Software Engineering work placement

At Azotel Technologies

Zhongyu Li
Work Placement Project 2010
M.Sc. Software Engineering

Department of Computer Science
National University of Ireland, Maynooth
County Kildare, Ireland

A thesis submitted in partial fulfillment
Of the requirements for the
M.Sc. Software Engineering

Head of Department: Dr. Adam Winstanley
Supervisor: Prof. Ronan Reilly
30th, January, 2011
Declaration

I hereby certify that this material, which I now submit for assessment on the program of study leading to the award of M.Sc. in Software Engineering, is entirely my own work and has not been taken from the work of others – save and to the extent that such work has been cited and acknowledged within the text of my work.

Signed  Zhongyu Li  Date 30th, January, 2011
Table of contents

Tables and figures: .................................................................................................................. 6

Preface ................................................................................................................................... 7

1. Overview of the company ....................................................................................................... 8
   1.1 Background .......................................................................................................................... 8
   1.2 Azotel technologies Ltd ....................................................................................................... 8
   1.3 Overview of the product ...................................................................................................... 8
   1.4 Actives Covered .................................................................................................................. 9
      1.4.2 Interface programming for Radio Equipment ............................................................... 9
      1.4.3 Network Protocol & Service .......................................................................................... 9
      1.4.4 Network Security ......................................................................................................... 9
   1.5 The mission ........................................................................................................................ 9
   1.6 My role .............................................................................................................................. 10

2 SS project ................................................................................................................................ 11
   2.1 SS organization structure .................................................................................................. 11
      2.1.1 The project lifecycle ..................................................................................................... 11
      2.1.2 The project staff and roles definition .......................................................................... 14
   2.2 Development workflow ....................................................................................................... 15
   2.3 Internationalization ............................................................................................................ 16

3 Project Architecture, development guide and Technical Overview ........................................ 17
   3.1 Architecture Overview ....................................................................................................... 17
      3.1.1 Architecture in Perl programming ............................................................................ 17
      3.1.2 System component Introduction ................................................................................... 20
   3.2 Software Development Guide ............................................................................................. 22
      3.2.1 Procedure to deploy updates/changes ....................................................................... 22
      3.2.2 Mechanics .................................................................................................................... 23
   3.3 Technical overview .............................................................................................................. 24
      3.3.1 Overview of the software environment ....................................................................... 24
      3.3.2 "LAMP" environment .................................................................................................. 24
      3.3.3 Third party software .................................................................................................... 25
      3.3.4 General applications .................................................................................................... 26

4 The Perl client .......................................................................................................................... 27
   4.1 The perl application ............................................................................................................ 27
      4.1.1 Contact book enhancement .......................................................................................... 27
      4.1.2 RT Integration ............................................................................................................... 29
      4.1.3 History list of recently viewed customers ..................................................................... 30
      4.1.4 Site/Equipment enhancement ....................................................................................... 32
         4.1.4.1 Using CPAN module Net::FTPSSL to allow a user attach files via a interface .... 32
         4.1.4.2 Google Map Interaction .......................................................................................... 33
         4.1.4.3 Using crontab(Linux task schedule application) to implement customer maintenance. 34
      4.1.5 An enhancement for interface usability ........................................................................ 36
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.1.6</td>
<td>Export customer report</td>
<td>38</td>
</tr>
<tr>
<td>4.1.7</td>
<td>Linux process interruption (Software-level)</td>
<td>39</td>
</tr>
<tr>
<td>4.2</td>
<td>Non-programming Task</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1</td>
<td>System Data backup and Disaster recovery</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1.1</td>
<td>Introduction to data backup and disaster recovery</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1.2</td>
<td>Common Data protection strategies</td>
<td>41</td>
</tr>
<tr>
<td>4.2.1.2.1</td>
<td>Data protection strategies and hardware structure in Project SS</td>
<td>42</td>
</tr>
<tr>
<td>4.2.1.2.2</td>
<td>Data backup</td>
<td>44</td>
</tr>
<tr>
<td>4.2.1.2.3</td>
<td>Data recovery</td>
<td>44</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Network Security</td>
<td>44</td>
</tr>
<tr>
<td>4.2.2.1</td>
<td>Introduction to Network security</td>
<td>44</td>
</tr>
<tr>
<td>4.2.2.2</td>
<td>Communication via secure tunnel</td>
<td>44</td>
</tr>
<tr>
<td>4.2.2.3</td>
<td>Firewall &amp; IP table configuration based on Linux OS</td>
<td>45</td>
</tr>
<tr>
<td>5</td>
<td>Software Testing in Project SS</td>
<td>47</td>
</tr>
<tr>
<td>5.1</td>
<td>Review of previous Testing procedure</td>
<td>47</td>
</tr>
<tr>
<td>5.2</td>
<td>Testing for Web based applications</td>
<td>47</td>
</tr>
<tr>
<td>5.3</td>
<td>Introduction to the new testing plan</td>
<td>48</td>
</tr>
<tr>
<td>5.3.1</td>
<td>Functionality test</td>
<td>48</td>
</tr>
<tr>
<td>5.3.1.1</td>
<td>Form Validation</td>
<td>48</td>
</tr>
<tr>
<td>5.3.1.2</td>
<td>Code / Script testing</td>
<td>49</td>
</tr>
<tr>
<td>5.3.1.2.1</td>
<td>Version control</td>
<td>49</td>
</tr>
<tr>
<td>5.3.1.2.2</td>
<td>Black box testing</td>
<td>49</td>
</tr>
<tr>
<td>5.3.1.2.3</td>
<td>White box testing</td>
<td>50</td>
</tr>
<tr>
<td>5.3.1.3</td>
<td>Link validation</td>
<td>50</td>
</tr>
<tr>
<td>5.3.2</td>
<td>Load &amp; Stress test</td>
<td>50</td>
</tr>
<tr>
<td>5.3.3</td>
<td>User Interface test</td>
<td>50</td>
</tr>
<tr>
<td>5.3.4</td>
<td>Compatibility test</td>
<td>51</td>
</tr>
<tr>
<td>5.3.5</td>
<td>Security test</td>
<td>51</td>
</tr>
<tr>
<td>6</td>
<td>Industry standard and Agile software development</td>
<td>52</td>
</tr>
<tr>
<td>6.1</td>
<td>Requirements Engineering and System Design</td>
<td>52</td>
</tr>
<tr>
<td>6.1.1</td>
<td>Overview of the Requirements Analysis and System Design</td>
<td>52</td>
</tr>
<tr>
<td>6.1.2</td>
<td>Comparison between Industrialized standard and Academic standard</td>
<td>52</td>
</tr>
<tr>
<td>6.1.2.1</td>
<td>Requirement Engineering in Project SS</td>
<td>54</td>
</tr>
<tr>
<td>6.1.2.1.1</td>
<td>Current requirement analysis processing</td>
<td>54</td>
</tr>
<tr>
<td>6.1.2.1.2</td>
<td>Requirement analysis proposal for Project SS</td>
<td>55</td>
</tr>
<tr>
<td>6.1.2.2</td>
<td>System Design in Project SS</td>
<td>57</td>
</tr>
<tr>
<td>6.1.2.2.1</td>
<td>Low-Fidelity Prototype</td>
<td>57</td>
</tr>
<tr>
<td>6.1.2.2.2</td>
<td>Overview of System design</td>
<td>57</td>
</tr>
<tr>
<td>6.1.2.2.3</td>
<td>System designing Proposal for Project SS</td>
<td>58</td>
</tr>
<tr>
<td>6.2</td>
<td>Coding in Project SS</td>
<td>62</td>
</tr>
<tr>
<td>6.2.1</td>
<td>Coding Style</td>
<td>62</td>
</tr>
<tr>
<td>6.2.2</td>
<td>Efficiency</td>
<td>63</td>
</tr>
</tbody>
</table>
Tables and figures:

Fig 1 Standard software lifecycle.................................................................11
Fig 2 An illustration of the spiral model for the Project SS.................................13
Fig 3 A pie chart for time spent on each phase in the lifecycle..........................13
Fig 4 Gartner Survey of S/W projects..........................................................14
Fig 5 Development tasks collecting and distributors.........................................16
Fig 6 MVC prototype.......................................................................................17
Fig 7 An example of the not-MVC architecture in Project SS...............................18
Fig 8 The procedure to import and use a module in Project SS.........................20
Fig 9 Procedure to deploy updates/changes.....................................................23
Fig 10 Customer Contact Form........................................................................27
Fig 11 Edit Contact Page..................................................................................28
Fig 12 Data transaction (To database)...............................................................28
Fig 13 Data transaction (From database)...........................................................29
Fig 14 Fields to define the target type and destination email addresses................29
Fig 15 Sample email.........................................................................................30
Fig 16 Recently viewed customer List...............................................................31
Fig 17 Interface for attaching files....................................................................32
Fig 18 Google map Interaction in Project SS....................................................34
Fig 19 Fields to get user input (interval, scheduled date).....................................34
Fig 20 Date processing......................................................................................36
Fig 21 the interface page which contains a table...............................................36
Fig 22 Comparison between the previous and current manner to manage a table..........................................................................................37
Fig 23 Process interruption procedure.............................................................41
Fig 24 Server backup strategy..........................................................................43
Fig 25 Communication via secure tunnel..........................................................45
Fig 26 Gartner Survey of Software Project.......................................................53
Fig 27 Communication tunnel for requirement analysis.....................................56
Fig 28 Comparing with project quality with or without system designing............58
Fig 29 A sample database table.........................................................................59
Fig 30 The relationship between the amount of system components and development costs.................................................................61
Fig 31 Using CSS to provide ‘user help information’.........................................62
Preface

The Master of Science in Software engineering course exposes graduates in computer science and related disciplines to the many facts of this complex area – the technical, the methodological, the organizational; so that successful participants will subsequently be able to lead major projects in software engineering in many industrial and commercial sectors.

The course ends with a 6-month industrial placement, with an associated thesis report to be submitted.

The M.Sc. Courses I attended in 2009 – 2010 are the following:

- CS605 Mathematics and Theory of Computer Science
- CS607 Requirements Engineering and System Design
- CS613 Advanced Concepts in OO Programming
- CS603 Rigorous Software Development
- CS619 Programming Comprehension
- CS608 Software Testing
- CS610 Human Computer Interface
- CS615 Web engineering and Internet Solution

In a company where many different technologies are used for the many projects that are implemented experience in project management and software engineering practices among the staff are really high. However, there are a number of differences when implementing a software project according to an industrial standard against that of an academic approach. Thus, it will be informative to analyze the development of the project with reference to the material taught in the modules on the M.Sc. program. With such a comparison, the advantages and disadvantages of both perspectives will be discussed in an effort to find a balance point between the industrial and academic standards.
1. Overview of the company

1.1 Background

As part of my M.Sc. in Software Engineering, I applied for a six-month placement as a consulting engineer in Azotel Technologies Ltd. The purpose of this work placement was to develop, maintain and upgrade the company’s main software product – the SIMPLer platform. Also, tasks associated with hardware equipment and networking were carried out during my placement.

1.2 Azotel Technologies Ltd

Azotel Technologies was established following the successful sale of Amocom and the transfer of its entire R&D (including the SIMPLer platform) to Azotel in October, 2002. Since then, Azotel started to develop and maintain the SIMPLer platform as its main software product. It promoted the platform and the associated hardware and service worldwide. Azotel now equips broadband operators with the technology and automated business processes required to build out commercially successful and highly profitable networks.

Azotel Technologies have now located her head office in Co.Cork, Ireland, and have managed regional offices in North America, Europe and Africa. The software development team is running in both the Cork office and another office in Poland. Thus, the software product is designed internationally and the daily work in Azotel involves international teams.

At the moment, the software development work is mainly derived from the requirements from customers. Feature development and customization account for a big percentage of work in Azotel’s software development team. Software engineers have much experience in activities such as collecting user requirements and refining high-level specifications when dealing with customer support issues.

