrator should be able to organize the different applications into folders and be able to rename, add, or delete applications. He also can add configuration items to a specific application and reorganize them into different folders.

The third part is the most ambitious in the whole project. It should present the company’s different datacenters. As with the two previous points, a hierarchical tree will list all the data centers. While clicking on a specific node, the user will access to a map of the data center that shows the position of all the elements included in the room. If the user has the corresponding rights, he can modify the map by editing the room’s characteristics (such as the number of rows and columns and the position of the different servers). He can also add or delete servers. Now, when a user clicks on a server, he will be redirected to a 3D view of that server. This view will show him the different configuration items that are present in the server. By clicking on it, the user will be redirected to the properties of that configuration item, as shown in the first part.

The fourth and last part will present an inventory of the different human resources available in the company. They will be organized according to a hierarchical tree into folders. An administrator can have the possibility to add and remove folders. These folders will represent the different departments of the company, and will allow a listing of all employees and functions in the company. A property view similar to the one from the first part will allow users to access the different data of an employee (like the hiring year) or a department (like the number of UNIX developers). Like before, an administrator will be able to modify all these information while a simple user will be read-only.

The application will also feature other minor functionalities like a login and sign up page, a user account information page, and a search bar to allow the user to look for specific properties or items more easily.

Now that we have seen the problem the application is trying to solve, as well as the different requirements that it should have, we will first focus on the work that was done before regarding CMDBs as well as the different software available on the market.

# State of the art

Configuration Management Databases, commonly referred to as CMDBs, are part of the ITIL (Information Technology Infrastructure Library) framework. Before describing what the purpose of a CMDB is, it is important to understand what ITIL is and what its specifications are in order to better understand the big picture.

## ITIL

The Information Technology Infrastructure Library framework, more commonly referred to as ITIL framework, consists of different tools and practices to “help individuals and organizations use IT to realize business change, transformation, and growth.” [4] More specifically, it focuses on providing different requirements in order to help IT services to focus on business needs.

ITIL consists of a succession of procedures and tasks to be completed by the IT department of each company in order to master the model described in a set of books. The first requirements of this kind appeared in the 1980s in the UK, before being released into a major version in 2000 as ITIL v2. The current version in use as today is ITIL v3, with an update which came in July 2011.

The advantage of following this model is to help companies achieve their goals in a more efficient and less costly way, as well as to grow and develop the size and budget of the company throughout the years. To achieve this goal, 5 key points are incorporated into the ITIL specifications:

* Strategy management for IT services
* Service portfolio management
* Financial management for IT services
* Demand management
* Business relationship management

Although understanding the specific role of each process is not necessary in the continuation of this report, it is nonetheless interesting to note where CMDBs are positioned in this larger scale.

The IT services strategy management includes 2 core points: service support and service delivery. CMDBs are part of the first one. This first service focuses on the user; it makes sure that he, whether a final consumer or a developer in the company, can have a positive experience and will be able to ask information, request update, or notify of an issue. The CMDB should be able to process any request and, if it fails, notify a third party which will analyze the problem and likely propose a solution.

## Configuration Management Database (CMDB)

Part of the ITIL process chain, a Configuration Management Database, is, as the name suggests, a database containing the references of the different assets of a company (servers, databases, networks, applications…). Each single item in this database is referred to as a configuration item. A CMDB’s task is to list all these assets along with their properties, thus allowing any user in the company to know which assets are available in a specific department.

Most of the time, there are two different views in a CMDB: administrator and user. The first one must be able to manage the different items, either globally or in a specific department. That includes being able to add different configuration items in the database when the company is purchasing new hardware, but also being able to update properties on specific objects. An administrator will also grant permissions of access to certain types of users according to the department in which they work. A global administrator can also be used to manage the different administrators at each department’s level.

A user, on the contrary, is only interested in viewing the contents of the database, and most specifically in knowing specific properties relative to his department. Most CMDBs will then grant a view permission only to a certain types of items related to the user’s department.

Having these two main roles in mind, it is important to note the different interests motivated by using a CMDB. On one hand, we have a user only interested in knowing certain properties on a specific configuration item. The CMDB must provide a very fast and intuitive interface so the search can be done quickly. On the other hand, there is an administrator that needs to frequently update the database and make sure each update will not have a negative impact on another configuration item. A CMDB implementation needs then to make sure each of these specifications is respected.

One might wonder what should be included in a CMDB. The answer to this question is obviously both large and subjective. Most of the time, the configuration items will be virtual objects such as a database instance, a server host, a network, a virtual storage instance, etc. Basically, it can be pretty much everything included in the company’s IT infrastructure. To this can also be added the different human resources as well.

Visually, the structure of a CMDB uses trees and graphs to show the different interactions between the components. For example, if an administrator wants to change a setting on a certain server, the CMDB gives him the possibility to see the different configuration items that would be impacted by this change. Visually, this solution is most of the time rendered under a tree or graph form to be obvious to the user.

The ITIL specifications describe the four major tasks that a CMDB must fulfill:

* Identification of configuration items to be included in the CMDB
* Control of data to ensure that it can only be changed by authorized individuals
* Status maintenance, which involves ensuring that current status of any CI is consistently recorded and kept updated
* Verification, through audits and reviews of the data to ensure that it is accurate. [5]

Today, the market presents different implementations for CMDBs, both in private and public domain. Although CMDBs were initially being developed as commercial products requiring licensing, some open-source CMDBs started to appear on the market by the mid-2000s; the main reason explaining this phenomena was the desire for some smaller companies to customize the CMDB to their needs, a difficult and even sometimes impossible task to achieve without help of a consultant from the developing company. Although not being as complete as commercial applications, open-source CMDBs give certain advantages to users. Below is a description of two of the most popular open-source CMDBs. [6]

The first one, Itop, is developed by Combodo, a French company. It offers a full customizable CMDB, along with a service desk, allowing users to create and access IT products of their company. The software also includes an error and incident management system allowing the administrator to keep track of the different errors in the system.

Another popular open-source CMDB is OneCMDB. Aimed at small- and medium sized businesses, it was developed in 2006 by Lokomo Systems AB, a company based in Sweden. The version currently available is a standalone edition coming with source code, and released under a GNU General Public License. The code is accessible on SourceForge and requires Java installed on the system in order to work. [7]

Although those two CMDBs are free to use and more easily accessible for small companies, they lack the ability to manage bigger flows of data such as bigger companies need to. In early 2015, Capterra, a website dedicated to help companies find the best business software available on the market, published an article entitled “Top ITSM Software Products”. For information, ITSM, or IT Service Management, refers to all the procedures in a company to manage IT. ITIL is a part of these procedures. [8]

The article pointed to Freshservice as the most popular solution used by companies. Developed by Freshdesk, a company based in San Francisco, it features a whole helpdesk system along with the basics of CMDBs described earlier. This solution is used by a number of big companies, such as Sony or Honda. As with the other commercial software, the annual licensing can get really expensive depending on the companies’ needs.

There are many other solutions on the market, sometimes some are being developed for an internal use only, meaning that the company which is developing it will be the one to use it. That solution can be a nice alternative to buying a license to classic CMDB provider, as it will be created to fit the company’ needs. However, it is first important to show how the CMDB can be a real advantage to the company, as it will take time and money to develop it.

In a nutshell, there are a lot of different CMDBs on the market to choose from, but most of them have the same problems: they are designed for big companies who can afford a high annual licensing cost; they require a special training to use as it takes time to ensure that everything is kept up to date; and, one of, if not the most point is that the complexity of the system can discourage users to take advantage of it, leading sometimes to cases where the CMDB is not used at all, and thus is not even kept up-to-date.

All these different problems ask for reconsidering the way to use and/or develop CMDBs. This is where the product that I developed during my internship comes as an alternative solution in order to prove that a CMDB can be both simple and necessary to use. I will now describe the solution that my manager and I came up with.

# Problem analysis and our solution: a cloud-based CMDB

In this part, I am going to describe all the steps we made in the thinking process to come up with the solution of a cloud-based CMDB.

During the first days of my internship, my manager explained to me the main flows of CMDBs. As a manager at Kohl’s, he understood the importance of having CMDBs in the company, both in a corporate and user point of view. But the reality of the market makes it harder to find a tool that is affordable, maintainable, and easy to use, both for administrators and users. To him, the most important point of a CMDB are summed up with three key words: MVP, or Minimally Viable Product. This notion of MVP is very important to understand for the rest of this report, and was the main thing to remember as I developed the application.

Basically, a minimally viable product is, as the name suggests, a product that is viable so it can do what it is supposed to in an efficient way with as few flaws as possible, but at the same time a product that is conceived in pure simplicity, or simple enough that a user should not need a manual or an explanatory note to know what to do with it. These notions should apply both to the final product, i.e. what the user will see, and the code, so that other developers who come after can understand precisely what was done. I now describe what this implies for both point of views.

## From a user point of view

For a user, as was previously written, the solution needs to be user-friendly. This goes from the main functionalities of the product to the simplest and tiniest details such as where the option bar is located or how a user signs up in the application. While developing an application that can fulfill these requirements, there are five main points to keep in mind. [9] Walker Fenton, CEO of Sepia Labs, the company behind the professional social network Glassboard, these are:

* **Understanding the context of the users that will use the application**

This point may seem obvious, but too often in the corporate world engineers and developers conceive applications and keep adding features to them while losing the view of what the application was originally intended for. It is really important to ask the following questions: what is the user expecting from the application? In which circumstances will he use the application? What main functionalities is the user expecting? With this in mind, it becomes easier to target the different options to include in the application and the ones to leave along.

* **Having a clean presentation**

The size of the screen on which the app will be run does not really matter, what matters is how to visually organize the different graphic elements of the app so that the screen will not be overloaded with components and functionalities. It is important to take the best advantage of the space available, with also leaving some empty room in order for the eyes of the user to ‘rest’ at some point.

* **Be intuitive**

This is probably one of the most important aspect to keep in mind. The user must not spend a few minutes to look for an obvious functionality. Everything needs to be smooth, easily accessible and understandable. If we take the example of Amazon, it does not take long for a user to know where to search for a product, how to add it to the cart, and how to checkout. Actually, this process is almost immediate. So it should be for any well-conceived application. While using the application, the question “What should I do?” should not be asked.

* **Have a great design**

Whether a clean presentation means using the space wisely and not overloading the screen with various components, a great design implies that the graphic used will not be of poor quality or too simplistic. In today’s digital age, users have higher expectations regarding graphics. While a need for high definition graphics will depend on the purpose of the application (videogames developers should obviously make sure this point is respected), it is still good to make sure that the images used are not blurry or of low-quality. It is actually better to not provide any graphics at all than to include low-quality images.

* **Responsiveness**

With the availability to run the same application on different devices (PC, iPhone, or tablet), the developer needs to make sure that the application will fit on the screen of each of these different devices while keeping a clean presentation. Depending on the application, responsiveness can sometimes be a challenge for the developer. That is why it is important to know how and on which devices the application will be used.

Ideally, these five points always need to be followed in order to provide the user with the best experience. However, the reality can sometimes be a little different from the theory. Sometimes, some of these rules will not be followed for a specific reason. But in most cases, one must ensure these properties will always be implemented in any application made available on the market.

While developing the solution, I made sure all of these aspects were followed. I will detail how later in this report.

## From a developer point of view

On the developer side, the minimum viable product implies that all the code that is being developed is essential to the core functionalities of the application. In other words, code that is used to prettify the application in terms of functionalities or design should not be written. As we previously saw, of course the application must be well-designed and intuitive, but this can be done while keeping code at a minimum quantity.

In terms of coding, it can be achieved by keeping simple functions that should be reused as often as possible. For each functionality, it is important to consider if something that has been previously written in the application can be reused and slightly altered to fit the new use case. This will allow to reuse as much code as possible, keeping a rather small number of lines in each file.

The other advantage of reusing same material allow the development to go much faster. This is a key point in a minimally viable product. Since the final product is more of a prototype than a final version, the development hours should be shortened so the product can be released sooner and thus get the first reviews in order to add new features to the product for a final version.

Testing is also affected: tests should be kept to a minimal number, sometimes the test consisting in checking that the application is just working well. Unit testing is not encouraged as it requires time. The only concern a developer needs to have is that the different functionalities are working and producing the desirable effect. If bugs are later discovered, they can just be fixed in the future releases of the product.

Whenever a developer needs to add a feature or modify the code, there is one question that he should consider: will the modifications affect the product in a way that the user will want it more? If the answer to that question is no, then the changes should not be made. Indeed, it is first necessary to prove that the basic concept carried by the minimally viable product can be sold on the market before improving it.

## Why a minimally viable product

One might wonder why we chose to develop a minimally viable product instead of a classic application which incorporates test cases and advanced functionalities, especially when developing a MVP is the exact opposite of what most companies expect. Usually, a company will keep asking more features to the product, thus leading to a final application that can sometimes double the size of the one initially intended.

The reason is that this cloud-based CMDB is not designed for a specific company; actually, the product was not the request of a client, or even a group of clients. It came up as an idea and, like every idea in the business world, they need to be verified.