1.3 Overview of the product

Since the SIMPLer platform has been developed and maintained for several years as a mature software product, unfortunately, I was not involved in this project from its very beginning. When I joined the company’s software engineering team, the basic framework of the project had been deployed by several engineers that had been on the team previously. The programming language, software environment and various third-party applications had been chosen as well. I started my placement by learning the existing parts of the project and interacting with a number of software applications that I was going to use later on during my placement. I then focused on the software programming area – particularly feature development, and worked on some other assignments which involved learning a variety of skills in radio equipment, networking, and operating systems to serve the project.
1.4 Actives Covered

The extent of skills required was very broad and covered multiple fields:

- Software engineering
- Radio Equipment Interface programming (Motorola Canopy, Ubiquiti NanoStation etc)
- Network protocol & service (RADIUS, VoIP, Payment Gateway)
- Network security
- Assistance into project ownership and contraction

A short introduction to the type of skills developed is interesting as it highlights the kind of knowledge I encountered in my various contributions to the project

1.4.1 Software engineering

The application engineering is mostly involved with programming in perl, CGI, JavaScript, AJAX, and database statement for developing the SIMPLer platform.

1.4.2 Interface programming for Radio Equipment

Several items of radio equipment are involved in this project to manage its client-side network functionality. Therefore, some particular skills of interface programming are required to manage the communication between the SIMPLer platform and the radio equipment.

1.4.3 Network Protocol & Service

There are several sub projects such as RADIUS\(^1\), and Payment Gateway\(^2\) that are involved in this project to enable the platform to communicate with such network protocols & services.

1.4.4 Network Security

The company maintains a number of production servers to carry the software product, and several stand-by servers and backup servers for data recovery, and data backup. Therefore, skills about network security (based on a Unix-like operating system) and network encryption are involved in this project.

1.5 The mission

After several years of promoting their software product – the SIMPLer platform - Azotel now have a large number of clients using the platform to manage their network. With Radius and hotspot business, the number of clients has rapidly increased. Meanwhile, more and more requirements for new

---

\(^1\) Remote Authentication Dial In User Service (RADIUS) is a networking protocol that provides centralized Authentication, Authorization, and Accounting (AAA) management for computers to connect and use a network service.

\(^2\) Payment gateway is an e-commerce application service provider service that authorizes payments for e-businesses, online retailers, bricks and clicks.
features, localization and customization need to be deployed for the platform. Therefore, the software development team is focusing on implementing new features according to customer requirements, localizing the platform for newly joined users and updating the existing functionalities based on the CIs (Customer Issues).

1.6  My role

As a software engineer, my main goal was to develop, maintain and update the software product – SIMPLer platform. Besides the software programming work, there were several tasks assigned to me including the setting up of servers, maintaining hardware and software testing.
2 SS project

Being different from a standard university software project, the Project SS is implementing an industry standard in its requirements engineering, system development and software architecture design, which gave me different experiences and knowledge from what I learned about software engineering during the modules for the MSc. Compared with a standard university software project, the project SS could face a number of problems and issues during its lifecycle. However, by understanding the concept of the project lifecycle, the advantages and disadvantages of implementing a software project according to a real industry standard can be demonstrated to be much more accurate.

2.1 SS organization structure

2.1.1 The project lifecycle

After spending a few weeks on this project, I have been able to give an outline of the project structure and lifecycle based on my understanding. From a software engineering point of view, the structure of the project is not designed to be the same as a ‘formal’ software project. However, considering the particular circumstances such as the size of the company, the size of the project and the technical support, the SS cannot be said to have a well designed system structure that could be illustrated by a theoretical project lifecycle.

A complete theoretical SDLC (systems development lifecycle) will include several phases as the following shows:

- Project planning, feasibility study
- Systems analysis, requirements definition
- Systems design
- Implementation
- Acceptance, installation, deployment
- Integration and testing
- Maintenance

Fig 1 Standard software lifecycle

Since I have not been involved in the first phase, I assumed that the SS followed a standard procedure at the first stage. After that, the project SS took a different and ‘informal’ (will be discussed later) evolution compared with a
A spiral model is suitable to describe the life cycle that the project SS has gone through so far. Compared to the other life cycle models a spiral model has its own advantages that make it suitable to be used in the project SS and the other small and medium size software projects as well. The benefits are:

1. The Spiral model is in contrast to the waterfall model in that the life cycle is divided into multiple iterations rather a single pass. The Waterfall model has a rigid sequence of phases from the beginning to the end of the project. At the end the software is complete and no further adjustments can be made to it. The sequence of the Waterfall model is:

   system analysis -> requirements definition -> system design -> implementation -> integration and testing -> maintenance.

   In essence the Spiral model transfers the rigid sequence of phases of the waterfall model into multiple iterations, and therefore the risk of the project is be reduced significantly. A risk analysis will ensure that the current iteration is well managed and everything is traceable. If the process is interrupted, the project manager can make the decision to stop the current iteration in time and move on to the next one.

2. Even though the initial target and requirements for project SS were specified at the first stage, but in a realistic circumstance, there are various further requirements that can be sent to the company to update or customize the software product, that can then become the main tasks for the software development team. The Spiral model, as illustrated in Figure 2, can perfectly fit this process because every iteration is specified to contain the following steps:

   a. Requirement analysis, scheme and backup scheme preparation

   b. Risk analysis,

   c. Evaluate the scheme and backup scheme.

   d. Implementation and testing

   e. Next iteration preparation

   f. Submit the proposal for next iteration.

3. Each iteration in the Spiral model can only include one or two phases of waterfall model instead of processing all phases. In project SS, the requirement analysis and implementation are normally the two phases that are emphasized.
When discussing the life cycle for project SS, there are several shortcomings that do not fully comply with the standard principles. These are:

1. The spiral model should still comply with the waterfall model in that every phase should be covered within one iteration. However, in the project SS, the System design and requirements analysis only occupies a small percentage of time in each iteration. Due to a time/cost pressure, both the project manager and the team might skip some steps to reach the target. On the other hand, the implementation and maintenance is the most important part, which takes the majority of time during the project life cycle.

However, from a the software engineering point of view, allowing too short a time for the system design and requirements analysis are the top factors which can lead to the failure of a software project. Table 1 has a list of the relative significance of the challenging factors that can cause a software project failure. We can see that the top three factors (lack of user input, incomplete requirements & specifications, and a changing requirements & specifications) contribute by a combined figure of 36.9% to the causes for a
project failure.

![Fig 4 Gartner Survey of S/W projects](image)

2. Risk analysis, which is the most important feature in spiral model, has been considered in the project SS but not sufficiently. The intention of Risk analysis is to assess each risk during a software life cycle in two dimensions: the probability of the risk happening and the consequences if the risk does happen. In the project SS, there is no formal methodologies to support the risk analysis so that the analysis is only based on the experience of the project manager or software engineer. Therefore, due to the fact that the engineer might not be able to make complete consideration of the whole project when assessing risk, one update may result in unknown or unpredictable effects to other parts of the software. Furthermore, the hidden effects may not be spotted by the standard software testing procedures.

**2.1.2 The project staff and roles definition**

At this point in the organization of the project, the staff can be divided into four parts. The management staff, the organization staff, the development staff, the user support staff. Another important participant is the project vendor\(^3\).

The management staff is the person in charged of the project and who takes all the critical decisions about costs, and the lifecycle of project.

The organization staff is in charge of the project to formalize it through the specifications requirements. Also, they are involved in testing the product from the user’s point of view because they have more experience dealing with the project user. They are the direct contact between the engineers and the end user during the project life cycle. Additionally, the organization staff is in charge of regular administration activities such as holding the weekly group meeting, task distribution and making decisions on details.

---

\(^3\) Technically the vendor is not a part of the project staff, I mention the vendor here to explain the project organization completely.
The development staffs such as the programmer, hardware engineer, tester and system designer are the most important ones in the project. They are in charge of not only technical support but some of them are in a transition where their roles mean that they have more opportunities to meet the end user and manage the project process. Due to the nature of the product, every engineer has been involved in some hardware issues so that the boundary between a software engineer and a hardware engineer is not specified very clearly in the company. In such cases, the benefit is that the engineers are able to have experiences in multiple areas.

The user support staff members have, in most cases, transferred from the software/ hardware engineers group. These are people who have first hand experience on the project coding and product interfaces, or were in charge of the product installation and technical support. Also, they are involved in part of the project documentation such as the project troubleshooting guide, product installation guide, user interface tutorials and so on.

2.2 Development workflow

According to the project life cycle, the main features and functionalities of the product have been developed completely. The current development requirements are mostly coming from particular users and concern customization, new features and localization. Both the management staff and development engineers are involved in requirement collecting and analysis. Then the project managers will list the requirement specifications on the task list and then assign these to the engineers. The Developers should process the requirements and submit a proposal that estimates the time-required, technical skills involved etc to allow the project manager/team leader manager to determine the priority of the task and to set its deadline, and to decide the numbers of engineers involved.

---

4 Most of the software engineers are involved in taking care of the testing work themselves in Project SS.
5 Please refer to the project life cycle to make further understanding of this group’s roles and definition.
A development task could come out from the discussion at the weekly internal meeting. The meeting is organized by the project manager and the development team leader. The aim of the weekly meeting is to discuss the distributed tasks with the engineers, trace the time-spend and propose internal features for the project.

The processing of Customer Issues is another route through which a new development task could arise. The User support staff will process feedback from the user, summarize the reasonable requirements and make a proposal for the necessary TODO features.

Important updates or changes for the project are distributed by senior managers according to the company’s business strategies. When managers decide to deploy new functionalities for the project, a series of processes such as marketing research, feasibility studies and requirement analysis will be managed by the project manager.

2.3 Internationalization

Because the company has a worldwide market and business, the engineers, organizers and user support staff are living and working in several different countries, but mainly in Europe. We are using English as the main contact language, talking to each other via email and online phone. The time difference can cause some inconvenient issues but the benefit is obvious: the help desk and user support are running all day and night.

Since the product is used and promoted on a worldwide basis, a language database is created that allow end users choose their language preferences, which is an easy way to approach language internationalization. The user’s language preference can be modified through the user-interface.

The engineers have to create the language hash table firstly on the page they are working on, and then replace each text label which will be displayed in the front end interface page with the get_language function $lng{xxx}, e.g. the text label “Go Next” will be replaced with $lng{134}, Therefore, once the corresponding column in the language database for other languages such as German, French etc are translated, the displaying text will be replaced by the chosen language.
3 Project Architecture, development guide and Technical Overview

3.1 Architecture Overview

Since I did not join the project at the beginning, the architecture has been deployed already by the engineers. From my own experience, I found several advantages and disadvantages to this. I will talk about each side of this deployment, focus on the problems I found and introduce my solutions to improve it.

3.1.1 Architecture in Perl programming

The most critical problem in this project is that there is no MVC architecture used for the perl programming during the development.

Model-View-Controller (MVC) is a software architecture in which a software development process can be divided into 3 core parts: Model, View and Controller.

By using MVC, the Model (user input) and the View (application output) can be totally independent, so that the output of the application can be represented by different formats. The Controller is used to make sure that the Model and View are synchronized; that is, the output should be updated when the input changes.

![MVC prototype](image)

Advantages to use MVC:

1. Since MVC is the separation of Model, View and Controller, so one model can be reused for multiple views. For example, normally there are several different user interfaces for one web application. Users want to send/receive emails via a browser, nowadays; they also want to send/receive emails via their mobile phones. To implement this, a web site has to be able to represent both an Internet interface and a WAP interface. Using MVC enables engineers to easily update the view of the web site without changing any model or controller. So in MVC, the view does not care about the underlying model, supporting multiple file formats is easier
by just adding a model subclass for each. From a software engineering point of view, MVC represents great code reusability and extensibility.