Selling a MVP is more about selling a concept or a prototype than a real product. The product, though conceived for users, will not be sold to users, but to investors, IT managers, and other businessmen who would welcome a solution to help them be more productive in their company. Steve Blank, a manager who specialized in startup management methods, puts it in this way: “*You’re selling the vision and delivering the minimum feature set to visionaries not everyone.”* [10]

That is a key concept to keep in mind for the rest of this paper. The cloud-based CMDB that I developed is not only an application, but a new vision of using CMDBs that has not been tried before. The goal is then to write a solution with all the functionalities needed to prove to potential investors the efficiency and potential of our solution. In the case where the product is not convincing enough, when time will not have been wasted in developing something unsuccessful.

As stated earlier, developing a MVP is not common for most companies, but it can be an advantage for startups, especially while looking for new investors. Those investors, who could also be called “earlyvangelists” for early adopter and internal evangelist, are people with five major concerns that can be schemed by the pyramid represented on figure 1. [10]



Figure 1: Earlyvangelist pyramid

The pyramid addresses the concerns of these potential investors:

* They have a problem.
* They understand they have a problem.
* They are actively searching for a solution and have a timetable for finding it.
* The problem is painful enough that they have cobbled together an interim solution.
* They have, or can quickly acquire, dollars to purchase the product to solve their problem.

These people should be the target of the product, because they have the means to provide financial support to help the company pursue its development. They do not necessarily care if the product is ready to be sold, but care more about what are the different functions available, because they can see the potential the product can have on the market.

Let’s take a well-known example in the business world: Facebook. At its beginning, it was just a simple social network that allowed people to access photos and basic information on people registered on the network. It was just a minimally viable product: it allowed people to be on the same network and share data. Rapidly, though, investors started being interested in the product and using it. At first it was mostly universities and campus colleges, but then the network expanded into the corporate world, until finally being able to reach everybody aged 13 years old and older with a valid email address. During all this evolution, developers kept adding features to the social network because it worked well. Investors were convinced, so they sponsored the product, thus allowing Facebook to grow and become the billion-user network we know today. [11]

This example proves the importance of a MVP. Before investing time and money in a product, it is first necessary to check if the concept can work and can find sponsors to allow it to grow into a larger application.

## Answering the flaws of enterprise CMDBs

As we have previously seen, CMDBs available on the market are hard to keep up-to-date and, as a result, employees do not rely on them. The only way to fix that is to create a simple application that people will want to use so it can stay updated most of the time. There are, however, two important issues.

First, it is evident that a minimally viable product will never replace an enterprise CMDB published by Oracle. Even if features were progressively added to the product, creating an enterprise CMDB will require time and money. Plus, if the solution was as ambitious as one of the software found on the market, the application will lose in simplicity and we would end up with another “big” CMDB which would have the same flaws that we were trying to fix.

Second, according to the size of the company, having a CMDB is sometimes not an option. Most big companies will require one, but this implies having an enterprise solution as described in the previous paragraph. Regarding smaller companies, a CMDB could be really useful but such companies obviously will not need an enterprise application like bigger companies do, so they usually end up with no CMDB at all.

With the application, the main focus was to answer these two flaws with two strategies targeting both environments. First, we decided to primarily target small- and medium-sized companies as the potential customers for our CMDB. Such companies are looking for something simple, which described perfectly the kind of application we want to give them. They just want to focus on the core functionalities of a CMDB, so this described perfectly the minimally viable product.

The other target could be small cells or simple employees in bigger companies. While we do not expect big companies to replace their complex CMDB by a simple prototype, the solution can be advertised as a local alternative in a smaller scale and be used in different departments. Instead of representing the structure of the whole company, the CMDB could simply show the architecture of a specific department and its subcomponents. Thus, keeping the CMDB up-to-date will not require a lot of time, and employees will be able to know exactly what the different interactions between configuration items are.

This strategy of targeting two types of users with different goals could actually work if it helps people become more efficient and productive. But for thus to be done, it is important to develop the solution using the latest technologies to help ensure a smooth experience to the user.

## The importance of cloud-based applications

After targeting the audience for the application, it was important to determine which platform would support the application. Originally, my manager proposed a Java application working with Google App Engine. The advantage of this solution would be the opportunity to run the solution on any operating systems. This solution would have followed OneCMDB, an open-source CMDB that was developed in Java.

When I started to select all the tools I would need to start the Java development, I encountered an issue with the Java version. Google App Engine did not support the latest Java version (Java 8) and it was necessary to develop the application using Java 7. Although it was not a problem to download and install that version, it made me think of another problem I encountered when I tested OneCMDB, the open-source application written in Java. This solution offered a visualization of the configuration items that I could not try on my computer, because the Java plug-in was made unavailable in the latest version of Google Chrome for security issues. [12] [13] In order for it to work, it was necessary to use a previous version of Chrome.

This little problem made me realize how crucial compatibility problems are in applications. It would be very unfortunate if users could not use the product just because the technologies used encounter compatibility problems or are outdated. Using another browser could have been a solution, but if we take a look at the statistics given by the Digital Analytics Program (DAP) from the Federal Government, we realize that Chrome is the most used browser, all devices included, with 34.7% of all visitors. [14] This means that it was crucial to make the solution work on Chrome if we wanted to keep using the same technologies.

Intra-browsers compatibility can be one of the toughest parts of web development. Indeed, it is important to ensure that the application being developed will be working with the majority of browsers on the market. Figure 2 shows the market shares among browsers in the United States. It is interesting to note that Chrome is the most used one today, with a difference of 14 points with its main competitor, Internet Explorer, as of September 4, 2015. [15]



Figure 2: Web browsers market shares in the United States

This schema shows the importance of taking into account the cross-browser compatibility during the development. Although Chrome is the most used browser, IE, Firefox, and Safari are nonetheless widely used and it is necessary to make sure the application that will be developed will be compatible on these browsers as well.

In our original plan, although most part of the application would be for desktop, managing the database had to be done through Google App Engine, so using a browser was required anyway. The other consideration I had was the restrictions offered by a desktop application. The original application, as imagined by my manager and me, would be run only on a desktop computer. That could dramatically limit how the application is used.

According to the Pew Research Center, nearly two-thirds of Americans own a smartphone. [16] This does not only mean that there is an important market available regarding smartphone applications, but it also implies that Americans will most likely use their smartphone at a most frequent time than they would with their laptop, due to the permanent access to their phone. The report shows also that people use their smartphone to do much more than accessing Facebook or reading emails, for example for online banking or accessing different services related to their business. Thus, an application that would be available to both desktop computers and smartphones could be a real advantage for the popularity of the app and for the user’s convenience.