2. The Model is separate from View and Controller, so it is very flexible to change the data source. For example, if a MySQL data source has to be relocated to an Oracle database or changing a RDBMS source to LDAP, the Model is the only part to be changed. Once the model has been deployed properly, no matter where the data comes from be it a database or a LDAP server, the presentation will be appear according to the View. From a software engineering point of view, MVC represents great encapsulation.

Following on from the advantages of using the MVC architecture mentioned above, there are also several disadvantages when building up a project SS without using MVC.

‘No-MVC’ programming in Project SS:

In terms of perl programming in the project SS, data layer coding such as database queries and presentation layer coding such as HTML / CSS scripts are all mixed up on the one page. Taking the user login page as example, the architecture of one .pl file is shown in the following flow-chart:

Fig 7 An example of the not-MVC architecture in Project SS

1. Every time the sub action is triggered by user input, the page scripts will be executed from the beginning, until it finds the proper entrance for the sub
function, which significantly increases the response time of a page. Engineers are currently trying to optimize the architecture by using various sub functions to create the simplest path; most of the scripts are executed at an unnecessary point. The only solution is to separate it into different layers by using MVC.

2. The un-separated layout, which is mixed with different scripts and programming languages makes maintenance and extension extremely difficult. Making changes and improvements to the code usually requires the engineers to trace and follow a previous engineer’s personal coding style and preference. Every single update and improvement will usually effect the entire page.

3. The un-separated layout reduced the code reusability. As Fig 6 shows, every page has its own database queries for particular purposes. Several high frequency database queries have been made into individual models\textsuperscript{6}, however, the engineers still have to make database searches on every single page.

**Why MVC is not used**

After I talked to our senior engineers about the shortcomings of our development, we found out some reasons why MVC is not used in the Project SS development.

First of all, the most obvious reason is the project architecture was not designed from software engineering point of view. The previous engineer might not familiar with MVC. This is not a fairly good reason to explain such a critical shortage in a software project but in real industry, this is a common issue. Given that there is only a small development team involved to start a small or medium size software project, one or two engineers could start the development straight away without a fully prepared system/architecture design.

Secondly, the size of the project was not estimated properly. In the case of the Project SS, it was started as a simple support platform, but because the company changed its business from hardware to software development, this platform came up to be one of the company’s main products and required various further improvements. More and more new features were deployed onto the existing architecture which makes improvement of the architecture much more difficult.

Thirdly, managers have not wanted to spend extra effort on improving the system architecture since the product is still working well. Due to financial issues and the schedule of the project, improving the system architecture would mean that the product would be out of service for a fairly long time. Considering the existing customers who are using this web product, the time for this change can never be scheduled.

\textsuperscript{6} Please see details on section 3.1.2 System component Introduction
3.1.2 System component Introduction

A good architecture will make the development much easier, but without a good MVC architecture, the programming architecture is messy. One solution is to generate individual models and packages to clean up the layout of the program. Models will be only loaded for particular purposes.

By doing this, the advantages are:

1. Code reusability and extendibility can get improved. Individual models are able to be reused over multiple pages instead of generating a new one in every single page.

2. Individual models perform better code encapsulation. After a model is loaded, engineers are able to activate this model’s functionality by calling the name of the model, and sending different parameters will lead to different result to meet their own purposes. Also, changes to a particular function will not affect those individual models. Changes to one individual model will not affect other models as well.

3. Since most of the frequently-used functions and database queries are deployed into individual models, the architecture becomes more clear and understandable.

Packages:

In terms of ‘SIMPLer’ development, modules which perform similar functionality are addressed together into a package. So when a module is required, we can load the package in the header part of a page, and then activate the module by calling the module name. For example,

```
Use packageUser;
PackageUser -> ModuleUserInfo;
```

*Load the package  *Using ModuleUserInfo by calling the module name ‘ModuleUserInfo’, this will return all the information for a user.

Fig 8 The procedure to import and use a module in Project SS

The Project SS has 4 major packages as following, wisp.pm, wisp_db.pm, wisp_web.pm, and amo_js.pm; the included modules in each package will be introduced in the following part.

Modules:

Since the reader might not be familiar with the development procedure of this project, I will divide these modules into different categories to make the introduction easier to follow. In this part, I will only discuss some of the
modules which can provide a standard experience to the reader. Modules which are developed / used for specific purposes will not be included.

1. Database modules – wisp_db.pm

This package includes common database access modules. All database related functionalities will be deployed inside this package.

a. DBI modules (Database independent interface for Perl), which is the standard database interface module for Perl. By using DBI, the application is able to communicate with one or more database sources. *DBI can provide a database interface independent of the actual database being used.*

Using DBI, it has a general set of database functions (add / modify / list / delete / execute etc), along with general database connect and disconnect functions. These should be generated independently for code usability and encapsulation purposes. Therefore, when a particular module is required during the development, the engineers should choose the corresponding package and module to access the database instead of creating repeat code on the page.

b. Customer information modules. As high-frequency used data, a customer_info module is generated for accessing customers’ information such as name, password, status, credits etc. Also, using the CGI::Session module can allow the platform to store/delete the session info.

c. Bash - database (postgresql) convert modules. To convert the argument string to a correct argument string for the postgresql insert query when the database statement has to be executed inside a bash script.

2. General_function modules – wisp.pm

General functions which will be used during the development will be deployed inside this package.

a. General functions such as date_format convert, email sending, findPath, runCommand (Unix-like OS), file_access, Ftp_service etc.

b. Regular expressions which are used for form validation are also included.

c. Database encoding functions.

3. Website presentation modules – wisp_web.pm

Include modules that are used for outputting website presentations which are generated by a combination of CSS and perl scripts.

4. JavaScript modules - amo_js.pm
All JavaScript functions which are involved in each feature will be deployed inside this package. By using the module findPath in the wisp.pm package, engineers will be allowed to separate JavaScript functions with the perl script, and use the corresponding functions by giving the right path.

3.2 Software Development Guide

In this part, a development guide about rules of coding in this project will be presented. The purpose is, to demonstrate a step-by-step guide how to make a software code update or change in this project. Also, this is to get a summary of how to process a generic software development procedure based on my experience.

3.2.1 Procedure to deploy updates/changes

1. Production servers

Production servers are used to store our software products and customer data. The software products and data sources are updated every day to get the latest version automatically by executing the update script. Data is transferred from the synchronize server to every production server.

2. ‘Synchronize’ server

The Synchronize server is the server that will be updated manually by engineers. Once changes and updates for our software products are deployed on this server, all production servers will receive the updates the day after. Also, engineers could run the last round test for their work before it is pushed to all production servers.

3. ‘Playground’ server

A Playground server is used for internal testing. Changes to the software products will be deployed on this server to get feedback from other engineers and testers.
3.2.2 Mechanics

1. Version control system

SourceJammer\(^8\) is a source control and versioning system written in JAVA, which is used to maintain the code in the project SS. The first step in making a code change is to make sure that the code on your development machine is up-to-date. In terms of developing code on a new page, this step could be skipped.

Then engineer should check out the required file in SourceJammer, which means that the other engineers cannot make changes on the checked out file until the file is checked in again. This can avoid conflicts caused by engineers changing the same files at the same time.

Version control in Team work:

If a required file is already checked out by another engineer, it may be still possible to deploy changes in a way that will not cause a conflict. If the change is small, engineers could send the change to the one who is currently using the code, and then push the change to the ‘synchro’ server without updating the file in SourceJammer. The other engineer can add the change to the file. So when the other engineers finish their work and run an update for this file, the file will not be overwritten improperly.

2. Apply all required changes

3. Test changes on local development machine

The first stage of testing requires that the changes are tested on engineer’s

---

\(^8\) More details can be found on SourceJammer’s homepage [http://www.sourcejammer.org/](http://www.sourcejammer.org/)
local development machine. If expected results appear, then the changes could be pushed to the development playground server for internal testing by other engineers and testers.

4. Add database updates and modules to ‘synchro’ server.

   If, as part of the code changes, it is necessary to add new database fields, modify existing database columns or install new modules, these must be logged on the ‘synchro’ server. After the database/module changes are updated properly in the synchronize command, they will be synchronized to all production servers automatically the day after.

5. Fix the code according to feedback

6. Upload changes to ‘synchro’ server

Once the code has been approved by another engineer, the code should be uploaded to the ‘synchro’ server.

7. Test changes on ‘synchro’ server

3.3 Technical overview

3.3.1 Overview of the software environment

   Basically, the software development team in the company is working on our ‘SIMPLer’ platform, so it is useful to introduce several specific third party software packages here to help readers gain a further understanding of this environment. Besides some specific commercial software, the company is also using some excellent open source software to deal with both the development and general issues.

3.3.2 “LAMP” environment

   LAMP is short for Linux (operating system), Apache HTTP Server, MySQL (database software) and PHP/Perl/Python, which is used in the company to build a viable general purpose web server. In this case, the LAMP is actually RedHat Linux (operating system), Apache, Postgresql (database software) and Perl. It is used by the software development team and is run on all of the servers.

   The benefits of using LAMP are:

1. Less Costs: The whole development environment is built using open-source software and an open-source operating system.

2. Flexible configuration: Open-source software can be configured for particular purposes. For example, we can easily access and modify configuration files like perl.ini and postgresql.ini.

3. Develop locally: Once the LAMP has been set up, an application can be built locally and then deployed on the web.
4. Easy to code: A web application can be created and run very quickly with Perl and Postgresql database.

However, based on my experience, the disadvantages could be

1. Software testing might be limited within a narrow range. Unexpected errors could happen when the product performs in different environments.

2. Unique environment might cause a low compatibility.

**The advantages of choosing perl in the Project SS:**

The advantages of using Perl as our main programming language are listed as follows:

1. By using Perl, it is possible to have more support for text/numeric processing and the manipulation of text files as this is a known strength of Perl.

2. Perl has good support for network programming and application development that requires database access and CGI programming.

3. Perl is a “simple” language, in our case; we need quick solutions for small problems.

4. It has good adaptability and flexibility.

3.3.3 Third party software

In this part, I will introduce several third party applications that form part of this system. Since they are all open source applications, I will list them with their names here and give some demonstration as to their real industry usage.

**Openvpn:** Openvpn is used to create a security tunnel to allow engineers or client-side technicians secure access to equipment or servers on a private network from any location with access to the internet.

In the Project SS, Openvpn is installed in a RedHat Linux system. After installation on both the server and client side, the Certificate Authority (CA) should be set up, and certificates and keys for an OpenVPN server and multiple clients should be generated.\(^9\)

**Virtual Machine:** This is a software application of a programmable machine, where the software implementation is constrained within another computer at a higher or lower level of symbolic abstraction.

**VirtualBox**\(^10\) is used in the Project SS to create a virtual Linux Operating System on another physical machines for use by the engineers. By using virtual machines, multiple OS environments can co-exist on the same computer in strong isolation from each other.

\(^9\) More details about OpenVPN can be found on [http://openvpn.net/](http://openvpn.net/)

\(^10\) More details about virtualbox can be found on [http://www.virtualbox.org/](http://www.virtualbox.org/)
SourceJammer: Source control and versioning system\textsuperscript{11}.

3.3.4 General applications

In this part, I will give a brief description of some general applications we are using in the company which basically relate to ‘Google apps’.

Documentation: Google Doc is used for general documentation. This will prevent version-difference issues for the documentation application and avoid extra costs.

Task trace: Google Calendar is used for task-tracing by staff and managers. Staff are able to self-trace their task list by editing their daily task log, and task logs can be viewed, edited and traced by managers.

GoToMeeting: A third party software for the company’s group meeting.

Skype and Email is the main system for communicating internally for the staff and for making contact with the clients.

\textsuperscript{11} Please see more details about SourceJammer on section 3.2 Mechanics, part 1 – Version control system
4 The Perl client

In this chapter, I will explain what I did in this project, and I will focus on some sub-projects that can represent a standard issue/solution in that area. Details for the solution will be given as thoroughly as possible.