However, writing a desktop application is not the same than writing a smartphone application. Also, the compatibility problems between operating systems would still be there. A solution that appeared to me was to create a cloud-based application. It was a perfect compromise, both for the developer and the user. For the first one, there is only one main application to develop, without worrying about which operating system or type of device will run the application. For the second one, it will be much faster to use. Basically, a user (here the concept of user can also refer to a company) will just need to register and will be ready to use the application, as there will be no need to download an installer and configure an app locally.

If we get into details, we can also see the advantages of using cloud computing compared to a more standard way of developing applications. Figure 3 presents 10 of these advantages. [17]



Figure 3: The 10 benefits of cloud computing

* **Flexibility:** the demand in bandwidth can be satisfied really easily, so there is no worries to have about a potential system overflow
* **Disaster recovery:** since everything is stored on the cloud on a database, there is no need to invest time and money into complex recovery modes
* **Automatic software updates:** the user will never have to update any software since all updates will be done by the cloud provider
* **Cap-ex free:** there’s no need for capital expenditure, since services are usually pay-as-you-go. This way, it allows users to know exactly how much it will cost per month. It is also more accessible to start-ups.
* **Increased collaboration:** no matter where the different employees are, they can synchronize their work at any time, allowing updates to be released faster. It also contributes to multi-tasking on a same project.
* **Work from anywhere:** all that is needed is an Internet access and then users can log in to their account and start working, no matter where they are in the world
* **Document control:** all documents are located in a central sport. Thus, if employees are not necessarily working in the same state or time zone, there is worries to have about sending the documents by email. They will be automatically updated.
* **Security:** if a laptop gets lost, there is no worries about the data on the computer since they are stored in the cloud. And since most cloud accounts will require a password, a theft will not be able to access the cloud account of the user, and thus will fail in retrieving any useful data from the company
* **Competitiveness:** cloud-computing makes companies more competitive and more dynamic, providing smaller companies the same advantage of access to the cloud than bigger companies
* **Environmentally friendly:** since cloud computing only uses the space needed for the applications to run, the carbon footprint is decreased. According to Salesforce, there is at least 30% less of energy consumptions than using on-site servers

These different advantages show the importance that cloud-computing plays in today’s IT world. It requires companies to rethink how they conceive and use their products in order to stay competitive on the market while saving money and eventually increasing their income.

After seeing all the advantages of cloud computing, we decided to head up in that direction and create a cloud-based CMDB. Historically, this would be the first of its kind, since enterprise solutions were too complex to transfer to the cloud. The advantage of having only an MVP to develop gave us more freedom for choosing the platforms and the technologies involved.

Let’s now see which technologies were used for developing the product.

# Technologies used

In this part, I will present the different tools that I used to implement the solution.

## Programming languages

The project did not have specifications regarding a particular language to use. My manager also gave me the opportunity to choose the language in which I am the more comfortable to develop. While doing web development, there is a lot of possibilities to choose from. Obviously, web pages would be written in HTML5 and styled with CSS3, using the framework Bootstrap for a better design. Regarding the server side, I would use PHP to communicate between the application and the server.

Regarding the different graphic elements and their interactions with the user, the code would be written in JavaScript and jQuery. This was chosen because of the tools that I found to represent the data from the server. Each of these tools will be explained in subsequent paragraphs. Since they were written in JavaScript, it would be easier to write the code in the same language. I also used jQuery to dynamically interact with some HTML elements, and also to asynchronously load different parts of the webpages, thus providing a more user-friendly interface. Finally, I used a MySQL database with the query language used being SQL with a syntax specific to MySQL.

## XAMPP

XAMPP is an Apache distribution of a full PHP development environment. It includes an Apache server, a MySQL database, FileZilla, Mercury, and Tomcat. For the project, I only used the Apache server to run the webserver, and the MySQL database. I chose XAMPP because it is a pretty neat tool; indeed, there is no need to configure a server and a database separately since XAMPP provides both. The configuration time is non-significant: all that is required is to download the installer and follow the instructions provided. Also, it is open-source, which dismisses any cost that a commercial license could have. [18]

## phpMyAdmin

Written in PHP as the name suggests, phpMyAdmin is a free MySQL database that provides a web interface to manage the different tables. One of the most popular MySQL databases, it is included in the XAMPP development environment, thus being more simple to handle. It includes a lot of features, like administering multiple servers, global search in a specific database or in different subsets, data import and export, and much more. [19]

Although I was free regarding the choice of my development tools, my manager and I were aware that, according to the data model of the application, a relational database needed to be used. Indeed, there are many interactions between the different tables, as we will see in the next part, so only a relational database could provide the best storing options.

The choice of a MySQL database was done because of its license type. Being open-source, it is easier to use for developing an application prototype because it is costless for the company. Moreover, installing a standard Oracle SQL database requires both time and money, and is something much too ambitious regarding the purpose of the application being developped.

## jsTree

jsTree is an open-source jQuery plugin providing interactive trees to include in web pages. The trees provided have hierarchical views and can be customized with personal icons. It also supports JSON data sources and AJAX loading. The project is still active on GitHub and involves a community of 24 contributors. Currently, the stable version is 3.2.1. [20]

I needed a tree to display the different configuration items into folders, and that plugin was exactly what I was looking for. It offers a great number of functions allowing the developer to incorporate his own functions regarding the different interactions available with the tree. It also has built-in functions such as renaming, creating, deleting, or selecting nodes, which can be easily called to handle events with the user. The API documentation is also plentiful.

## Vis.js

Vis.js is a visualization library that provides different interfaces to interact with dynamic data. The library includes different components allowing to visualize data on different representations. It includes interactive 2D and 3D graphs, different network views used to show the interdependence in a graph, a timeline allowing the creation of events as specific dates, and a data set, a data structure to format the data imported into the different components. The Vis library is a JavaScript component compatible with all modern browsers and allows the user to dynamically interact with the data. It is an open-source project available on GitHub that is still being developed. [21]

Using the library is fairly easy; the developer includes it in his work repository, references it in the html page he wants to use it on, and gives a specific id to the HTML component in which he wishes to integrate the visualization. Vis.js being developed in JavaScript, the different functions to interact with the library need to be written in JavaScript as well.

## Aptana Studio

Aptana Studio is an open-source web development IDE. It provides support for many different languages, including HTML, CSS, PHP, JavaScript, jQuery, Python, Ruby, and Rails. Considered the world’s most powerful open-source web development IDE, it offers an Eclipse-like interface as well as different tools allowing the developer to be more productive. The software is being developed by the company Appcelerator. The current version is Aptana Studio 3. [22]

Before choosing an IDE, I wanted to make sure that it would offer support for all the languages I needed, but I also wanted something that would not be too cumbersome to use. For example, I could have used Visual Studio to develop the application, because that IDE also supports web development. However, it would not have been necessary to use such a complex tool if there was an alternative.