My major implementation in this project is ‘feature development’, which is to develop Perl and JavaScript applications to serve the ‘SIMPLer’ platform. Applications are normally required from the client side, but critical changes on the platform are deployed internally through the group meeting.

Besides programming, my work will cover some other areas which include system security, network protocols, hardware installation and software maintenance documentation.

4.1 The perl application

The application has been split into small projects and assigned to every engineer in the software development team. Here are listed the tasks assigned to me, each task can be considered as the main component of that feature.

4.1.1 Contact book enhancement

This project is my first task, which is implemented by using Perl and JavaScript. The aim is to enable adding multiple contacts, and editing contact details for each customer.

Technical Overview:

To enable users to perform their operations in the same page, JavaScript is used here to grab user’s input strings, display a string which is composed from user’s inputs from each field (Name, Phone, Role and Email) in the dropdown list. A Java function will be activated when a button such as ‘Add’, ‘Delete’, ‘Edit’ and ‘Set as preference’ is clicked.

Feature Description:

---

12 In this system, each user has specific ‘user rights’ which will manage users’ behavior. In the following part, ‘user rights’ will be ignored unless needed in particular circumstance.
When the user fills their contact information into the contact list as shown above, information will be combined into a string and displayed in the dropdown list when the ‘Add Contact’ button is clicked. Buttons such as ‘Delete’ and ‘Set as preferred’ will target to the current chosen content in the dropdown list. By clicking ‘Edit Contact’, a new page will be displayed as follows:

Fig 11  Edit Contact Page

Behind the Scenes:

Form validation based on JavaScript will be performed once forms are submitted. Several checking functions such as checking for a duplicated username, Phone number etc. will be performed as well.

Critical Points:

Data transaction and data storage is very critical in this project. To store the user’s input into the database, the procedure is shown as follows:

<table>
<thead>
<tr>
<th>Combine the input in separate lines into a string.</th>
<th>Name, Phone, Role, Email</th>
</tr>
</thead>
<tbody>
<tr>
<td>Convert the string to fit the data format of the database.</td>
<td>Name:::Phone:::Role:::Email</td>
</tr>
<tr>
<td></td>
<td>Name1:::Phone1:::Role1::Email1</td>
</tr>
<tr>
<td></td>
<td>&gt;&gt;Name2:::Phone2:::Role2:::Email2;...</td>
</tr>
<tr>
<td></td>
<td>{Name1:::Phone1:::Role1::Email1 &gt;&gt;Name2:::Phone2:::Role2:::Email2;...}</td>
</tr>
</tbody>
</table>

*Note: This piece of data is stored as a ‘text[]’ type in database, so ‘{ }’ is added to fit the format.

Fig 12  Data transaction (To database)
To retrieve the data from database and display the data the following sequence
4.1.2 RT Integration

The aim of this project is to enable the ‘Automata Email Sending’ function in ‘SIMPLer’ platform interact with the ‘RT’ system which is part of our client’s platform.

Even though I kept contact with the client side manager to confirm their requirement before the implementation, the specification of this project was changed at least 3 times. This resulted in a long development period.

Technical overview:

It uses the Perl:: MIME::Lite Module to enable automatic Email sending. It approves the content of the email for the interaction with ‘RT’ system. JavaScript is used to collect multiple email addresses.

Feature description:

From the ‘SIMPLer’ side: Once a maintenance ticket is created in ‘SIMPLer’, an email will be sent automatically to the pre-defined email addresses.
From the ‘RT’ system side: to interact with the ‘RT’ system, the sent email will contain information as follows.

Critical Points:

Using CPAN Module Perl::MIME::Lite to implement Email sending functionality in Project SS. This module can be used as the following example shows:

Use MIME::Lite;

# Create a new message
My $msg = MIME::Lite -> new (  
   From   => 'testSupport@azotel.com'
   To     => 'testCustomer@azotel.com'
   Cc     => 'testOperator@azotel.com'
   Subject => 'test message'
   Data => 'This is a test report.'
);

# Send out message via default
$msg->send;

4.1.3 History list of recently viewed customers

This project became such a popular feature in the ‘SIMPLer’ platform that I had to keep working on this project to create a template for this function.

The most interesting point for this project was the creation of the template. In our case, this function is highly required from our clients, and through the communication with our clients, my manager and I figured out that if several features which will represent the similar functionalities could be deployed in ‘SIMPLer’, the usability will get be significantly improved.

Development Key idea:

The key idea is to add a timestamp to the database when a user accesses the entry of a customer detail page. Meanwhile, a database query will be executed to get the customer id and the timestamp that are then displayed
ordered by timestamp. This list shows the user the customer’s name and the last accessed time.

**Feature Description:**

In SIMPLer, an operator can view, modify, add, and delete their customers’ details. In this case, the customer detail page is a reasonable entry to trigger the recording of a customer’s actions, which means that every time this page is accessed / refreshed, a timestamp will be recorded in the database.

![Fig 16 Recently viewed customer List](image)

**The Template Module:**

From a software reusability point of view, an application which is deployed regularly could be developed as a template. The template could be reused in the future for implementing various different purposes. In this case, the improved template has a user friendly interface to receive various parameters. By changing the value of the parameters, the template can record particular user actions, and display a list of user-action-history, such as the recently viewed customers.

To use this module, the *template module* should be loaded first, and then by calling the template name and passing proper parameters, this application will be activated to record particular user actions. For example,

```php
use module;
$my_module = new module;
$my_module->recentlyViewed($param1; $param2; $param 3));
```

**CSS issue in this application**

According to the application requirement, a floating box should be displayed in every page’s bottom right corner. To implement this, CSS(Cascading Style Sheets) are used here to control the position of the display box. However, a CSS compatibility issue appeared during the development.

Even though W3C has been working on promoting a standard CSS architecture worldwide, with various browsers support for the W3C CSS standard still needs to be improved. For example, a CSS code *Position: float* is used to fix the position of the displaying box. However, some versions of ‘Internet Explorer’ do not support it very well so that extra pieces of code have to been added against Internet Explorer and some other browsers.
4.1.4 Site/Equipment enhancement

Several new technologies are involved in this project, which are:

1. Using CPAN module `Net::FTPSSL` to allow a user attach files via an interface

2. Using the Google Map API to obtain a GPS coordinates when an address is entered into the address fields, and then displaying a position marker in Google Maps.

3. Using crontab (Linux task schedule application) to implement customer maintenance events.

4. Using JavaScript to create a table, add and delete a row in the table, modify the content of a cell and store the data.

4.1.4.1 Using CPAN module `Net::FTPSSL` to allow a user attach files via a interface

Feature Description:

![Interface for attaching files](image)

A user can attach files via a user-friendly interface to his/her account. To deploy this, a module from CPAN `Net::FTPSSL` is used to implement attachment management. This module is a class implementing a simple FTP client over a Secure Shell Layer (SSL) connection written in Perl. The basic usage is:

```perl
$ftp = Net::FTPSSL->new(FTP_Address, Encryption=>'E',Debug=>1)
```

To implement the attachment feature using `Net::FTPSSL`, the key point is to use the CPAN module `File::Basename` with the file name (file path) as a regular Expression. The file name can be processed as follows,

```perl
my ( $name, $path, $extension ) = fileparse($filename,’”[^.]*$’);
my $file = $name.$extension;
$name =~ tr/ /_/;
$extension =~ tr/ /_/;
my $safe_filename_characters = “a-zA-Z0-9_.-”; # file name constraint
```

---

13 Encryption – The connection can be implicitly(IMP_Crypt) or explicitly (EXP_CRYPT) encrypted. In explicit cases the connection begins clear and became encrypted after an “AUTH” command is sent.
Debug – This set the debug information option on/off.
$name =~ s/[^$safe_filename_characters]/\g{0}/g;
$extension =~ s/[^$safe_filename_characters]/\g{0}/g;
Then $filename = $name.".".$extension;

Then the processed file name can be used to represent the path of the attached file, and can be used when adding or deleting attachments.

**File extension issue:**

A minor issue was spotted during the testing which was that the suffix of a file might be changed among the different software versions (e.g. MS Office word 2003 will take .dox as the suffix but word 2007 will take a .docx as the suffix). Therefore, a different file suffix will cause a fail when downloading the attached file from the FTP server because the different suffixes will be considered as two different files (different paths). However, this is one of those “beyond the code” issues, I suggested that this attachment interface takes only one standard file extension against software which has multiple suffixes possible in their updated version. However, this is not a reasonable solution from a user-convenience point of view.

**4.1.4.2 Google Map Interaction**

**Feature description:**

A Google map app has been used on certain pages in the platform to provide a more intuitive experience when the addresses or GPS co-ordinates are displayed. On top of that, the Google map app can also automatically provide the GPS co-ordinates when the addresses are recorded in the platform so that the proper co-ordinates will allow the platform to provide the address information more accurately. Some related features which depend on the GPS co-ordinates such as network coverage are then possible to implement.

By investigating Google map APIs, the front-end interface is generated as the following image shows. When user types in or changes the address information in the address fields, a confirm box will pop up to allow the user confirm their inputs. Once OK is clicked, the background process will query the GPS co-ordinates from the Internet resource and then display the proper results of the co-ordinate. The user can also drag the marker to adjust the location.
Development key ideas:

First of all, the background process has to monitor the address fields to get the address information. Once the value of the address is updated, the 'onChange' function will be activated to submit the updated address information to the resource. If the inputs are valid, the getLatLng function can be called to detect the user’s co-ordinates information.

4.1.4.3 Using crontab (Linux task schedule application) to implement customer maintenance.

Feature description:

A feature in this project requires that the platform can manage customer maintenance events based on the maintenance interval. This could be quarterly, semi-annually, or annually. To implement this application, a crontab script will be executed on the server every day to schedule the events by processing the current date and scheduled date.

---

All related APIs are offered by Google maps JavaScript API. See details on http://code.google.com/intl/en/apis/maps/documentation/javascript/basics.html
For example, in Figure 18 the maintenance interval is specified as ‘Annually’ and the Next maintenance Date is specified as Nov, 15th, 2010, and the automate Maintenance ticket is enabled. Then on the day Nov, 15th, 2011, which is the scheduled next maintenance date, a maintenance ticket will be created for the site automatically, and the Next maintenance Date will be updated to Nov, 15th, 2012 automatically.

Cron & Crontab script:

Cron is a time-based task scheduler in Unix-like computer operating systems. Cron enables users to schedule jobs (commands or shell scripts) to run periodically at certain times or dates. In this case, auto-updating the data information (next maintenance date) will be activated by monitoring the due date and current date.

Crontab usage:

```
* 6 * * * root /usr/bin/perl .../.doUpdate.pl > /update.log
```

<table>
<thead>
<tr>
<th>①</th>
<th>②</th>
<th>③</th>
<th>④</th>
<th>⑤</th>
<th>⑥</th>
<th>⑦</th>
<th>⑧</th>
</tr>
</thead>
<tbody>
<tr>
<td>minute (0-59)</td>
<td>hour (0-23)</td>
<td>day of month (1-31)</td>
<td>month (1-12 or jan,feb…)</td>
<td>day of week (0-6 Sunday = 0/7 or sun,mon…)</td>
<td>user name</td>
<td>command to be executed</td>
<td>output log file</td>
</tr>
</tbody>
</table>

The cron script will be executed on the servers 6 a.m. every day, the datetime processing procedure is showed as follows:

4.1.5 An enhancement for interface usability

Tables are used regularly on a web interface page in the project SS, but since a MVC architecture is not used in this project, when a table is required on the interface page, the engineers are still using pure perl scripting to manage a table's layout. Thus, the interface page will be reloaded when a user adds or deletes a row in a table because the add/delete operation will reloaded in the page to perform a form submission. From a usability point of view, a dynamic way to manage the table would offer a better user experience than a forced page reload.