The other reason for choosing Aptana is the open-source license. It is easier and cheaper to work with open-source licenses than to purchase a Microsoft license. Indeed, since my project is for a company and not for myself or a school, I could not have used my Visual Studio student license to write the application, since a professional license would have been required by the law.

All these different tools were crucial in developing the application. I will now enter into the core part of this report, namely how the solution to the initial problem was implemented.

# Implementation of the solution

I will now present the implementation process of the cloud-based CMDB. Beforehand, it is necessary to mention that the first month of my internship was spent in research and brainstorming with my supervisor in order for me to better understand the problem, master the aspects that a CMDB should include, and have the vision on how to develop the product with choosing the right technologies. This first step was essential to start the development of the application in the best possible conditions.

## Data model

The first step in the development was to write a data model for the database. This would be the building block of the application. It was very important to keep this data model simple in order to follow the specification of a minimally viable product. The following three figures show the different interactions between the tables of the model.

Figure 4 represents the user table. The first step of the application is for user to register themselves. There are two different views: user and administrator, represented in the table with the column *isAdmin* (0 for user, 1 for administrator). Although a generated id is used for the primary key, the unique identifiers used during the authentication process are the email address and the user name. These two fields are unique according to the user, and interchangeable during the log in process (this means that one can either use his/her email address or username to log in to the application). Regarding the password, it is encrypted (cf. next part on user security).



Figure 4: User table

Regarding the application data, figure 5 shows the interaction between the different tables to manage the configuration items and their properties. Each configuration item is contained within a class. For a reminder, a configuration item is viewed as a specific item within the company, for example a Linux server with a certain hostname. In order to organize them, similar configuration items are placed into a class if they share similar properties. For example, all the servers will be included in the class Server. A class can have a parent class as well; if we take our Server class example, it can have two child classes: Linux Server and Windows Server. The identifier for each parent’s class is saved in the child class table; thus, knowing a specific configuration item, it is easy to get the whole set of classes in which the item is included in. It is important to note that a configuration item cannot have children. To understand that, we can view a class as a folder and a configuration item as a file.

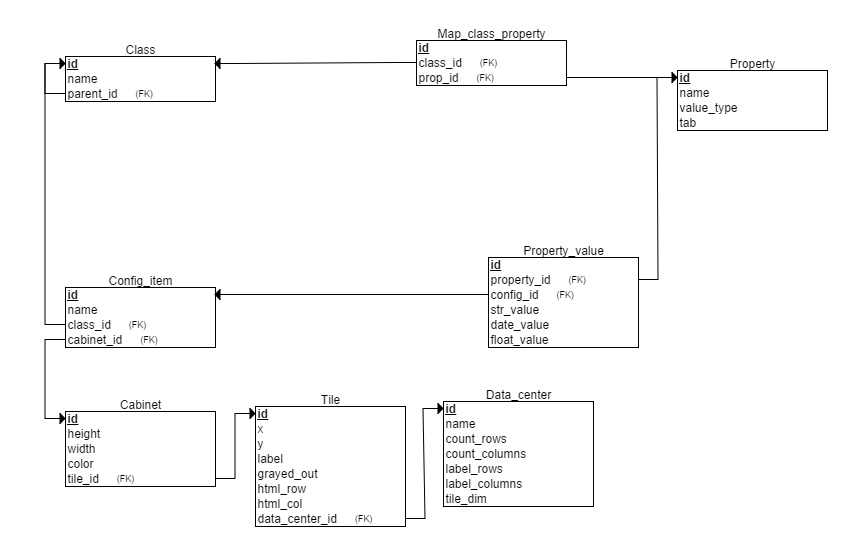


Figure 5: Configuration items data model

Classes contain properties as well. Each configuration item in a class will inherit its parent’s properties. Since a same property can be used by multiple classes who are not linked together, it was necessary to place the columns in a different table, the *Property* table; mapped to a specific class through the *Map\_class\_property* table in order to know all the properties references to a specific class.

A property is defined by a unique id and a name known to the user (like the other tables’ id, the user will never access the id unless he goes directly to the database). Each property also has a value type that can be a float, a string, or a date. The last field determines which type the properties is assigned to. Indeed, properties can be of three different types: general (for example a hostname), financial (like the maintenance cost), or labor (like the number of employees using the configuration item).

Since the same property can be used with multiple classes, there cannot be a value field in the property table, otherwise all properties would have the same value no matter what class it would be assigned to. This calls for another table, *Property\_value*, which links the value to its property and to the configuration item that will receive the value. To store it, the table includes three different columns according to the property type. When a new value is created, it is inserted in its corresponding type field and the two other columns are set to NULL. The choice of having three columns instead of a single one was done regarding the data formatting. Each field has a special input format specific to its value type that was necessary to follow when retrieving the value to load it onto the web page.

All these tables are used to display the configuration items properties on the first panel in the application. The second part of the data model focuses on the data centers. The specifications of the project included the possibility to manage views of the different data centers used in the company. That is what the tables *Data\_center*, *Tile*, and *Cabinet* are there for.

In real life, a room where a data center is located is divided into tiles. Each tile is identified by a specific row and column that allows someone to know precisely where a specific server is located. The table *Data\_center* will record the properties of the room, *Tile* will, as the name suggests, record the properties of the tile, and the table associated to a cabinet will represent the physical object positioned in the data center. It can be a server or something different; for example, ventilation systems can be represented in the room as well.

The first table specifies the properties of the data center with a number of rows and columns, and a name. The labels are used to know which letter or number will start the columns or rows. Most of the time, it will be a letter and a number like A1. Sometimes though, it can be different and use either two letters and a number (like AA1) or numbers not starting at 1 (like A30). It is then essential to give the user the opportunity to choose this start label to match the reality. The last column specifies the dimension of the tiles. Since all tiles are squared, only one field is needed for the dimensions.

The *Tile* table has a certain numbers of parameters, all essential to get all the important information in the application. The coordinates (x, y) will give the exact positioning of the tile, taking into account the tile dimensions. It means that, if a tile is 2x2, the coordinates of the first tile on the upper left will be x=2 and y=2. These columns ought not to be confused with *html\_row* and *html\_col*, which are the coordinates used by the application to locate the cells. In the previous example, the upper left cell parameters would be html\_row=1 and html\_col=1. The label will use the room notation systems as defined in the *Data\_center* table. If label\_rows=1 and label\_columns=A, then for that same tile we would have label=A1. Lastly, the *grayed\_out* column specifies if the tile is accessible, meaning if a cabinet can be positioned on the cell. The values possible are either 0 or 1, the first case being if the tile can be used.

The *Cabinet* table represents a component positioned on a tile. Most of the time it will be a server, but it can also be an air conditioning cabinet or something else as well. The different properties recorded here (height, width, and color) are for information only and will not be rendered on the grid. Something important, though, is the reference to a configuration item. If the cabinet is a server, then the system must know which configuration items are saved on this server. Thus, a user, knowing the reference to a specific cabinet, will also be able to get the references to all the configuration items saved on that server.

Another part of the data model is shown on figure 6. It represents the different applications used within the company. An application can either be a software purchased by the company (like Oracle Application Express), or a service website with an annual fee (like Microsoft Azure). Each application incorporates a network graph which shows the interaction between an application and the configuration items using it. Each application can be used by one or more items, and the user can organize them into folders in the graph.

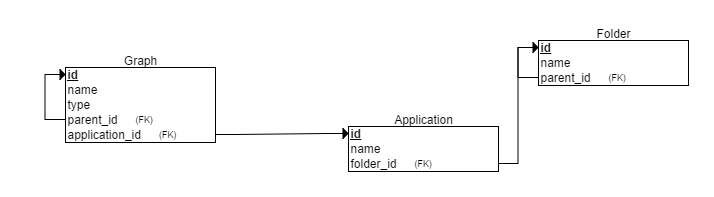


Figure 6: Applications data model

As with the other tables, each application is identified by a unique identifier automatically generated upon insertion in the database. It is not known from the user. Applications are grouped into folders according to their type (for example all IT applications will be grouped inside the same folder). The graph table is used to represent the network. The architecture of this graph is interesting because of the way it is built. A graph is none other than a node with a reference to its parent and a type, which can be an application, a folder, or a configuration item. It is important to note that the folder in the graph is totally different from the *Folder* table of the data model.

The root node is the application, so its *parent\_id* is NULL. Then, the user can add a node, which will create a new graph with a reference to the previous node in the *parent\_id* column. The graph is then built on this same pattern. Adding nodes this way makes it really easy to maintain a simple data model that is efficient with lesser complexity. It is important to note that only the root node can be an application, and that a configuration item cannot have any child nodes.

Now that we have seen the different components of the data model, we can go deeper into the different functions of the cloud-based CMDB while looking at the code of the core functions.

## User administration

One of the first functions to implement were the ones related to the user administration. It was important to have the two different views at first in order to know exactly what view would be available to the user and the administrator. Obviously, security was a key factor in order to make sure maleficent users could not steal passwords or other information. When accessing the application for the first time, a user will be directed to the log in screen showed on figure 7.



Figure 7: Log in screen

If the user is not registered, he can click on sign up and will be redirected to the following page.



Figure 8: Sign up screen

This asks for the user’s basic information as most websites would do. There is a password check field in order to make sure that the user did not make any typos in his password when creating it. The type of user is determined with the admin field. Each field, as it is with the login page as well, is processed to make sure that what was input in the field corresponds to the type of data it should contain. This allows the servers to not get useless queries that could overload it. Once the information entered match the data type and the different specifications (for example, an email address matches a specific format), the data is sent to the server and processed according to the following function.



Notice the use of the PHP function *bindparam*. It is used to create **prepared statements** *(see glossary)*. Doing this has two advantages:

* It can execute the same query with different sets of parameters with a higher efficiency
* It helps preventing **SQL injections** *(see glossary and addendum I for more information on SQL injections and how to avoid them)*

Similar techniques were used during the whole project to make sure any inputs from the user was securely processed before being sent to the server, thus making the application free from any SQL injections.

An important notion to consider in the user administration is how passwords are managed. For security reasons, they cannot be plainly recorded in the database. The idea is to apply a hashing function and record the hash value in the database. In the previous code snippet, this is done with the function *password\_hash* that takes two parameters: the password entered by the user and the hashing algorithm. PHP provides different algorithms; the one used here is the bcrypt algorithm *(see addendum II for more information regarding password hash and the bcrypt algorithm)*.

Once the user has successfully registered, he can then log in to the application using the screen as previously shown on figure 7. The PHP login function will query the database to know if a user with either the email address or the username entered by the user exists. Then, the built-in PHP function *password\_verify* makes sure that the given password matches the hash recorded in the database. All information needed to verify the password are included in the function, which makes things easier for the developer.



Once a user is logged in to the application, a session is created for that specific user, which will allow the application to know which users are currently logged in to the application. Thus, the system can manage the different authorizations, because not all users will have the same access privilege to certain web pages. This is important to make sure that a page reserved to an administrator cannot be accessed by a simple user, and vice versa. In order to do that, each page will include at its beginning a few lines to:

* Check that the user viewing the page is logged in
* Make sure the user logged in has the right to view the page

The following snippet shows how the application first checks that the user is logged in, then retrieves his information from the database. Now the permission of the user is known, allowing the system to redirect the user to another page if the visibility of the current page is not intended for the user type. This is the code included at the beginning of the administrator configuration item page. Each page has a similar heading as well.



The last feature necessary for a smooth user management is the ability to change passwords. It is the only parameter that can be directly changed by the user. The choice of not allowing the user to change anything else was done regarding the notion of MVP. I thought it was not necessary since the email address or the username will rarely be changed. The user cannot change the administrator rights as well, but this makes sense because, if he is a simple user, he will need the authorization from the database administrator in order to be promoted and access to more restricted information. While on the interface, a menu will allow the user to access his personal information where he can modify his password, as shown on the following figure.



Figure 9: Account information page

The data sent by the user to the server is processed with a PHP function similar to the *login* function seen earlier. As previously seen, the new password is hashed before being recorded in the database.

This concludes the part regarding users’ administration. Now, let’s study the interaction between the application and the database.

## Interaction with the database

In this part, I will present how the application is interacting with the MySQL database, in this case being phpMyAdmin.

### Connection to the database

In PHP, there are two ways to communicate with a database: PDO or MySQLi. One of the main differences between these two is that PDO supports twelve different database drivers, whether MySQLi only support MySQL driver. [23] Although the application will be running with a MySQL database, what if a future client uses an Oracle database? The communication protocol would then need to be rewritten with PDO to match the client’s database. Another advantage of PDO is the possibility to write prepared statements. As we have seen in the previous part, prepared statements are more efficient and essential for security purposes. For these two reasons I adopted PDO as the communication protocol in the application.

To start communicating with the database in PDO, the first thing to do is to open a session (as shown in the following code snippet). This is done when the user first accesses the application (accessing the application does not mean logging in, but simply going to the login or signup webpage).

*db\_connect.php*: function to connect to the database



First, a session is created with the function *session\_start*. The database parameters are then given to the system (they are kept confidential here for security purposes) and used to create a PDO object that will represent the connection. Then, the user file containing the different functions previously seen (login, signup…) is included, allowing the creation of a user being based on the current connection. This simple script allows a user to connect to the database, either by creating an account or logging in.

### User session

Once a user is logged in, the application needs to remember who he is when he accesses the different pages. On each web page of the application, the system needs to make sure that:

* The user has been identified
* The user has the permission to access the current web page

For this, it is necessary to create a session per user to grant him access to the different parts of the website. This is done by inserting the following code snippet on each page.



If the user is not logged in, he is redirected to the login page from which he can either enter his credentials or create a new account. Once he is logged in, a session is created, allowing him to stay connected on the website as long he is active. Indeed, after a certain period of inactivity, he is automatically logged out. That security measure helps to prevent unauthorized account access in case the user, for some reasons, leaves his computer with access to the website.

The application also checks the permission. Each part of the website comes with two different pages: one for the user, and one for the administrator. Thus, according to the permission, the user will be redirected to the adequate page.

The choice of having two views for each page was done for simplicity reasons. Indeed, administrator pages call for numerous functions to process the different events and interactions with the database. Adding different events to manage the scenarios when the user is administrator or not would have complicated the code, so I decided to separate the two processing views. They will be discussed in the continuation of this report.

### Retrieving data with Ajax

To load data from the database into the application, Ajax is a common tool to be used. It is a client-side script that communicates from the database to the application and vice-versa. There are four main advantages to using this technique:

* **Asynchronous calls:** the client browser can make asynchronous calls to the server, which allows the user to keep interacting on the webpage without waiting while all the data is being loaded
* **Callbacks:** data can be sent or retrieved from the server without having to send all form data or reload the whole page. It is very useful in order to reload only a specific part of the webpage.
* **User-friendly:** because of the callbacks, the web page will be more responsive and faster to load.
* **Speed:** Ajax allows major improvements over speed in an application thanks to the callbacks. There is no need to wait for the page to reload each time there is an update or a request to the database. [24]

For these reasons, Ajax was used whenever there was an update to the database or a request to retrieve data from it. For all the operations available to the user (updating/retrieving data, creating new items…), a PHP script would include the SQL query. The following example shows how the configuration items from the database are loaded.

*app\_db\_loadConfigItems.php*: function to load the configuration items



After including the file necessary to open the connection, the header specifies the format in which the data will be returned. JSON was used as it is the most convenient format to process. The SQL query is then written, taking into account the security specifications talked about earlier. The result is then encoded in JSON and sent to the client. The previous file returns the following results from the server.



This JSON can then be processed on the client side. The following snippet shows how Ajax is used to call the PHP script and handle the JSON sent back from the server.



The Ajax query takes different fields into account. First, the type specifies the http method used to send and retrieve information to the server. There are usually two main methods, POST and GET. POST is used here because it is more secure then GET; indeed, in the last one, data is sent as part of the URL, which makes it very unsafe to use if some private information is sent in the query.

The URL parameter defines the PHP file which will be executed with the Ajax request. If the file is successfully executed without throwing any exceptions, then the success field is executed. Here, the callback function is specified and tells the application what to do with the data retrieved from the server. In the example, it creates an input control for each configuration item in the system.

### Updating data with Ajax

When sending data to the database, the application will use the same techniques as retrieving data, with some differences in how the data is handled. First, the SQL query will be included in a PHP file as previously seen, but no data will be returned from the server this time. Let’s take the example of renaming a configuration item.

*db\_renameConfigItem.php*: function to rename a configuration item



The data posted by the user are processed in a secure way to avoid an SQL injection before executing the query. Then, the Ajax callback function is executed.



Another field that was not included in the previous example is the data field. It includes the different parameters that are sent to the server in order to be process the query (retrieved by the $\_POST function in the PHP script). If the script is executed successfully, the tree showing the configuration items is refreshed without reloading the whole page.

This shows the importance of using Ajax to process data while communicating with the server in a web application.

# Conclusion

# References

|  |  |
| --- | --- |
| [1] | Prime Resources LLC, "What We Do," May 2015. [Online]. Available: http://www.primeresources-llc.com/index.php. [Accessed 1 September 2015]. |
| [2] | Internal Revenue Service, "Limited Liability Company (LLC)," 5 August 2015. [Online]. Available: http://www.irs.gov/Businesses/Small-Businesses-&-Self-Employed/Limited-Liability-Company-LLC. [Accessed 1 September 2015]. |
| [3] | Hudson Business Lounge, "The Hudson," 2014. [Online]. Available: http://www.hudson-business-lounge.com/. [Accessed 1 September 2015]. |
| [4] | AXELOS Global Best Practice, "ITIL - IT Service Management," [Online]. Available: https://www.axelos.com/best-practice-solutions/itil.aspx?utm\_source=itil-officialsite&utm\_medium=redirect&utm\_campaign=redirects. [Accessed 1 September 2015]. |
| [5] | M. Rouse, "Configuration management database (CMDB) definition," December 2006. [Online]. Available: http://searchdatacenter.techtarget.com/definition/configuration-management-database. [Accessed 1 September 2015]. |
| [6] | Riccardo, "6 Open Source CMDB," Linuxaria, 19 February 2011. [Online]. Available: http://linuxaria.com/article/6-cmdb-open-source. [Accessed 1 September 2015]. |
| [7] | OneCMDB, "Home Page," 2 April 2012. [Online]. Available: http://onecmdb.org/wiki/index.php?title=Main\_Page. [Accessed 1 September 2015]. |
| [8] | Capterra, "Top ITSM Software Products," [Online]. Available: http://www.capterra.com/itsm-software/. [Accessed 1 September 2015]. |
| [9] | W. Fenton, "Top 5 Rules for Creating User Friendly Mobile Apps," 31 May 2012. [Online]. Available: http://www.cmswire.com/cms/customer-experience/top-5-rules-for-creating-user-friendly-mobile-apps-015841.php. [Accessed 1 September 2015]. |
| [10] | S. Blank, "Perfection By Subtraction – The Minimum Feature Set," 4 March 2010. [Online]. Available: http://steveblank.com/2010/03/04/perfection-by-subtraction-the-minimum-feature-set/. [Accessed 1 September 2015]. |
| [11] | S. Phillips, "A brief history of Facebook," The Guardian, US Edition, 27 July 2007. [Online]. Available: http://www.theguardian.com/technology/2007/jul/25/media.newmedia. [Accessed 1 September 2015]. |
| [12] | Oracle, "Java and Google Chrome Browser," [Online]. Available: https://java.com/en/download/faq/chrome.xml. [Accessed 2 September 2015]. |
| [13] | Google, "NPAPI deprecation: developer guide," January 2015. [Online]. Available: https://www.chromium.org/developers/npapi-deprecation. [Accessed 2 September 2015]. |
| [14] | S. J. Vaughan-Nichols, "Most popular US web browsers, according to the federal government," ZDNet, 26 March 2015. [Online]. Available: http://www.zdnet.com/article/the-most-u-s-popular-web-browsers/. [Accessed 2 September 2015]. |
| [15] | Clicky Web Analytics, "Web browsers (US marketshare)," 4 September 2015. [Online]. Available: https://clicky.com/marketshare/us/web-browsers/. [Accessed 4 September 2015]. |
| [16] | A. Smith, "U.S. Smartphone Use in 2015," Pew Research Center, 1 April 2015. [Online]. Available: http://www.pewinternet.org/2015/04/01/us-smartphone-use-in-2015/. [Accessed 2 September 2015]. |
| [17] | SalesForce, "Why Move to the Cloud? 10 Benefits of Cloud Computing," 2015. [Online]. Available: http://www.salesforce.com/uk/socialsuccess/cloud-computing/why-move-to-cloud-10-benefits-cloud-computing.jsp. [Accessed 2 September 2015]. |
| [18] | Apache Friends, "XAMPP," 2015. [Online]. Available: https://www.apachefriends.org/index.html. [Accessed 5 September 2015]. |
| [19] | phpMyAdmin contributors, "Bringing MySQL to the web," 2015. [Online]. Available: https://www.phpmyadmin.net/. [Accessed 5 September 2015]. |
| [20] | jsTree, "What is jsTree?," [Online]. Available: https://www.jstree.com/. [Accessed 5 September 2015]. |
| [21] | Almende B.V., "vis.js," 2015. [Online]. Available: http://visjs.org/. [Accessed 5 September 2015]. |
| [22] | Appcelerator, Inc., "Aptana Studio 3," 2014. [Online]. Available: http://www.aptana.com/index.html. [Accessed 5 September 2015]. |
| [23] | D. Marjanovic, "PDO vs. MySQLi: Which Should You Use?," tuts+, 21 February 2012. [Online]. Available: http://code.tutsplus.com/tutorials/pdo-vs-mysqli-which-should-you-use--net-24059. [Accessed 9 September 2015]. |
| [24] | W. Fote, "What is Ajax and Where is it Used in Technology?," Segue Technologies, 13 March 2013. [Online]. Available: http://www.seguetech.com/blog/2013/03/12/what-is-ajax-and-where-is-it-used-in-technology. [Accessed 10 September 2015]. |
| [25] | w3schools.com, "SQL Injection," Refsnes Data, 2015. [Online]. Available: http://www.w3schools.com/sql/sql\_injection.asp. [Accessed 8 September 2015]. |
| [26] | Wikipedia, "Hash function," 7 August 2015. [Online]. Available: https://en.wikipedia.org/wiki/Hash\_function. [Accessed 8 September 2015]. |
| [27] | S. K. Bansal, "Securing Passwords with Bcrypt Hashing Functions," The Hacker News, 10 April 2014. [Online]. Available: http://thehackernews.com/2014/04/securing-passwords-with-bcrypt-hashing.html. [Accessed 8 September 2015]. |

# List of figures

[Figure 1: Earlyvangelist pyramid 20](#_Toc429663955)

[Figure 2: Web browsers market shares in the United States 23](#_Toc429663956)

[Figure 3: The 10 benefits of cloud computing 24](#_Toc429663957)

[Figure 4: User table 29](#_Toc429663958)

[Figure 5: Configuration items data model 30](#_Toc429663959)

[Figure 6: Applications data model 32](#_Toc429663960)

[Figure 7: Log in screen 33](#_Toc429663961)

[Figure 8: Sign up screen 33](#_Toc429663962)

[Figure 9: Account information page 36](#_Toc429663963)

[Figure 10: Principles of a hash function 51](#_Toc429663964)

Glossary

**Prepared statement:** database feature used to execute the same query at multiple times with different parameters and a higher efficiency. Parameters in the query are not specified but replaced by variables. The database then compiles the query and optimizes it. Then, when the application specifies or *binds* the variables to the database, the query is executed.

**SQL injection:** technique used by hackers to attack data-driven applications. It consists of inserting SQL statement inside a field sent to the server in order to interact with the database (for example to get the user’s email address or login information).

Addenda

1. SQL injections and how to prevent them

A SQL injection is a technique used by malicious users to destroy the database or get data from it. If not secured, they can be highly dangerous as they compromise the security of the application. A SQL injection can be created using a text field, for example on a search field or a login page. The W3 schools website provides a good example of a typical SQL injection. [25]

Let’s consider a login page where the user is requested to enter his user name and password.

User Name:  


Password:  


To process the data entered by the user without any security features, a typical request would look like this:

uName = getRequestString("UserName");  
uPass = getRequestString("UserPass");  
  
sql = "SELECT \* FROM Users WHERE Name ='" + uName + "' AND Pass ='" + uPass + "'"

If the user is honest, it would work fine. But someone with bad intentions could get the names and passwords by inserting a SQL query inside one of the fields, like " or ""=".

The SQL statement executed by the server then becomes:

SELECT \* FROM Users WHERE Name ="" or ""="" AND Pass ="" or ""=""

This statement will always be true, and then the server will return the list of all users with their names and passwords. This is the reason why it is necessary to secure the inputs send by the user to the server.

To do this, there needs to be some steps added into the PHP file. In the following example taken directly from the CMDB application, the user is asked to rename a field.



We can notice that the user’s inputs are not directly inserted into the query. First, the SQL statement is prepared, it means the system parses it once and then it can be executed many times with different parameters without the need to parse the query another time. Then, it is executed with the parameters entered by the user. If these are invalid or potentially dangerous as previously seen, then the query won’t be executed and there will be no risk of data corruption.

This technique was used for the whole development of the application to maintain a secure application, safe to use, both for users and administrators.

Hash functions

When securing a password to store in a database, it is common practice to use a hash function. The purpose of a hash function is to encode the passwords in order to prevent a maleficent user who would get in possession of the hash to retrieve the original password. The following schema shows how a hash function works. [26]



Figure 10: Principles of a hash function

The result of this function – the hash – is then stored in the database instead of the real password. There are currently different algorithms that produce a hash. Some are more secured than others. PHP offers different hash functions. The one used in the application is based on the Bcrypt algorithm, a “method of hashing passwords [that] is solid enough for most web applications that stores users’ passwords and other sensitive data.” [27]

<Dos de couverture>

Résumé

Mots-clés :

Abstract

Keywords :