*Note 1. Using datetime module to process datetime math
*Note 2. Using keyword ‘interval’ (postgresql) to update next scheduled date. E.g. Current_date + interval ‘1 year’, current_date + interval ‘2 month’
creation and manage the rows/cells in the table to improve the user experience. This is now promoted in the team as the standard manner to manage a web interface page which contains a table.

A brief introduction to the JavaScript function:

Elements of HTML DOM and JavaScript are used here to add or delete a row of a HTML table and manage the content of the cells. The HTML DOM defines a standard way for accessing and manipulating HTML documents. By using the combination of DOM and JavaScript, any modifications for a HTML table can be processed within the current page to avoid a page refresh.

To add a row in a table, we can use the DOM method `table.rows.length` to get the length of the current table. Therefore, every time the ‘add’ button is clicked, the addRow function will get the length of the table and insert a row as the last row in the table. To create cells in the row, the cell type will be defined firstly as any of the valid HTML elements such as ‘input field, radio’ etc. Also, the name/id of the cell element can be defined in a sequence (row_0, row_1, cell_0, cell_1...), so that the content of the cells can be accessed by the DOM method `getElementById()`. Then, the DOM method `document.createElement(type)` is invoked to create the cell and `appendChild()` to insert the cell into a row.

Deleting a row is a bit more complex than adding rows because the management of the name/id of the cell elements has to be considered when deleting a row in the table. The problem of deleting a row can be shown by example: if there are 3 rows in the table, the ids are row_0, row_1 and row_2. After deleting row_1, row_0 and row_2 will be left but with unsequenced ids. Thus, it becomes impossible to access the cell element by using the DOM method `getElementById()` because the script could not determine which row has been deleted. In order to solve the above problem, a re-order function can
be used to go through the table and rename each row in the table while the delete function is being called. Therefore, when the ‘delete’ button is clicked, the script will firstly delete the current row, then reorder each row in the table to keep the name/id in the correct order.

Based on the discussion above, each new added row will be assigned with an id in alphabetical order. The ids will be re-assigned for each row in the table after a row is deleted. This enables the DOM method `getElementById()` to process each cell element by simply looping through every element in the table.

4.1.6 Export customer report

Data reporting is a very important feature in the Project SS for both the administrator and the end user. Tracing data and generating reports are necessary in every sub project, e.g. generating customer invoices, customer billing, and operator management. The platform can display a variety of information and provide a user-friendly interface for accessing the data. However, the web interface only is not enough and for several particular tasks the product has to perform a data collecting and reporting functionality to meet the user’s requirement.

**Feature description:**

Generating a file export template `wisp->export`, the template should be able to collect required data from the database and export a file to represent the information in a proper format (that is, .xls, a Microsoft Excel file).

**Technical overview:**

The CPAN module `Spreadsheet::WriteExcel` is used to create the file-export template. Another CPAN module MIME::Lite is involved to send the exported file to a target email address as an attachment.

A brief example of the module `Spreadsheet::WriteExcel` usage of is introduced as following:

```perl
# Load module
use Spreadsheet::WriteExcel;

# Create a new Excel workbook
my $workbook = Spreadsheet::WriteExcel->new('perl.xls');

# Add a worksheet
$worksheet = $workbook->add_worksheet();

# Output data to the worksheet
$worksheet->write();
```
4.1.7 Linux process interruption (Software-level)

Introduction to Process interruption – based on ‘SIMPLer’ platform

Process interruption means a background process which is activated by user such as downloading, exporting various data, etc that will be controlled by a process checking script for their start/end time and the duration of execution duration. The reason to monitor the background process is that the execution of multiple processes at the same time will cause a significant reduction in server performance. An even worse outcome is that the physical hardware will stop working because of overloading. Therefore, monitoring background processes will help the platform manage concurrent processes and improve hardware performance. Process interruption includes monitoring process start/end time, making restrictions against concurrent processes that are activated at the same time, and removing restrictions when necessary.

Case study:

Recently, one of the production servers was reported to exhibit unstable performance. After investigation our engineers noticed that the reason to cause such low performance and long response time was that, on the one hand the MRTG (Multi Router Traffic Grapher)\(^\text{16}\) took a large amount of memory when generating the Grapher. However, this was not totally beyond our expectations when considering the increasing number of customers on this server. Updating the physical memory and making further configuration adjustments to optimize the application could be the solution. On the other hand, when several concurrent processes are being executed on the server, the hardware is extremely overloaded so that the server is not able to give a quick response, and sometimes even stops working. Therefore, how to monitor the background process and make an interruption when multiple concurrent processes are being executed by the user lead to the following issues that arose during discussion:

1. The solution will be based on which layer? Performed by the platform itself or performed by Linux commands?

After a team discussion, we decided to perform a process check associated with the platform but not Linux commands. The benefits were:

a). Portability: Performing the process interruption as a perl script rather than a Linux command will improve the portability of the platform. Implementing the functionality in Linux commands will cause a failure when the Operating System get changed in the future.

b). Easy to approach: A Linux-like operating system provides various interfaces for the end-user to approach process management according to their own purposes, the process management command such as \textit{sar, ps},

\(^{16}\) MRTG is a Third Party application used for monitoring and measuring the traffic load on network links.
etime\textsuperscript{17} etc could help the user access the process information in detail. However, considering our own case, deploying the process interruption at a script level would be much easier. Also, a Linux command solution would take our engineers much more time as they would have to investigate shell programming and Linux commands usability. Thus, they could not be deployed as a quick solution for the effected server.

c). More efficient: As we know, any Linux command will more or less occupy some cpu and memory resources itself, especially when executing several processes to monitor the system all the time. A script level process checking is easier to manage, causes less cost of system recourses and is more efficient.

2. How to stop some similar issues happen again in the future?

Rather than deploying a process interruption functionality, the script code which is triggered by a background process such as downloading, generating an invoice etc has been updated to output log files if it is not logged already. Log files will be analyzed periodically to report any unexpected data so that engineers will receive notice of such issues in time.

**Development Key idea:**

For the project SS itself, the process interruption can be deployed as following:

Taking the downloading process as an example, when this process is activated by the user, a process interruption procedure will be executed at the beginning of the downloading script page. First of all, it will require a piece of information from the database to check the ‘process entry’. If there is no entry in existence, then it will perform the downloading process. However, if there is one already, then it will interrupt the current the process, display some warning information and pause current process. The process will return to being live until the corresponding ‘process entry’ is removed from database.

It is worth mentioning that because the execution order of perl is in a plain sequence, such as with the C language, from the beginning to the end\textsuperscript{18}, we can therefore address the process interruption at the beginning of the script. The perl compiler will ensure that the process checking will be executed every time when the script is activated.

\textsuperscript{17} Linux command, Sar: output information of cpu&memory utilization. Ps: process monitoring and management. Time: output process executing during.

\textsuperscript{18} In case that OOP programming is not used in perl.
4.2 Non-programming Task

Since I was involved in several tasks that were beyond software programming such as system data backup and disaster recovery, and system security and network security during the placement, I will talk about these tasks briefly in this chapter to explain the breadth of my experience completely.

4.2.1 System Data backup and Disaster recovery

4.2.1.1 Introduction to data backup and disaster recovery

Nowadays, data is the most sensitive and critical resource for an IT company. Without a good system backup procedure and disaster recovery plan, a software company might face to a real disaster when a natural or human induced error that damages data happens. I took part in making the system backup and disaster recovery plan for the company during my placement, and I will present what I learned in this area.

4.2.1.2 Common Data protection strategies

To protect their data, various strategies have been deployed by the companies to approach a perfect data protection plan. The most common ones...
are\textsuperscript{19}:

1. Backups made to tape and sent off-site at regular intervals.

2. Backups made to disk on-site and automatically copied to off-site disk, or made directly to off-site disk.

3. Replication of data to an off-site location, which overcomes the need to restore the data (only the systems then need to be restored or synced). This generally makes use of storage area network (SAN) technology.

However, these strategies above may not be able to guarantee perfect data protection, and most obviously, once a data recovery is required, issues such as re-building the OS, installing various drivers and applications, and environment configuration will cost lots of time and money. Therefore, two critical questions still need to be addressed:

1. How to get a guaranteed procedure for data protection.

2. How to implement a recovery plan as quickly as possible.

\textbf{4.2.1.2.1 Data protection strategies and hardware structure in Project SS}

\textbf{a) Various backups to guarantee Data Correctness and Consistence}

We are using a combination of strategy 1 and 2 above to generate various backups for critical data, which are:

1) Each production server will be arranged to choose another one as its backup server based on the location\textsuperscript{20}.

2) Each production server will make a daily backup to a different partition on its own hard disk.

3) Each production server will make a daily backup to the backup server.

4) Each production server will make a 30-day increment backup to backup server 2.

\textsuperscript{19} http://en.wikipedia.org/wiki/Disaster_recovery

\textsuperscript{20} The company’s servers are addressed all over the world, so basically if there are two servers in one country, they will backup to each other to reduce data loss during the transfer and utilize the fastest network.
b) Using RAID (Redundant Array of Independent Disk) to improve Fault Tolerant and Stock Rate

RAID is a technology that combines multiple relatively low-cost, less-reliable disk drive components into a logical unit where all drives in the array are interdependent. In the company’s backup strategy, we use an open source application called RAIDTools\textsuperscript{21} to combine two SATA drives into a RAID1 scheme on each backup server. The purposes and advantages using RAID1 are:

1) A RAID1 scheme provides the ‘mirror’ functionality. Data which is written into one disk will be 100% copied to another disk. When the data is read, the RAID 1 will be the primary data source. If a failure occurs, the RAID controller will skip disk 1 and consider disk 2, which is the backup disk, as the data source. So, the RAID 1 scheme will allow for one disk to fail while still operating fully.

2) All critical data (databases, logfiles, configuration files, etc.) is backed up periodically to two backup servers which are using a RAID1 scheme. Two backup servers can provide separate physical locations and RAID1 can provide separate physical hard drives.

c) Maintaining Standby servers to provide a hot-swap scheme

\textsuperscript{21} A Linux like OS built in application, opensource.
The company maintains several standby servers to be hot-swapped in for any failed server. Standby servers have the latest software loaded and are fully configured by the engineers. When a failure occurs, the client-side technicians will be able to process a hot-swap between the failed server and the standby server following the company’s disaster recovery document. No professional experience is required.

4.2.1.2.2 Data backup

After an enhancement to data backup procedure, it can now be processed by executing a shell script rather than following the original step-by-step data backup guide. With a simple configuration for both backup server and production server, the script will provide a user-friendly interface to help engineers or client-side technicians manage the whole backup process. This is so that the backup procedure can be maintained automatically by executing the script at a scheduled time.

4.2.1.2.3 Data recovery

a) The data recovery procedure has also been improved by following the guide to execute the data restore script. The script is able to provide a user-friendly interface displaying the backup-data source list, auto-creating log files and handling errors.

b) The following post recovery actions will be executed to verify the data restoration process.

   1. Verify that partition where the data is stored is encrypted.
   2. Verify that the recovered software is performing as expected.
   3. Ensure that the recovered server has the latest software running on it.
   4. If the recovery process has executed correctly, any temporary folders need to be removed.

4.2.2 Network Security

4.2.2.1 Introduction to Network security

The company is using Linux as the operating system on the server side. Linux, as an open source OS, has various advantages in its functionality and costs. However, because it is open source, there are some security risks if the Linux OS is not configured properly. Here, I will talk about the security issues I met during my placement and the most common and useful solutions to configure the OS for an enterprise-level application.

4.2.2.2 Communication via secure tunnel

A VPN (virtual Private Network) is a computer network that uses a public

---

Please refer to chapter 3.3.3 Third party software: Third Party Application OpenVPN
telecommunication such as Internet to create a secure and stable access to the organization’s network. The company is using a third party application called OpenVPN to build a VPN tunnel communicating between equipment, servers, local networks and business partners.

The secure tunnel created by using OpenVPN allows engineers or client-side technicians secure access to equipment or servers on a private network from any location with access to the internet. With a proper firewall configuration and SSL (Secure Sockets Layer) encryption, the network security is improved to a commercial level.

Fig 25 Communication via secure tunnel

1. The only way to gain access to any of the production servers (equipment) is connect to the access server (WIB).
2. Each connection is secured by SSL encryption.

4.2.2.3 Firewall & IP table configuration based on Linux OS

Iptables is a built-in firewall that is normally installed by default on most of Unix-like Operating Systems. The default iptables allows all incoming and outgoing traffic. Therefore, we need to configure the iptable rules to make it perform in a desirable way.

Referring to Fig 25 that illustrates communication via secure tunnel, the iptable is configured according to the data flow. For both head office and regional office network, the restrictions will be the same, internal communications are allowed among individuals on the office private network. Access to production

---

23 http://en.wikipedia.org/wiki/Virtual_private_network
24 Only general configurations will be discussed here due to confidential reasons.
servers is only allowed through the access server. Access to equipment is only allowed through an individual WIB. Connections from business partners are only allowed if and only if they are allowed in the Access server’s iptables. Moreover, ssh is allowed in each server for a service account to login but not for the root. Other services such as FTP, telnet, email etc are enabled / disabled according to their particular purposes.

Iptable rules will be defined as following format:

```
iptables -A -i eth0(eth1) -p tcp --dport 22 -j ACCEPT (DROP)
```

① parameters, ② eth0: local network, eth1: public network ③ port, ④ accept/drop statement

For example,
```
nftables -A -i eth0 -p tcp --dport 22 -j ACCEPT & nftables -A -i eth1 -p tcp --dport 22 -j ACCEPT will open port 22, which is the port for ssh service.
nftables -A -i eth1 -s 192.168.1.0/24 -j DROP will block any particular connections from subnet 192.168.1.0 – 192.168.1.255
```
5 Software Testing in Project SS

During my placement, I have been working in the software testing team for testing a sub project. The main goal was to improve the existing software testing procedure. In this chapter, I will present my experience during this period in a real industrial software testing environment.

5.1 Review of previous Testing procedure

For a web-based project, testing and validation is an important and challenging work. Being different from the traditional software testing methodologies, a web-based testing procedure has to perform not only functionality testing, but also needs to perform a front-end testing for the users’ visual experience due to the availability of various browsers and operating systems on the client side. The most important testing in the project SS was for security, accessibility and usability for end-users. However, the unpredictable nature of user behavior significantly increased the difficulty of testing for web-based applications.

Recently, a critical software error occurred after an upgrade for our main software product. The managers decided to pay more attention to this by deploying an enhancement for the existing software testing procedure which has not been updated fully for a while. The existing procedure had covered most of the test cases and had been updated for several times during the development according to the troubleshooting guide. But unfortunately, it had not been formalized sufficiently. Various test cases were dispersed throughout the testing procedure which lead to a less efficient, discontinuous and un-traceable approach to test. Therefore, it is necessary to make an enhancement to the existing procedure, which will focus on test case formalization and test flow standardization.

5.2 Testing for Web based applications

First of all, traditional software testing normally takes months, even years. However, there might be only several days available to design, implement, and deploy a section of a web application. Correspondingly, web-based software testing also has a shorter lifecycle and faster implementation time.

Secondly, traditional software testing will pay more attention to functionality testing; both white-box testing and black-box testing used to verify the operation of functions rather than other aspects of the software. Drawing on my own experience, the code/scripts that are implemented for web applications are more practical and straightforward than standard software that is developed using Java or C++. Therefore, although functionality testing still plays the most important role during the whole testing procedure, less attention is paid to it.
Thirdly, different methodologies are used in a web-based software testing scenario. Since these pay attention to different aspects, the most popular methodologies used in traditional software testing, such as boundary testing or partitioning testing in Black box testing, or branch testing and statement coverage in White box testing, would not produce useful test results when being implemented for testing web applications. On the one hand, this is because of the use of different programming languages (C++ or Java) versus scripts (Perl or PHP). There are significant differences in the programming architecture, language structure and data storage. On the other hand, high code coverage is much more difficult to achieve when implementing web based testing. An industry standard web application is normally associated with a database module, third party applications and scripts such as HTML/CSS. The scripts for web applications do not have the same consistency as traditional software code, and thus the percentage for code coverage could be much lower in a web based testing scenario.

5.3 Introduction to the new testing plan

After a few days study and investigation, we generated a new software testing plan based on the original one. Also, two documents were updated along with the analysis of various test cases, called the Trouble-Shooting Guide and the Software Development Guide. These three documents present the software testing procedure, trouble shooting procedure and solutions, and detail how to avoid common errors and bugs during development. In the following part, I will pay attention to the new testing plan.

In the new plan, the testing procedure is divided into 6 categories:

1. Functionality test
2. Load & Stress test
3. User interface test
4. Compatibility test
5. Security test

5.3.1 Functionality test

5.3.1.1 Form Validation

Forms are a very popular front-end interface in web applications. These are where user specified information is submitted, such as registration, login etc, Form validation has to be performed to verify the consistency and validity of the information, and prevent potential attacks such as an SQL injection\textsuperscript{25}, and Cross-site scripting\textsuperscript{26}. Form validation is usually performed using JavaScript and Regular Expressions, or a combination of both.

\textsuperscript{25} Please see details in http://en.wikipedia.org/wiki/SQL_injection

\textsuperscript{26} Please see details in http://en.wikipedia.org/wiki/Cross-site_scripting
Since form validation is performed regularly, a JavaScript function has been generated to provide generic validation performance, such as verifying email addresses, mandatory fields, trimming, and arrayToString operations. Regular expressions are more convenient for specific validation during development.

Example of Using Regular Expressions in JavaScript

Taking date checking as an example, when we want to check the submitted value meets the date format or not an isDate checking will be performed as follows when the value is submitted:

1) Using the DOM to get the value, object1_val.
2) Define a Regular Expression to check the desirable format, e.g.
   \n   Var exp = /^([0-2]?\d[0-1])\[/(0?\[1-9]\|1[0-2])\[/\][0-4]$/;
3) Using a JavaScript build-in function test() to check the if value follows the above expression or not. e.g.
   exp.test(object1_val) will return true(if the pattern is being followed, otherwise false)

5.3.1.2 Code / Script testing

5.3.1.2.1 Version control

A different version of programming / script language could cause serious problems in both server side and client side. For example, using different versions of HTML and CSS standard will cause a display issues, and different jQuery versions will cause different usages for built-in functions so that the functionality could fail to perform in a desirable way.

5.3.1.2.2 Black box testing

In a similar manner to its use in traditional black box testing, several common methodologies are used to perform black box testing here. By generating test cases, the submitted value and the possible output can be verified. The types of black box testing methods available are

Equivalence Partitions: These provide a minimum level of black box testing, at least one value has been tested from every input and output partition using a minimum number of test cases.

Boundary Value: at the expense of approximately twice the number of tests, the minimum and maximum value of each Equivalence Partition is tested at least once using a minimum number of test cases.

Truth Table: Using Cause-Effect Graphing and Truth Tables to identify a minimum subset of possible combined conditions, so that all different behaviors of the function will be tested. In a typical circumstance the Causes Effect Graphing will generate a very large number of test cases for the function. So, to minimize test cases, some conditions might be omitted when
considering a specific function.

5.3.1.2.3 White box testing

As discussed above, Statement testing, Branch testing and other white box testing methodologies which are used in traditional software testing may not be able to produce reasonable testing results. Also, the percentage of code coverage could be very low for a typical function in web based testing area. Therefore, it is worth to run white box testing for some specific functions but not for all of them.

5.3.1.3 Link validation

Link validation is one of the most important aspects in web application testing. Link validation can be processed as the following steps, 1. verify that all links are defined with the right path, 2. the right page will be approached by clicking the link, that is, the target page exists in the web system. 3. verify that there is no isolated page in the web system, which means that each page can be approached by clicking a link.

Link validation can be processed automatically by third party tools. The theory is that links on all of the front-end pages will be analyzed via reading the HTML script (normally processed by perl script). The page – link associations will be logged as output, and also, isolated pages will be recognized because the target page has no links associated with it. However, the automatic link validation can not tell that the page–link association meets its initial purpose or not. This step still needs to be processed manually.

5.3.2 Load & Stress test

Load & Stress testing is actually measuring and monitoring server performance. Since server performance makes a direct impact on user experience and the response time of the web applications, when the server is working under a heavy load, it can easily lead to a hardware disaster, even a crash. Therefore, load & stress testing will generate useful information about the maximum load & stress that a server can take. The web application administrator can decide then if it is necessary to limit the number of users and the amount of applications available on the web system.

Load & Stress testing can be processed automatically by third party applications such as ab (Apache build-in testing tools) and openSTA (system architecture testing tools). Load testing will focus on testing the boundary data such as the maximum/minimum load, Stress testing will more likely use some designed test cases to make the system crash on purpose.

5.3.3 User Interface test

A well-designed user interface will significantly improve the user experience when interacting with the facilities of the web based product,
though the meaning of well-designed more depends on a user’s “preference”. In the project SS, user interface tests are normally processed by the staff who have good experience in marketing, end-user requirements and design, so that they can give the engineers more accurate details before they make any enhancements. User interface tests will also cover testing with various popular browsers. Problems such as the CSS issue, mentioned earlier, have to be rectified. A history list of recently viewed pages will be reported to engineers for fixing. Also, usability and accessibility testing will be processed manually as well to achieve a better performance.

5.3.4 Compatibility test

Compatibility testing includes the monitoring of the performance of applications when running on different operating systems such as Windows, MacOS, and Linux. It is an essential step when moving the software to a new server, or running a brand new server. Though all production servers are maintained by the company itself, the team manager has decided to spend more time on improving the compatibility of the software product, which will enable all applications to interact with more operating systems.

5.3.5 Security test

Processing a Security test for a web based software, especially for the web based product which performs functions such as exchanging information, online shopping etc, is a critical step before the product is released to the public. As I know, we are using some third party applications to carry out the security testing, but unfortunately I was not involved in this part.
6 Industry standard and Agile software development

Modern Software engineering has developed over the years with various developed methodologies. However, in the software industry lots of individuals and companies are still developing their software according to their own styles with no concern for modern software engineering approaches or technologies. Even though the “programmer” has been replaced by the “software engineer”, and the latter has become a more popular title in the software industry, many development approaches such as software prototyping, software quality management systems and various applications has been applied successfully in the industry; there is still more research and investigation that needs to be done in software engineering for real industrial practice. The industry needs to be supported by a more mature and more precise software engineering theory.

In this chapter, I will review the M.Sc. lectures and discuss the lectures which were related to the project I was involved in. What is more important is, I will try to figure out the difference between a “academic standard” software project and the “industry standard” one. Considering what I learned from the M.Sc. academic studying, there are several critical points hiding in the project SS that may cause a series of failures at some stage. Through the following discussion, several common factors that exist in the software industry which can lead to software failures will be presented so that I, or any other software engineer, should be aware of during their development career. In the end, I will introduce my understanding of agile software development – which is a methodology to approach a better match between the industry standard and the academic standard.

6.1 Requirements Engineering and System Design

6.1.1 Overview of the Requirements Analysis and System Design

The target of the requirements analysis is to generate a high level specification for the software project and to make significant refinements to the specifications through a requirement engineering process. The final goal is to ensure that the created software product is able to meet the customers' needs. Therefore, requirements engineering is such an important aspect which can always determine whether the software project will be get a success or failure. The Gartner Survey of S/W projects shows us in Figure 25 that requirement related factors are responsible for 28.9 percent (15.9%+13.0%) of the total that go into making a successful software project.
Table 1: Project Success Factors and % of Responses

<table>
<thead>
<tr>
<th>Project Success Factors</th>
<th>% of Responses</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. User Involvement</td>
<td>15.9%</td>
</tr>
<tr>
<td>2. Executive Management Support</td>
<td>13.9%</td>
</tr>
<tr>
<td>3. Clear Statement of Requirements</td>
<td>13.0%</td>
</tr>
<tr>
<td>4. Proper Planning</td>
<td>9.6%</td>
</tr>
<tr>
<td>5. Realistic Expectations</td>
<td>8.2%</td>
</tr>
<tr>
<td>6. Smaller Project Milestones</td>
<td>7.7%</td>
</tr>
<tr>
<td>7. Competent Staff</td>
<td>7.2%</td>
</tr>
<tr>
<td>8. Ownership</td>
<td>5.3%</td>
</tr>
<tr>
<td>9. Clear Vision &amp; Objectives</td>
<td>2.9%</td>
</tr>
<tr>
<td>10. Hard-Working, Focused Staff</td>
<td>2.4%</td>
</tr>
<tr>
<td>11. Other</td>
<td>13.9%</td>
</tr>
</tbody>
</table>

Good system design is the way to help the software project achieve its requirements. Software cannot be developed on the fly. The industry has paid more attention to creating better programming languages and finding talented programmers in the past, but with the development of software engineering technologies, more managers and practitioners realized that these are now the essential factors which are driving the industry. Millions of software failures show the industry that they are more important factors rather than just strong programming skills of a team. However, requirements engineering and system design is still given less attention in industry currently, especially for the fairly small or medium size software projects.

The reason can be given as following (based on my own experience).

1. There is a shortage for professional engineers in this area. Nowadays,
most of software engineering department in the universities and training organizations still focus on develop an individual's programming skills primarily. People can get limited experience for a piece of an assignment or small project from such educational organizations, but they do not acquire the capabilities to manage an industry standard software project. In fact, a significant number of software developments team in industry is organized to have talented programmers but without a good system design engineer. Referring to the lifecycle of the Project SS, the fact that software maintenance took a large percentage of work and made slow progress is caused by the missing “1”. In this case, “9 + 1 > 10” shows us 9 talented programmers plus only 1 system design engineer will definitely be able to manage the software project better than 10 talented programmers developing this project on the fly.

2. System design is not able to make visible profits currently. In fact, it is easy to organize some software engineers to start working on a project, and the product might be good enough to use. However, in a long term view, a good plan of system designing will ensure the software project runs smoothly along its whole life cycle. If the project manager could not realize this, then the project might start quickly enough but might also become unstable.

3. Good requirements refinement is not easy to achieve. Requirement engineering is termed engineering because the basic technique of collecting customer requirements through a couple of emails is too inaccurate. A Requirement analysis has to be processed by a series of standard methods, documented formally and possibly negotiated several times. Unfortunately, many software companies are still processing their requirements analysis in a naive way, so that the engineers have to spend plenty of time identifying, recording and tracking the changing user requirements.

6.1.2 Comparison between Industrialized standard and Academic standard

Referring to what I learnt from the M.Sc lectures, I want to represent a possible way to process the requirement analysis and system design for the project SS.

6.1.2.1 Requirement Engineering in Project SS

6.1.2.1.1 Current requirement analysis processing

Requirement collecting and analysis in the Project SS is processed in such a casual way that the collected requirements from users are assigned to the task list without any analysis. The way to collect information from users is also informal. During the requirement collecting process only limited documentation is involved to record what the user requires. In such a case, the final specifications which the development team can get are too ambiguous to
follow. This is abnormal from a software engineering point of view but it is a regular practice in most companies and in other projects of this particular company. An ambiguous, non-analyzed set of requirements will cause the following issues I found during the development.

1) **Ambiguous requirements lead to development failure and Software Quality reduction:** Since there is no high level refinement for the collected requirements, engineers can only develop the functionalities following their own tastes but not the users’. In fact, engineers may use a “technically” easier way to implement the feature when there are no rules to follow, resulting in the software quality being invisibly reduced. Also, the generated functionality may need to be updated again if it is not able to meet the real specifications after review by the users. Even worse, the finished functionality has to be re-built because of a small issue, and sometimes, plenty of internal changes are required for such a small update.

2) **Non-analyzed requirements lead to non-generic functions:** In the case of the project SS every new piece of the requirements arrives when a new customer purchases the product. Then the project manager will assign it to an engineer to create this functionality. Therefore, engineers are working on their functionalities individually and each generated components are isolated from the others. The consequences are that the software product is being modified and updated all the time and most of the functionalities are particularly designed for only one or several customers in a non-generic way. The product becomes more and more tedious to alter in order to meet every single user’s needs. Maintaining such a tedious list of functionalities costs most of the time and money in this project.

6.1.2.1.2 Requirement analysis proposal for Project SS

1. **The target of requirement analysis:**

The target of a software requirement analysis is to define a high level refinement of functionalities and performance for the final product before implementation. Additionally, it will create definitions for all constraints and interface details for the product.

Only the end users are the target source for processing the requirement analysis. On the one hand, although it is necessary to fully understand each piece of requirements from the end user, it is not necessary to accept every piece of requirements. (In some cases, strictly following a piece of requirement will not suit with the product because of the overview consideration, but an alternative solution might be acceptable for both sides). On the other hand, only an accurate specification is able to present users’ needs, which is the foundation of everything else. Any ambiguous specification will eventually cause a software failure because of the non-matching with the users’ needs.

2. **The process for requirement analysis:**
I will propose a requirement analysis for the Project SS in the following steps

1) Identify system requirements

First of all, after receiving the requirements from the end user, engineers and project managers should cooperate with each other to identify the system requirements, make an evaluation of the implementation duration, the technologies involved, the security level and any interface details needed to achieve current requirement. Also, the quality should be discussed and estimated.

Secondly, it is necessary to build up a communication link between the development team and the requirements source in case that the requirements have to be negotiated during the development.

![Communication tunnel for requirement analysis](image)

2) Refine specifications

The step will need the engineers to make high level refinements for the current requirement specifications on a global scope. They must identify the connection between the existing functionalities and the new ones to find out if any existing modules and components can be reused. They must make alternative solutions for those which are too non-generic or unsuitable for the whole project. They must then generate the final specifications for the user requirements with the user’s agreement.

3) Documentation

When documenting the specifications for software development team, the details should be provided from a development engineer’s point of view as
accurately as possible. If more details are provided to guide the development work, a higher quality of the software will be ensured in the final product.

6.1.2.2 System Design in Project SS

6.1.2.2.1 Low-Fidelity Prototype

First of all, I would like to talk about Low-Fidelity prototypes and the way to build up such prototypes I learned from the M.Sc. lecture – Human Computer Interface that creating prototypes should happen in every step of the system design procedure. Using prototypes allows designers, users and engineers to work efficiently together for the system design. Also, prototypes can allow the involved parties to identify the quality of the product, provide feedback and engage in conversation with the actual prototype material.

With respect of the module and the lecturer, I introduced what I learned about the Low-Fidelity prototype to the team and got a great response from my workmates and my manager.

Low-fidelity prototypes are useful because they tend to be simple, cheap, and quick to produce\(^\text{28}\). Referring to the materials from the HCI lecture, some prototypes for new features and functionalities in the Project SS were built up with PowerPoint.

(For example see: [http://www.boxesandarrows.com/view/interactive](http://www.boxesandarrows.com/view/interactive))

6.1.2.2.2 Overview of System design

While Requirement Engineering answers the question “what to do”, the next step will deal with “How to do”. System designing concentrates on developing prototypes for the project and individual components. Through these generated prototypes, the final product can be abstracted according to the user requirements. System design can be defined as a sequence of processes which are data structure design, system architecture design, interface design and integrated process design, which will now be discussed individually.

System design is the foundation to ensure a good software quality during development. Evaluating product quality becomes possible because system design provides a way to represent the product.\(^\text{29}\) Without system design, a project could not be built in a stable way. Referring to the Project SS, as discussed before its lack of system design brought up a number of issues during the development and later maintenance such as the messy script / code layout, limited extendibility and redundancy functionalities as illustrated in Figure 27.

\(^{28}\)Sharp Helen, Rogers Yvonne, Preece Jenny, Interaction design: Beyond human computer interaction (2\(^{\text{nd}}\))

\(^{29}\)Zong MingDi, Cai Ying, Liu XuDong, Li Xiangyuan, A Multi-Perspective Hierarchical Division Approach for a Product in Its Modular Design (In Chinese), Transactions of Beijing Institute of Technology, Oct,2003
6.1.2.2.3 System designing Proposal for Project SS

Since I was not involved in the System design phase for the Project SS, and based on the discussion above, the Project SS has some shortages in its system design and project architecture, so I will make my own system design proposal for the Project SS.

1. Refine the system environment

The first step will require the project manager and development engineers to build a rough prototype for the entire system based on the required functionalities. Through a study of the requirement, a list of possible skills/technologies which might be needed during further development should be discussed. Things such as the development language/script, database, and operating system should be decided. Based on the project SS, it would be better if any special skills e.g. AJAX, JQUERY or Google App, for a particular functionality can be located by experienced engineers before the development. A top-down refinement methodology can be used to create the system prototype, so that the requirements will be analyzed further and deeper until the functionality can be illustrated by programming language.

2. Coding framework

Referring to the M.Sc. lecture Web engineering & Internet solutions, it is highly recommended to build up a website using a suitable framework, such as Ruby on Rails, Zend or Codeigniter. The advantages of using a framework for the development are:

1) MVC: Using a framework can really help engineers manage their coding layout. The most popular coding frameworks are all using a Model-View-Controller architecture. As the coding layout becomes

---

30 Codeigniter is a php framework I used before for web development. More details can be found on http://codeigniter.com/
organized, it will be therefore be much easier to manage the project and generate prototypes. Therefore, if a MVC architecture is the foundation of a web based project, then using a suitable framework will make the foundation more stable. Also, a MVC enables layout and content to be processed separately which significantly improves the extendibility, readability and reusability.

2) Speed of development: The amount of code required to get a web based application running by using a framework is significantly lesser than developing it in plain programming. For example, only one line of code can enable making a query to data source when coding within a framework, $result=$module_name->$function_name since the functions have been pre-designed separately. This could be approached without a framework, but obviously, there are a significant number of functionalities that would have to be generated by the development team itself.

3) Cost: Most of the popular frameworks are open source and free to use.

3. Data structure design

Data structure is a representation of the logic relationships between each of the elements. A data structure design should define data organization, data accessing methods, and connections with each other. The topical data structures include scalar, vector (array, stack etc) and n-dimensional vector. In web development (especially programming in Perl), the most popular data structure is the Hash Table.

In web-based development, a Hash Table provides an easy access to data in a database table. The content of a database table can be fully described by a corresponding Hash table, especially, using an inline array and inline Hash table, A Hash Table describes a complex data table structure. Taking the following database table in Figure 28 as an example, a corresponding Hash table can be created as shown below.

<table>
<thead>
<tr>
<th>Username</th>
<th>Password</th>
<th>Email</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>John Doe</td>
<td>123456</td>
<td><a href="mailto:JohnDoe@a.com">JohnDoe@a.com</a></td>
<td>Admin</td>
</tr>
<tr>
<td>John Smith</td>
<td>123456</td>
<td><a href="mailto:JS@a.com">JS@a.com</a></td>
<td>Operator</td>
</tr>
<tr>
<td>Sherry</td>
<td>123456</td>
<td><a href="mailto:Sherry@a.com">Sherry@a.com</a></td>
<td>User</td>
</tr>
</tbody>
</table>

Fig 29 A sample database table

Creating a corresponding Hash Table in Perl script:

```
my %user_info;
$user_info{JohnDoe}[pwd] = 123456;
$user_info{JohnDoe}[email] = JohnDoe@a.com;
$user_info{Sherry}[status] = User;
```
# store Hash Table in an array
my @user_infor_array = (  
  { username => 'JohnDoe', pwd => '123456', email => 'JohnDoe@a.com', status => 'admin' },  
  { username => 'JohnSmith', pwd => '123456', email => 'JS@a.com', status => 'operator' },  
);

# Access data
$user_info_array[0][username];  # 'JohnDoe'
$user_info_array[1][email];  # 'JS@a.com'

Note: A Hash Table can be much more flexible when being defined and used in practice than the example above shows.

### 4. System Component Design

Once the requirement analysis has been processed, engineers will be able to make a plan about system components. The System component is the representation in detail of the system layout, which means, the system is a combination of a set of system components which are named and addressed individually. The System components will be assembled into the whole system to meet the user requirements.

In terms of a large size software system, the complexity of the whole system makes it difficult to understand, implement and maintain, which means, solving two combined issues is much more difficult than finding two solutions for each individually. Therefore, the whole system should be divided into a set of system components so that each component can be treated as a simpler problem in itself. The system then can be built up by the combination of each component.

Based on the discussion above, if each component is individual to each other, when the number of system components becomes smaller, it costs less to be implemented. However, when the number becomes bigger, the connections between each component will increase and it costs more to define and deploy these connections for the system. Therefore, the number of system components should be optimized during the component designing process to reduce the development costs as shown in Figure 29.
5. Interface Design

In the Project SS, every software engineer has to take in charge of interface design work for their generated functionalities. With respect to Kyle Sollenberger’s 10 User Interface Design Fundamentals, I will talk about how those principles are followed in the Project SS with my own experience.

*It’s no great mystery that truly great user interfaces are the ones that are engineered to stay out of the way. ‘Staying out of the way’ means not distracting your users. Rather, good UIs let your users complete goals.*

--- Kyle Sollenberger, 10 User Interface Design Fundamentals

1. Stay consistent / Pay attention to patterns

Since there are a number of existing functions and interfaces in the Project SS, the consistent deployment of new user interfaces will allow users to get better experience of how things will work, increasing their efficiency. During the requirements collection, it is better that some of the users would like to provide more details on the interface they want to see, such as saying ‘it will be very similar like that (an existing feature)’. Most of the time, however, there are not enough details included in the requirement specifications. Thus, the engineers have to find out a pattern from the existing interfaces. *By using familiar UI patterns, you will help your users feel at home.*

2. Provide feedback

A combination of CSS and AJAX has been used in every possible place in the Project SS to provide feedback and information for the user. CSS is used to
provide ‘user help’ information. When a user moves his mouse above the
question mark, the corresponding information will be displayed.

AJAX is used to create interactive web applications on the client-side. In the
Project SS, it is particularly useful for calculations, e.g. calculating the tax or
discount for a product.

3. Be forgiving

A small feature is used in Project SS to improve the tolerance of the user
interface when a user has to deal with varied inputs. The ‘old value’ will be kept
for the user when the interface page is refreshed but could not be submitted
successfully due to it being an invalid input. Also, information will be displayed
to ensure that the user knows how to prevent the error from occurring again.

6.2 Coding in Project SS

Referring to the M.Sc. lectures, I will talk about my own experience about
coding principles and how I generated my code when developing the Project
SS. In contrast to working on some small software projects/applications in the
college, programming in a team-work based, large size software project
requires not only a strong coding skills, but engineers also have to be aware of
the ‘team style’ when programming. As a member of the development team,
besides implementing the required functionality, the code has to exhibit a good
quality of reusability, readability, extendibility and maintainability. The following
discussion will present my own experience of how to fit into the development
team and improve my skills during my development.

6.2.1 Coding Style

Since the Project SS has been developed for years and many features
have already been deployed, most of the coding work at this time was to
maintain and upgrade the product. Thus, the engineers were spending more
time reading existing code rather than generating new code. Therefore, the
readability and reusability of the code becomes the most important factor to
consider when programming. To improve the quality of coding, the engineers
should follow some rules to approach a better ‘coding style’. In the article How
to improve your programming Style, the author introduced the following
aspects to approach a better ‘coding style’:

1. Document the source code. They should choose meaningful names for
each variable, function and package. In case of using abbreviated names,
it is better to follow the former engineers’ style to keep the naming - rules
consistent. Comments have to be documented clearly. Spaces and line feeds can be used to improve the readability of the source code.

2. Structure the source code. The readability of the source code should be considered by engineers firstly, avoiding a one-side pursuit of efficiency that results in complex code. The efficiency of the program should be approached by creating better algorithm and using existing components and functions instead of generating new ones.

3. Improve the source code. New code should be generated instead of fixing old code. Even an existing function can still be used, but if that piece of code is already out of date and does not fit the global style, then it is better to spend some time to generate a new one.

6.2.2 Efficiency

Code efficiency is represented by the time taken when executing the code, and the memory this piece of code occupies. In the Project SS, especially in the perl program, the efficiency of a database query always acts on the efficiency of the whole program. From my own experience, one common issue about using too many database queries in the code is that it significantly decreases the efficiency of the whole program. It is better to create one query to retrieve data from database using SQL commands such as 'JOIN or ANY' to build up a combined query statement.

6.3 Non-formal testing

As discussed before, a formal testing plan has been defined for the testing procedure. However, the Project SS still has a bottleneck in testing. On the one hand, managers in the company do not want to set up a formal testing team for the Project SS according to their own experience. This because the Project SS was already built step by step from a small project; product development went smoothly in the past and the previous engineers did not make any big mistakes. So, most of the software engineers who are involved in this project are also involved in the testing and bug fixing for their own developed functionalities. This makes the testing process for the Project SS to be insufficiently professional. Even most of the tests in the testing plan can be processed by automated testing software the functional testing requires more professional experience and knowledge because both white-box and black-box testing have to be done with precision. Unfortunately, due to reasons such as deadlines or the cost that are so important in the real industry, most of the functionalities have not been tested completely for the functional testing, although it has been defined in the testing plan.

Another problem of the testing process in the Project SS is more or less a particular case. Since the software product is performing its services on a web-based interface, the project manager would like to consider the mass of users as a tester group. The huge number of users can provide the most
accurate feedback for any new upgrades for the product. From the managers’ point of view, it can greatly reduce the cost of what should be spent on testing since most of the users are involved in testing the product voluntarily. Furthermore, because the user support desk, engineers can give a quick response if any software errors are found. In my opinion, this could be considered as a business strategy to reduce development costs. However, ‘making changes first, fixing errors after’ might give a bad impression to users of the product, and also puts the product in an unstable situation.

6.4 Documentation

The Project SS is being upgraded and customized all the time but the manager staff has not paid enough attention to updating the documentation when there are new features and functionalities deployed for the product. This means that the corresponding part of the document can easily be out of date.

Instead, for the Project SS itself, the project managers have paid more attention on product documentations in terms of their business usage and the development documentation is not well organized and maintained. In fact, from a software engineering point of view, development documentations should have the same priority as product documentations. A one-side pursuit of business factors has created bottlenecks in several aspects of the Project SS.

6.5 Agile software development

From the above discussion, a conclusion can be draw that it is still different when implementing a software project using the ideals of the academic software engineering modules or using the reality of the industry modules. There is no doubt that the ideals of the academic software engineering modules can offer a better way to manage and develop a software project. But software projects might not always follow the ideal way due to the reality in practice, such as deadline, budget and lack of knowledge etc. Therefore, it is interesting to seek for a better match between the ideal modules and the industry-standard modules.

Agile software development is a group of software development methodologies based on iterative and incremental development, where requirements and solutions evolve through collaboration between self-organization, cross-functional teams. Agile software development can be summarized as the following aspects:

1. Approach a minimum requirement analysis and system designing at the very first stage;
2. Develop software applications iteratively;
3. Reconstruct as early as possible;

31 More details about Agile software development can be found on http://en.wikipedia.org/wiki/Agile_software_development
From my point of view, Agile software development can offer a good balance point to combine the academic SE modules and the industry modules. Taking Project SS as an example, the traditional SE modules such as Waterfall module could not fit this project very well. Traditional SE module divide the software lifecycle into different phases such as requirement analysis, development and testing etc, and the lifecycle will be approached phase by phase. But in Project SS, it is difficult to divide such phases neatly. Most of the time, user requirements are collected along with the development, and requirement changing happens all the time. So it seems that the ideal SE modules has to be a bit more “adaptive” to fit the industry environment.

For Project SS itself, it is more or less following a rough Agile software development methodology to process its lifecycle. The word “rough” here means Project SS is following a adaptive pattern to process each phase in its lifecycle but it is just an unconscious behavior due to the industry environment, and there is no theoretical guide to direct this behavior. Therefore, to find a better match between the ideal SE modules and the industry modules, I will use Agile software development methodologies to make a proposal as following:

a) Iterative development. Software lifecycle is divided into several iterations. Each iteration contains all the phases such as requirement analysis, system design, development and testing, so each iteration can be considered as a small Waterfall module. Therefore, software development process in each iteration is able to take advantages from the ideal SE module, but the global lifecycle is more adaptive to fit the industry environment.

b) Delta delivery. Software product can be delivered after each iteration instead of delivering a final product. Each delta delivery will be deployed and interacted with the existing functionalities, so that the earlier delta version can be delivered, the earlier end users can provide feedbacks and fix their requirements if necessary.

c) Development drivers by both developer and user. Agile software development advocates involving both client side and software developers. It is also meet the ideals of the academic theory about requirements engineering.

d) Development drivers by testing. Test cases can be generated before development, and then code can be generated to approach a pass of test cases.

e) Optimum communication. In the article Agile Modeling: Effective Practices for eXtreme Programming and the unified Process, the author noted that

---

32 Please refer to section 2.1.1 – the project lifecycle.
33 Scott W. Ambler, Ron Jeffes, Agile Modeling: Effective Practices for eXtreme Programming and the Unified
low-fidelity prototype can be used to approach an efficient communication instead of using Case tools and official documentation. In fact, documents which are generated before or during the development will always be updated beyond recognition. To find a better solution, quick communication tools such as low-fidelity is suggested.

In conclusion, Agile software development is a way to implement a better match between the ideals of academic software engineering modules and the industry modules. Using such methodologies can offer a better management of the software project to meet the industry standard, meanwhile, software project can also take advantages from the involved academic software engineering modules.
7 Conclusions

In this thesis, I introduced my experience of developing, maintaining and testing an industry-standard software project in which I was involved for my work placement from July 2010. With the comparison between the industrial standard and the academic standard I learned from the M.Sc. lectures, the advantages and disadvantages of implementing a software project in each way has been discussed. Through the comparison, the reviewing of the M.Sc. lectures helped me acquire a further understanding of what I learned from the modules I took. It was also very helpful for my research to see the real industry standard software project and examine it from a theoretical software engineering point of view.

During my work placement, all the software applications developed by me has been deployed successfully in the business for the company's software product. The other assignments such as taking part in the software testing team, maintaining and updating the hardware also got a measure of recognition from the project manager and my colleagues.

During my work placement, my capabilities of programming in perl, CGI and Java, implementing websites, operating Linux OS, requirement engineering and system designing have been improved to a higher industry-standard level. Meanwhile, my skills of team working, communicating and problem solving were also developed greatly during this period.

In conclusion, the Project SS which I worked on during my work placement has been implemented using an industry approach to software engineering that deviated from the academic theory I learnt during the MSc modules. The industry standard does have its own advantages for a software project in such an industrial environment but there are also some shortcomings that can appear. In my opinion, seeking a better equalization point between the industry standard and the context standard is a valuable topic for further software engineering research.
References:

[1] Niklaus Wirth, Program development by stepwise refinement, Communications of the ACM; April 1971
[8] Changlin Wang, Yue Zhang, Requirement engineering and Project feasibility study (In chinese), online tutorial, September, 2008

Website: