**Hydra Data Protection Tool**

Project Documentation Submitted

To the Faculty of School of

Computer Science and Information Technology

Of

Asia Pacific College

In Partial Fulfillment of the Requirements for the subject

Applied Projects 2 or Software Development

By

Alberca, Reginald John Steven R.  
Miculob, Kent Michael P.   
Laureta, Maria Letty M.

**Table of Contents**

[**Project Proponents** 4](#_Toc468648962)

[Project Team 4](#_Toc468648963)

[Project Adviser 4](#_Toc468648964)

[Project Consultant 4](#_Toc468648965)

[Project Professor 4](#_Toc468648966)

[**Abstract** 5](#_Toc468648967)

[**I.** **List of Figures, List of Tables, List of Notations** 6](#_Toc468648968)

[**II.** **Introduction** 7](#_Toc468648969)

[A. Background of the Problem 7](#_Toc468648970)

[B. Statement of the Problem 8](#_Toc468648971)

[C. Project Context 8](#_Toc468648972)

[D. Purpose and Description 8](#_Toc468648973)

[E. Objectives 8](#_Toc468648974)

[a. General Objectives 8](#_Toc468648975)

[b. Specific Objectives 8](#_Toc468648976)

[F. Significance 9](#_Toc468648977)

[G. Scope and Limitations 9](#_Toc468648978)

[**III.** **Related Literature** 10](#_Toc468648979)

[Data Corruption 10](#_Toc468648980)

[Causes of data corruption 10](#_Toc468648981)

[Back Up 11](#_Toc468648982)

[File System 12](#_Toc468648983)

[Hash value 12](#_Toc468648984)

[An Analysis of Data Corruption in the Storage Stack 12](#_Toc468648985)

[**IV.** **Technical Background** 13](#_Toc468648986)

[**V.** **Design and Methodology** 14](#_Toc468648987)

[**VI.** **Results and Discussion** 14](#_Toc468648988)

[**VII.** **Conclusion and Recommendations** 14](#_Toc468648989)

[**VIII.** **Appendices** 15](#_Toc468648990)

[A. Diagrams 15](#_Toc468648991)

[a. Flowchart 15](#_Toc468648992)

[b. Event Table 18](#_Toc468648993)

[c. Context Diagram 19](#_Toc468648994)

[d. Data Flow Diagram 20](#_Toc468648995)

[e. Entity- Relationship Diagram 22](#_Toc468648996)

[f. Data Dictionary 22](#_Toc468648997)

[g. Class Diagram 23](#_Toc468648998)

[h. Object Diagram 24](#_Toc468648999)

[i. Use Case Diagram 25](#_Toc468649000)

[j. Use Case Full Description 26](#_Toc468649001)

[k. Activity Diagram 31](#_Toc468649002)

[l. Sequence Diagram 37](#_Toc468649003)

[m. Communication Diagram 37](#_Toc468649004)

[n. State Diagram 38](#_Toc468649005)

[o. Timing Diagram 39](#_Toc468649006)

[p. Package Diagram 40](#_Toc468649007)

[q. Component Diagram 41](#_Toc468649008)

[r. Composite Structure Diagram 42](#_Toc468649009)

[s. Deployment Diagram 43](#_Toc468649010)

[t. Work Breakdown Structure (WBS) 44](#_Toc468649011)

[u. Activity List 44](#_Toc468649012)

[v. Gantt Chart 45](#_Toc468649013)

# **Project Proponents**

## Project Team

|  |  |  |
| --- | --- | --- |
| **Name** | **Position** | **Email** |
| Kent Michael P. Miculob | Project Manager/Developer | [kpmiculob@student.apc.edu.ph](mailto:kpmiculob@student.apc.edu.ph) |
| Reginald John Steven R. Alberca | Project Researcher | rralberca@student.apc.edu.ph |
| Maria Letty M. Laureta | Project Editor/Developer | mmlaureta@student.apc.edu.ph |

## Project Adviser

|  |  |  |
| --- | --- | --- |
| **Name** | **Position** | **Email** |
| Mr. Justin David Pineda | Professor | justinp@apc.edu.ph |

## Project Consultant

|  |  |  |
| --- | --- | --- |
| **Name** | **Position** | **Email** |
| Mr. Jacob Catayoc | Professor | jacobc@apc.edu.ph |

## Project Professor

|  |  |  |
| --- | --- | --- |
| **Name** | **Position** | **Email** |
| Mr. Manuel Sebastian Sanchez | Professor | manuels@apc.edu.ph |

# **Abstract**

This project aims to create a forensic tool capable of preventing data corruption. The existence of data corruption began since the beginning of the automated technologies, and in the world manipulated by virtual information it is necessary for us to secure the information that we have. Information stored virtually are so fragile that in fact a single error or point of failure could cause your information to be corrupted. Data corruptions are something you’ll definitely hated if you’re one of those people who doesn’t do backup or does backup occasionally. The efforts that you’ve inserted and the time allotted for the file or document would gone to waste, and how devastating could the event if it we’re to happen to you? The researches of this paper would want to create a solution on this. And they figure out that the data corruption itself are inevitable since there are external factors that aren’t preventable. As the researchers continue the study, they’ve realized that data corruption prevention aren’t just limited on how could you avoid the event of corruption, but you could minimize or neglect the impact by considering the fact that it’s unavoidable and be always prepared of it.

# **List of Figures, List of Tables, List of Notations**

* Figure 1: Flowchart
* Figure 2: Event Table
* Figure 3: Context Diagram
* Figure 4: Data Flow Diagram
* Figure 5: Entity-Relationship Diagram
* Figure 6: Data Dictionary
* Figure 7: Class Diagram
* Figure 8: Object Diagram
* Figure 9: Use Case Diagram
* Figure 10: Use Case Full Description
* Figure 11: Activity Diagram
* Figure 12: Sequence Diagram
* Figure 13: Communication Diagram
* Figure 14: State Diagram
* Figure 15: Timing Diagram
* Figure 16: Package Diagram
* Figure 17: Component Diagram
* Figure 18: Composite Structure Diagram
* Figure 19: Deployment Diagram
* Work Breakdown Schedule
* Activity List
* Gantt Chart

# **Introduction**

## Background of the Problem

Nowadays, technology is inevitable. From personal matter all the way to school works to professional works, researchers are inputting information to computers and other gadgets, thus, safeguarding is vital and crucial. They need to protect it not just from the hackers, who wanted to steal information for their own benefit, but also from being corrupted that will steal all the hard works and left you nothing but despair. These scenarios wherein the data itself becomes corrupted are due to hardware and software failures. Computers don’t have the ability to prevent nor retrieve because to them, it appears as the same data.

There are applications such as File Repair which identifies the corrupted data in your hard drive and creates a new usable file where it extracts much of the recovered data as possible. There is also an application called BadCopy pro where it recovers deleted files, formatted drive, or data loss due to damage, media errors, bad sectors and other reasons for flash drives and save the recovered files into a directory you specify. Another is application that offers safeguard is Microsoft Document, the approach that the Microsoft document does is deferent from file repair & other recovery applications such as BadCopy pro, since the Microsoft Document provides backup for your document in case it might be corrupted if the computer suddenly turns off, and it is a good practice since prevention is greater than cure. (BadCopy Pro - Flash Drive Data Recovery Software, n.d.) , (File Repair, n.d.)

As we have read the information, we can conclude that there’re a lot of application which fix and recover corrupted data, there are also application where they provide prevention mechanism to avoid having corrupted data. But the researches still think that all of this application is not enough.

## Statement of the Problem

Corrupted data had always been a problem in the computing industry, since most of the information are now stored in a mechanical device. A lot of factors can result to data corruption, such as software and hardware failures. There are identified tools or applications that can be used for in case of having corrupted data such as specified above. The researchers would like to know, how could they enhance the existing applications and how could they provide a better way in safeguarding your file from corruption.

## Project Context

Data corruption has always been a problem in computing. With the imminent implementation of internet of things, solutions on data corruption must be address to avoid future complications regarding data loss.

## Purpose and Description

People are expose to different security issues that revolves around the Security, Integrity and accessibility. The main essence of this research is to address availability issues that concerns data corruption. And to produce a tool capable of preventing the occurrence of corrupted data that can be easily used by non-IT people.

## Objectives

### General Objectives

* This Project aims to create a security algorithm that could save critical information from data loss.

### Specific Objectives

* To be able to provide an algorithm that could save someone’s information.

To be able to give confidentiality control on the user’s information.

* To be able to innovate a low-learning curve security algorithm that is applicable to most people.

## Significance

A lot of applications that were made related to this project are mostly for corrupted data fixation and retrieval. The essence of this research is to provide a way on how to prevent having corrupted data by means of having an application that is capable of detecting and fixing corrupted data, preventing having corrupted data by providing a way to backup, and minimizing the existence of corrupted information in real-time. The issue of corrupted data already exists along with the advancements of mechanical technology. As said by Microsoft “No computer is failure proof, but preventive measurements and strategic planning can make a computer more failure resistant”. And this innovation offers more than that.

## Scope and Limitations

This Project aims to help students and/or office-workers to secure their information or data on availability issues, whom are using windows OS- 7 and higher. The project primarily focuses on data corruption and prevention. It also assumes that the interaction is just between the players of the system such as the admin, system, and user. The prevention of data corruption is only focus on text documents, and it is beyond the systems capability if the flash drive itself became corrupted since it should be the users’ responsibility to protect the hardware itself.

The flash drive that will be used should have a NTFS file structure with a memory space not below 4 gigabytes. The optimal performance of the application could be achieved if the memory space of the flash drive is less than 90%.

# **Related Literature**

## Data Corruption

The data corruption is commonly problem of people that uses computer in such of programmers, office worker and even students. The data corruption can identify when files and folders are relocated or missing, open file error such a file invalid error in which you cannot open a file, file name renamed by a gibberish file name, file permission and attribute are modified, computer frequently crashing without any reason, slow disk operation and disk activities appear to be very busy even there is not much going on in the computer. (Data Corruption Info, 2009)

## Causes of data corruption

The data corruptions have causes which these causes have bad effect in the computer. These are the causes of the data corruption: First, the data is corrupted due to of bad programming, this happen when your computer is shutdown or restart improperly like intentionally or unintentionally. For example, switching off the computer in a hurry or sometimes the computer occurs to the power shortage it may cause the system to freeze in which the user going to restart the computer manually and because of this event, the file in the computer can be lose or be corrupt due incomplete bit. Second, the Malware, this may have different approach to the corruption of the file, this may occur a loss bot on the data or can’t be open or it might be lost in the computer, when this thing happens to the computer it might became prone to viruses.

These kind of malware can enter to the computer in which cause the many hardware failures. Lastly, the poor software and hardware, these issue can also affect to data because even into these modern age technologies it is still rapidly changing so fast that some file that we try to save into our computer cannot compatible in which also lead to hardware and software failure. (Rita, 2004)

## Back Up

Many people of this modern technology may prone to lost their data and for some, it is hard to make their work do it all over again. Maybe it can change its work to better or worse and one thing it can solve its problem when it is the file is backed-up, now other people may wonder what is a back-up. Back-up is an operation or procedure of copies data to an alternative location, so it can be recovered if deleted or it becomes corrupted.

There are many ways on how to back up files, there are two types on how to back up file. First, by using an external drive it can be used to store another copy of the file. Second, it can be back-up files using the internet it has many potentials such as cloud storage device that can store the file in Google Drive, Dropbox and Microsoft OneDrive these services can help to back up the file just in case the file is deleted or even worse got corrupted. These kinds of back up may help to recover the file. (Stepanovska, 2015)

Checksum

The checksum is used for ensure of the integrity of the file on which it can transmitted into one device to another. This can happen through the internet or simply through two different computers that are connected into a single network simply to ensure the transmitted file to be exact as the source file. Checksum is the best option to use also not only to secure the transmission of the file between two or more devices but also it is responsible for calculating on hash function on which normally posted when downloading.

This is also responsible to verify and calculate the downloaded file for making sure that there the same. The most common checksum that can see in your computer are MD5 and SHA-1, there are both have a weak point, this means that any malicious tampering can be lead to two different files having same computed hash. Due to these issue the newer SHA-1 is the best cryptographic hash function so far due to no issue that has been attacked. (Kishore, 2015)

## File System

File system is a method of an operating system which it uses to track the file inside the disk of the computer also the way on how the disk organize the file in it. It is created to initialize or format the hard disk. There are different types of file system and the commonly used are FAT, FAT32, NTFS. The newest one is the NTFS or New Technology File System it supports by Microsoft Windows Server 2003, this allows the user to gain the maximum benefits of need to the today’s enterprise business environment. By using this file system, the larger the hard drive, the larger the default cluster size. Since, would be increase the performance at the expense of the amount of space efficiency. (NT file system; sometimes New Technology File System, 2008)

## Hash value

A hash value is a numeric value of a fixed length that uniquely identifies data. Hash value represents a bigger amount of data as much a smaller numeric values, so they are used with a digital signature. It can put a hash value more efficient than signing a large amount of its value, also the this is more useful for the verifying the integrity of the data that send through the insecure channels. The hash value of received data can be compared to the hash value of the sender to determined that what data is altered. (Ensuring Data Integrity with Hash Codes)

## An Analysis of Data Corruption in the Storage Stack

The difficult part of creating and designing a storage system is that is providing the reliability and the availability that the user expects and it is use to be persistent forever and perpetually available but the machine has not of flaws, there will always be a problem that appears. There are lot of number of problem that, if not dealt with can cause data loss in storage system, unfortunately one of the primarily cause of the data loss in the storage system is disk drive unreliability it is well known that hard drive are mechanical, moving device that can suffer a mechanical problem leading to a drive failure and data loss. (Bairavasundaram, 2008)

# **Technical Background**

Nowadays, due to the fast evolving modern technology, things are easier to access and control. As the people continue to make life easy, convenient, and comfortable different Security issues occur in the global community, that can risk someone’s life or even threaten the national security. When we talk about security everyone is involve and everyone are at risks of losing their future. Improving security from different aspects could help us reduce our exposure on threats that might end up taking all our assets or even our lives.

The most powerful tool in today’s society are computers, they are multi-purpose gadgets that are capable of manipulating binary’s, which are used to communicated all types of gadgets. But even though it’s powerful it is still vulnerable to data loss or corruption.

The researchers conducted a survey on how many college students have experience data corruption, and the results shows that 88% of the students have experience data loss due to factors such as hardware failures, software failures, malwares, and Unintentional deletion of file. 29% of the corrupted files was caused by a malware, 27% are due to software failures and 44% was caused by hardware failures.

# **Design and Methodology**

# **Results and Discussion**

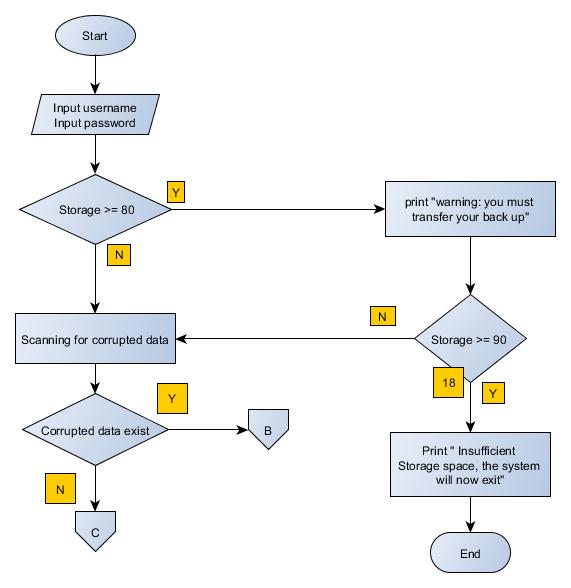
# **Conclusion and Recommendations**

# **Appendices**

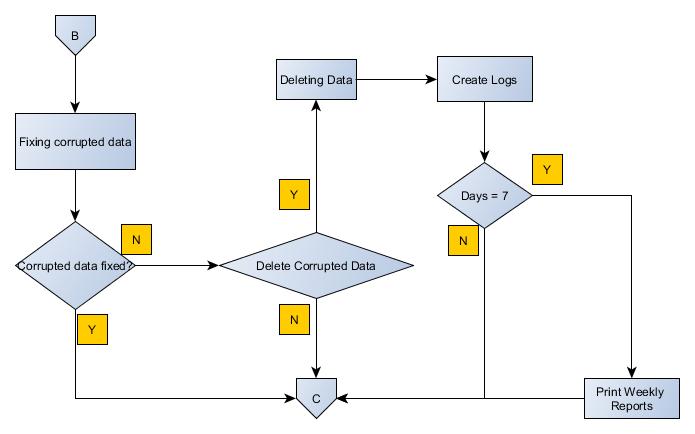
## Diagrams

### Flowchart

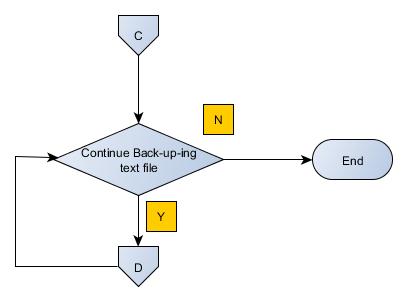
Main Program of Hydra Data Protection Tool



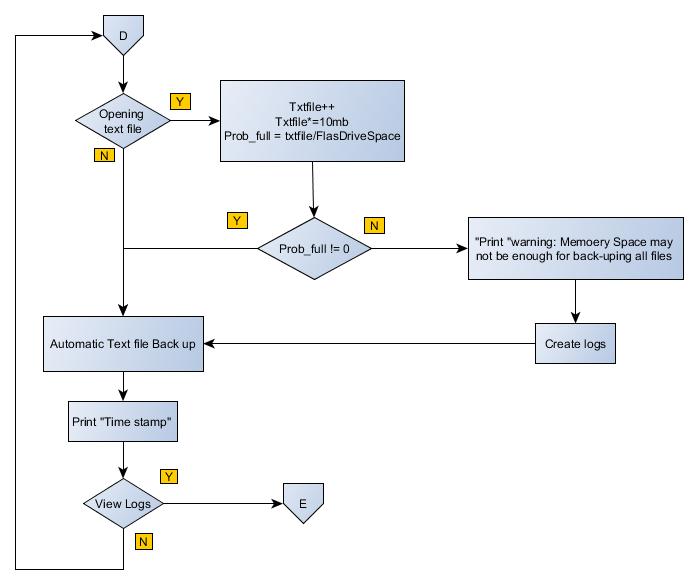
Fixation of Corrupted Text file



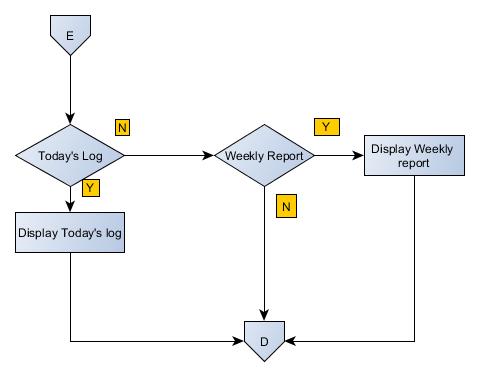
Continuation of Back-up



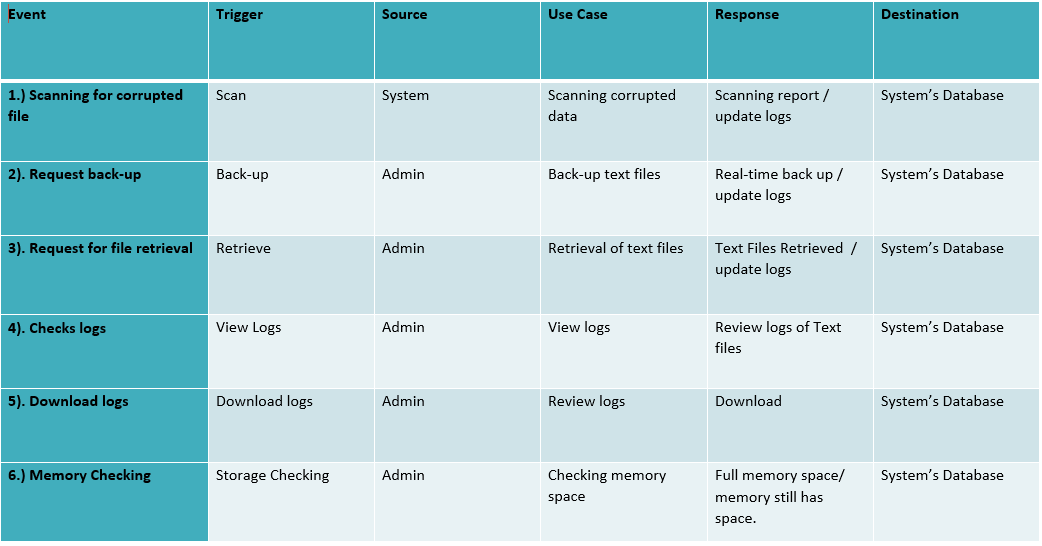
Flash Drive Space



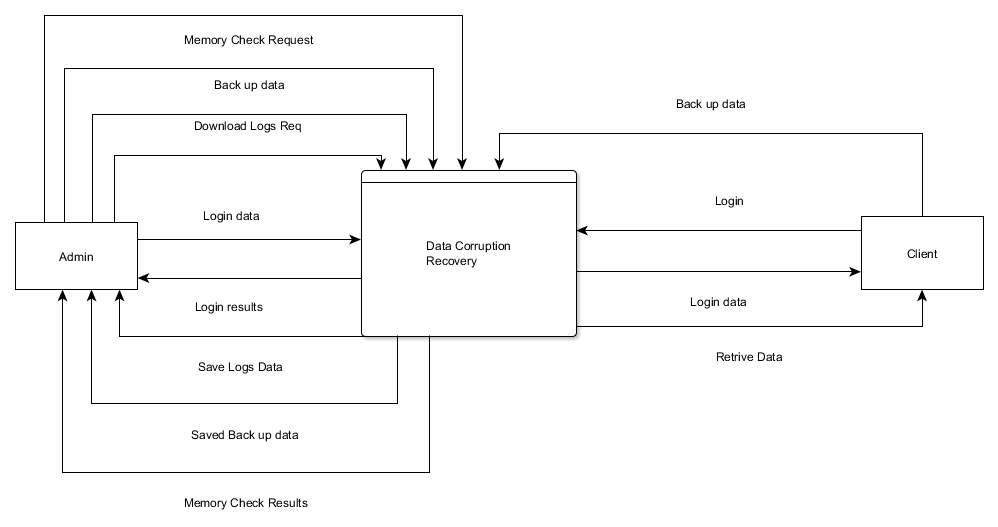
View Logs



### Event Table

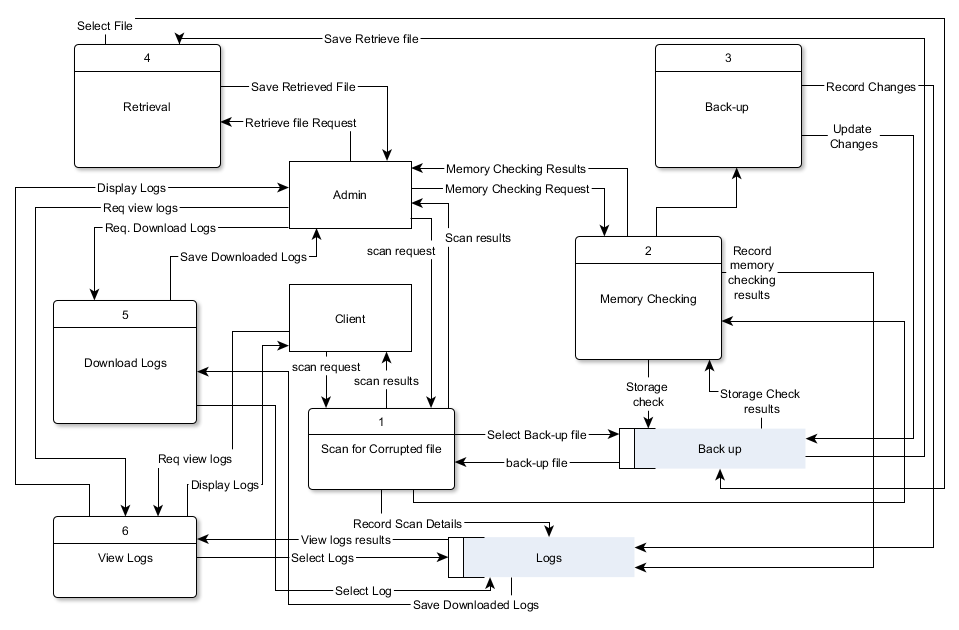


### Context Diagram

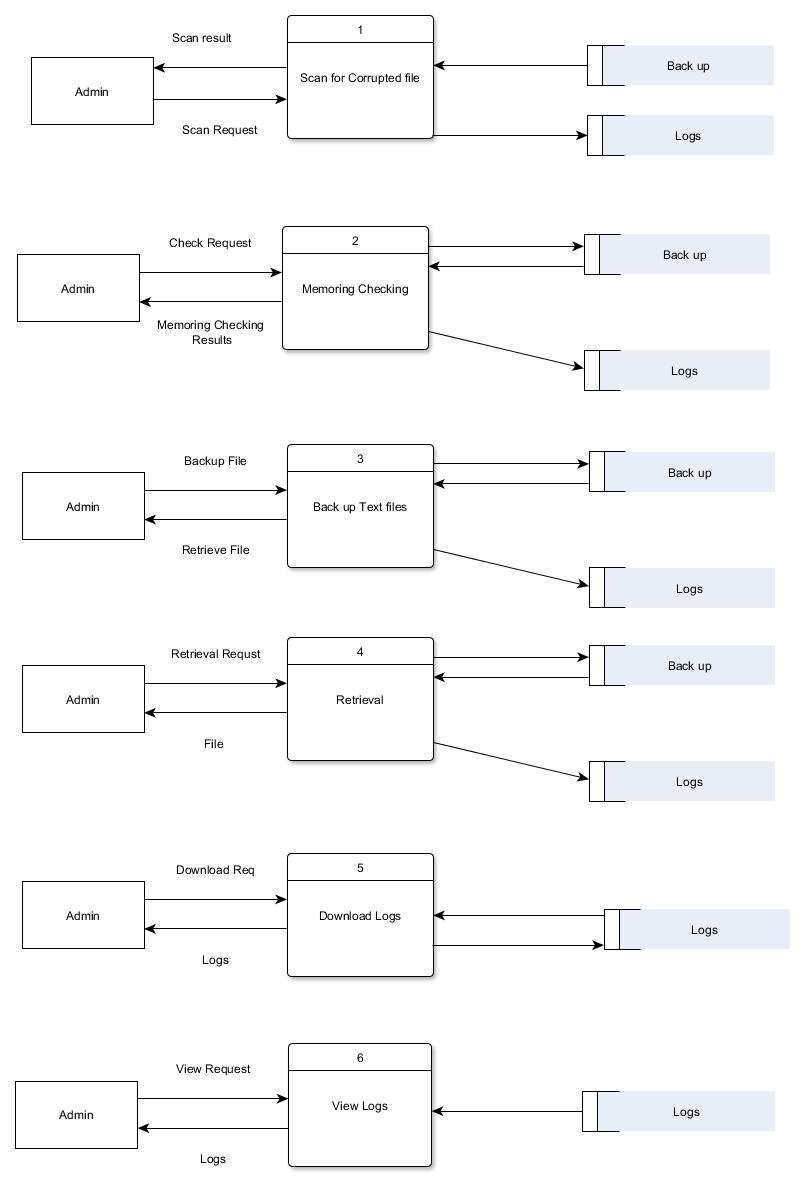


### Data Flow Diagram

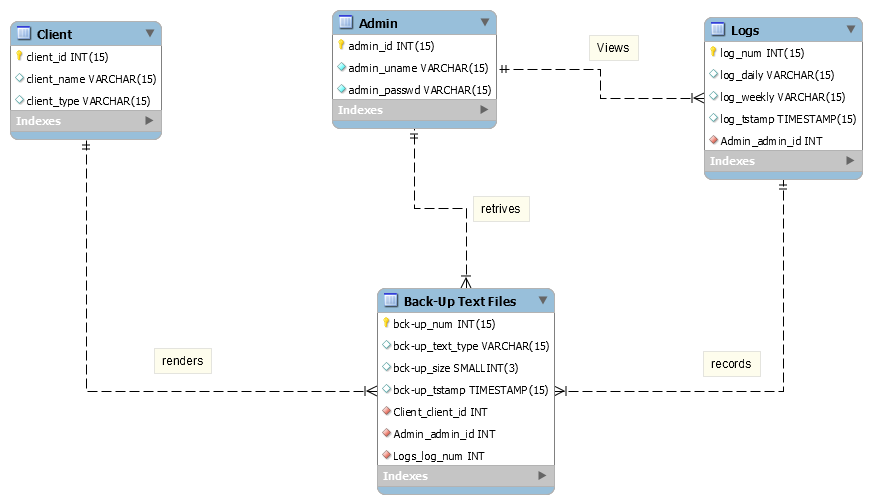
#### Level 0



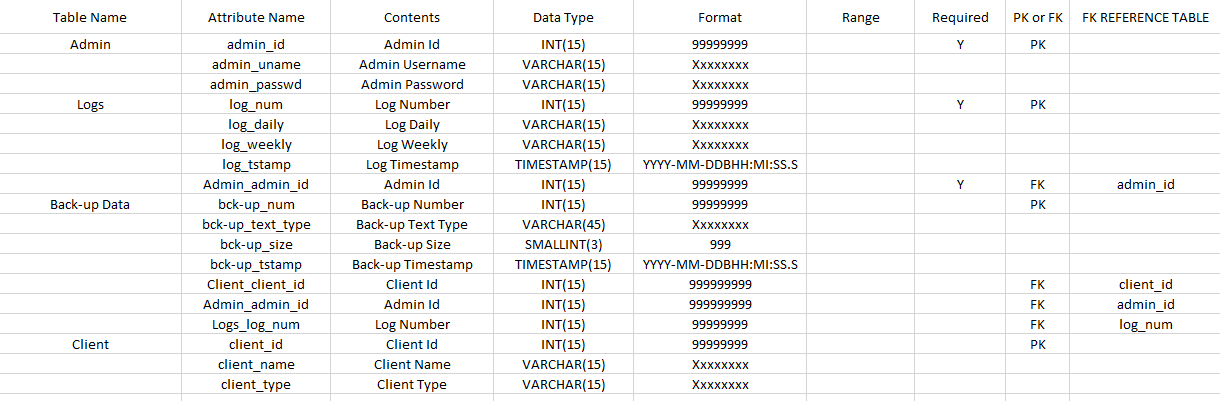
#### DFD Fragments



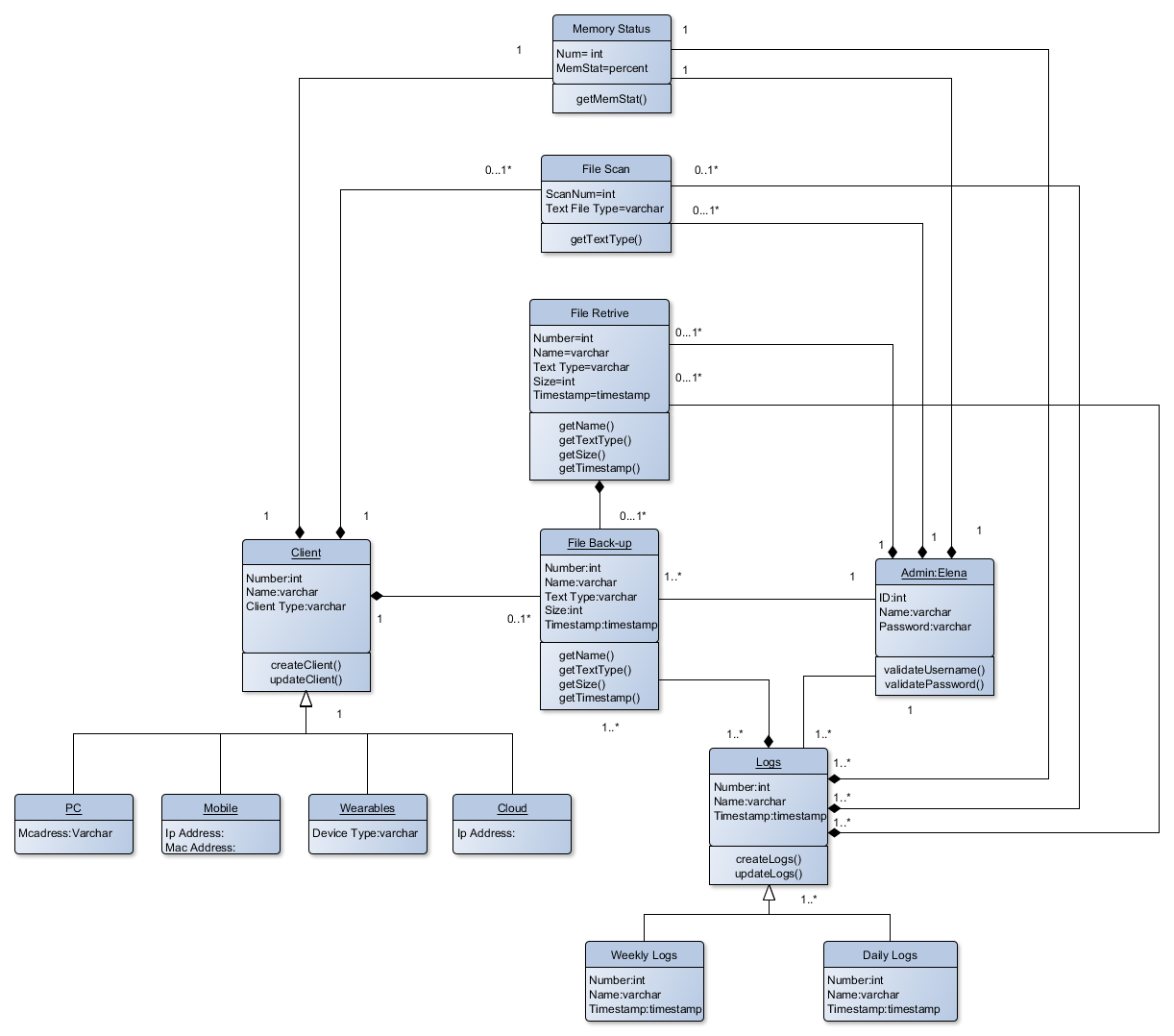
### Entity- Relationship Diagram



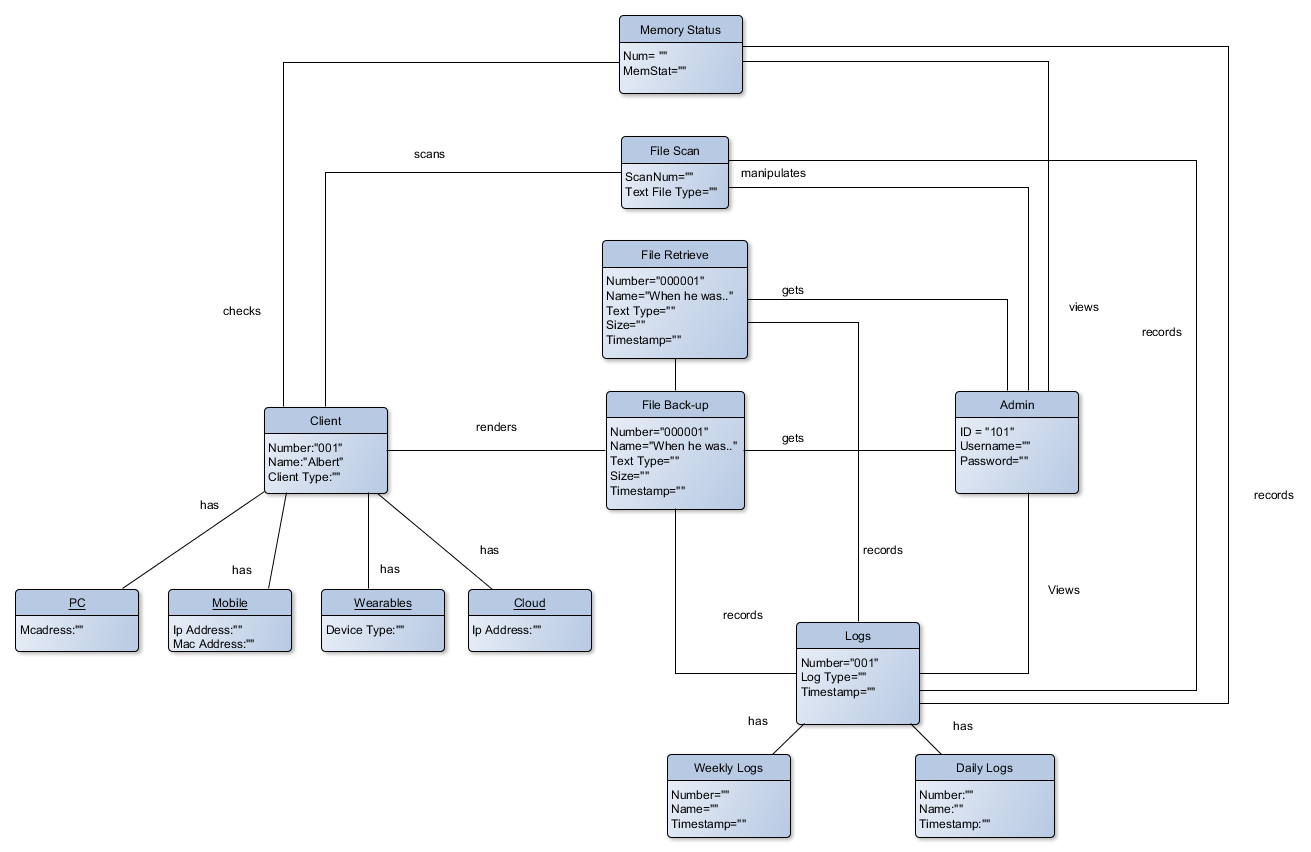
### Data Dictionary



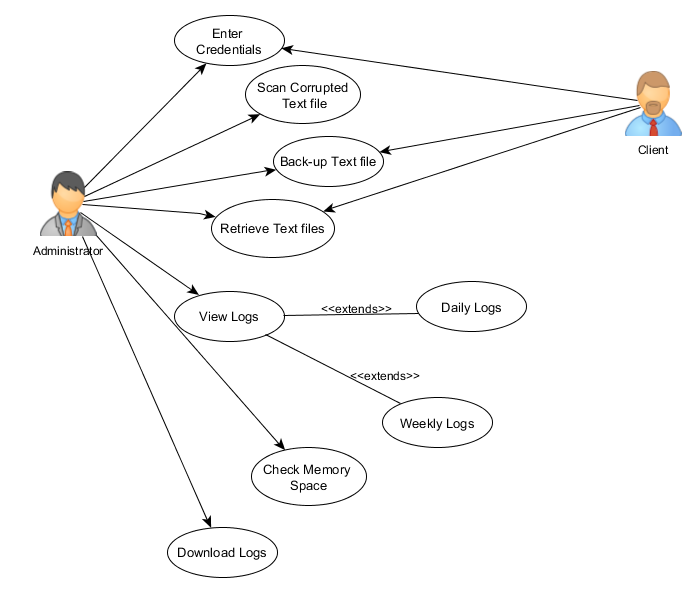
### Class Diagram



### Object Diagram



### Use Case Diagram



### Use Case Full Description

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC01 | |
| Use Case Name | User Authentication | |
| Scenario | User wants to Log in to the system | |
| Triggering Event | Log In | |
| Brief Description | When the admin access his files authentication is required. | |
| Actor(s) | Administrator | |
| Related Use Cases | ---- | |
| Stakeholders | Administrator, Client | |
| Precondition | User must plugin the Flash drive to the PC | |
| Post Condition | Administrator must able to access the system | |
| Basic Flow: | Actor’s Action | System’s Response |
|  | 1: Opens the application | 1.1: System displays the login page of the application |
| 2: User types its username/password |  |
| 3: User clicks login button | 3.1: System checks the database if the username and password is correct |
|  | 3.2: username and password matched System will message (SM01) |
|  | 3.3: System appears the main page of the application |
|  | 3.4: System scans for corrupted text file |
|  |  |  |
| Alternative Flow: |  | 3.2: If username and password of Administrator didn’t match |
|  |  | 3.3: System displays message (SM02) |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC02 | |
| Use Case Name | User back-ups text files | |
| Scenario | User request for back-up | |
| Triggering Event | Back-up Text files | |
| Brief Description | As a prevention mechanism on data corruption, an automatic real-time backup is running on background. | |
| Actor(s): | Administrator | |
| Related Use Cases | User Authentication | |
| Stakeholders | Administrator, Client | |
| Precondition | User opens an text application | |
| Post condition | User was able to back-up the text file automatically while working on text application | |
| Basic Flow: | Actor Action | System Response |
|  | Step 1: User Plug’s in the Flash Drive | 1.1: System checks if the flash drive is less than 80% of the storage |
|  | 1.2: System scans for corrupted text file |
|  | 1.3: System will identify If the corrupted data can be fixed or not. |
|  | 1.4 Fix corrupted data |
| 2: User opens text file application | 2.1: System will start executing automatic real-time back up |
|  | 2.2: System will calculate if the storage can accumulate the back-up file. |
|  | 2.3: System backed-up text file |
|  |  |
|  |  |  |
| Alternative Flow: |  | 1.3: System will identify If the corrupted data can be fixed or not. |
|  | 2. User Agrees to delete corrupted data | * 1. Delete Corrupted data. |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC03 | |
| Use Case Name | User views log | |
| Scenario | Admin views logs | |
| Triggering Event | View Logs | |
| Description | The changes made in the Flash drive are recorded in the logs. | |
| Actor(s) | Administrator | |
| Related Use Case | Use Authentication | |
| Stakeholders | Administrator | |
| Precondition | User is logged in as Admin | |
| Post Condition | User is able to view the in/out of text file | |
| Basic Flow: | Actor Action | System Response |
|  | 1: User view logs. | 1.1: System displays “Daily logs” and “Weekly logs” |
| 2. User select “Daily logs” | 2.1: System displays “Daily logs” page |
| 3. User select “Weekly logs” | 3.1: System displays “Weekly logs” page |
|  |  |  |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC04 | |
| Use Case Name | User retrieve text file | |
| Scenario | User wants to get the back-up text file | |
| Triggering Event | Retrieve Text File | |
| Brief Description | Admin retrieves files that were backed up. | |
| Actor(s) | Administrator | |
| Related Use Case | User Authentication, User Back-Up | |
| Stakeholders | Administrator, Client | |
| Precondition | User has a back-up text files | |
| Post condition | Admin was able retrieve the back-up files | |
| Basic Flow: | Actor Action | System Response |
|  | 1. Admin will Select & Copy the file to retrieve |  |
| 1. Admin will Paste the retrieved file. |  |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC05 | |
| Use Case Name | Scan Corrupted Text file | |
| Scenario | User wants to know if there’s any corrupted text file | |
| Triggering Event | Scan | |
| Brief Description | System scans for the corrupted text file | |
| Actor(s) | Administrator | |
| Related Use Case | User Authentication | |
| Stakeholders | Administrator, Client | |
| Precondition | Admin is logged in | |
| Post condition | Admin was able to log out the system | |
| Basic Flow: | Actor Action | System Response |
|  | 1. Log In’s the application |  |
|  | 1. Automatically scans for the corrupted text file |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC06 | |
| Use Case Name | Download Log | |
| Scenario | User wants to review logs | |
| Triggering Event | Logs | |
| Brief Description | To review the systems activity | |
| Actor(s) | Administrator | |
| Related Use Case | User Authentication, View logs | |
| Stakeholders | Administrator, Client | |
| Precondition | User must have logged in as Administrator | |
| Post condition | Admin was able to download the logs | |
| Basic Flow: | Actor Action | System Response |
|  | 1. Admin clicks “View Logs” | 1.1 Displays the View Logs page |
| 1. Admin clicks “Download Logs” | * 1. System download logs |

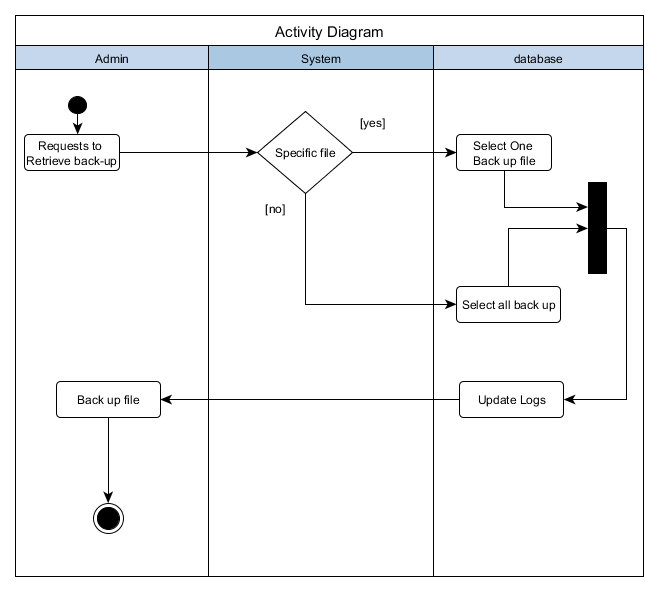
|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC07 | |
| Use Case Name | Memory Check | |
| Scenario | User wants to know the memory space of the flash drive | |
| Triggering Event | Memory Checking | |
| Brief Description | Able to identify the available memory space of the flash drive | |
| Actor(s) | Administrator, client | |
| Related Use Case | User Authentication | |
| Stakeholders | Administrator, Client | |
| Precondition | User must have logged in | |
| Post condition | User was able to know the memory space of the flash drive | |
| Basic Flow: | Actor Action | System Response |
|  | 1. Enter credentials | 1.1 Automatically Scans for the corrupted text file |
|  | 1. Scanned Complete |
|  |  | 1. Checks the memory space |

|  |  |  |
| --- | --- | --- |
| Number | PICDS-UC08 | |
| Use Case Name | Log out | |
| Scenario | User wants to log out from the system | |
| Triggering Event | Log out | |
| Brief Description | Log out from the system | |
| Actor(s) | Administrator | |
| Related Use Case | User Authentication | |
| Stakeholders | Administrator, Client | |
| Precondition | Admin is logged in | |
| Post condition | Admin was able to log out the system | |
| Basic Flow: | Actor Action | System Response |
|  | 1. Clicks the Logout button |  |
|  | 1. Destroys the session and logs out from the system 2. Displays the Log In page |

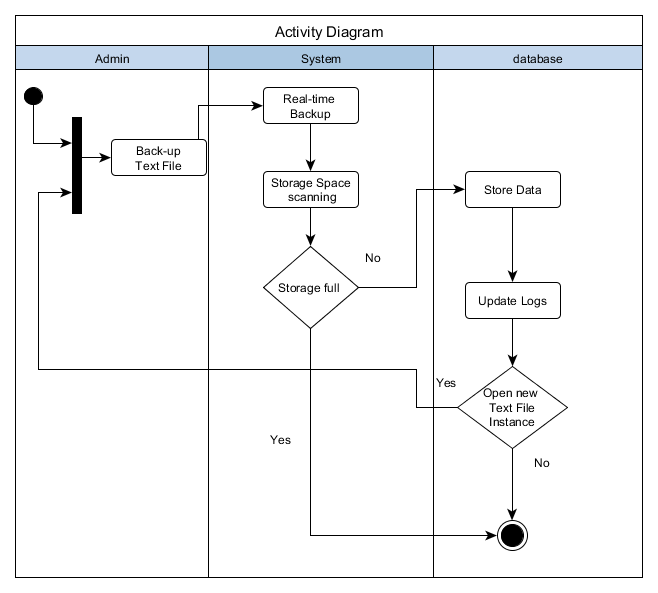
|  |  |
| --- | --- |
| System Message | |
| SM01 | You logged in successfully! |
| SM02 | Username or password is incorrect! |

### Activity Diagram

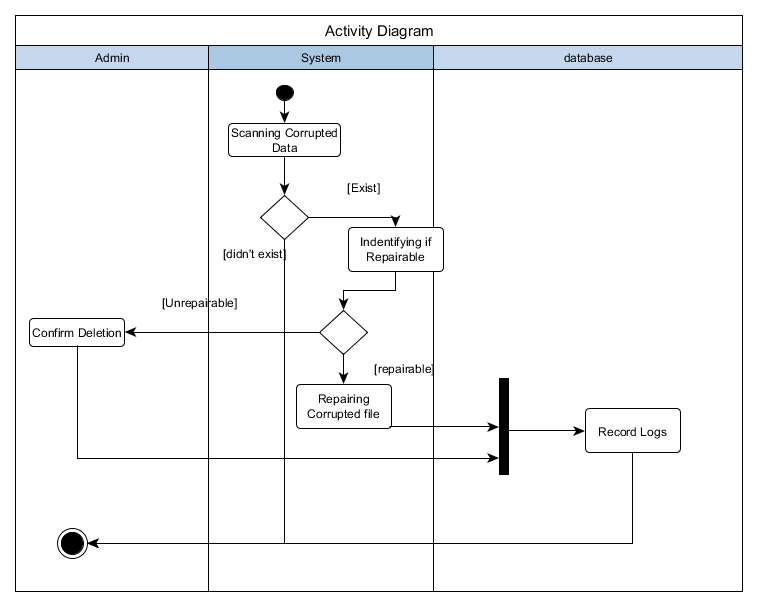
Back- up



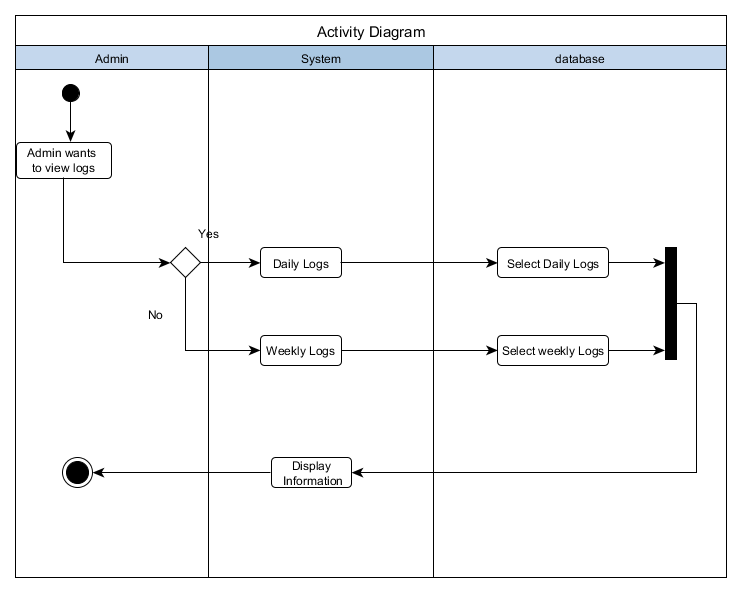
Memory Check and Scanning



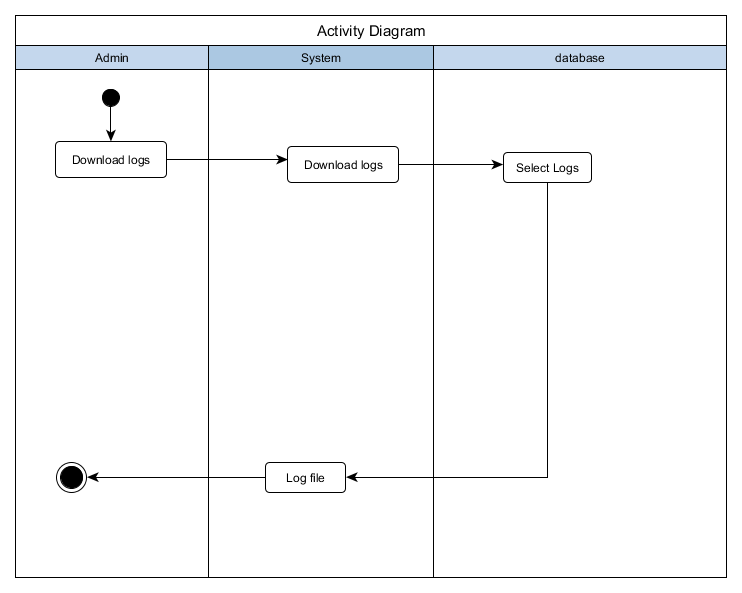
File Scanning



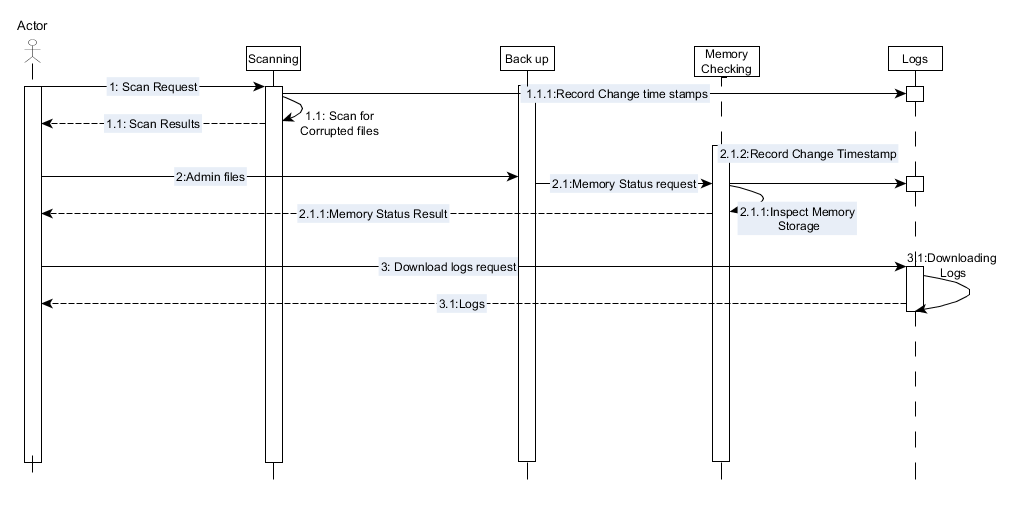
View Logs



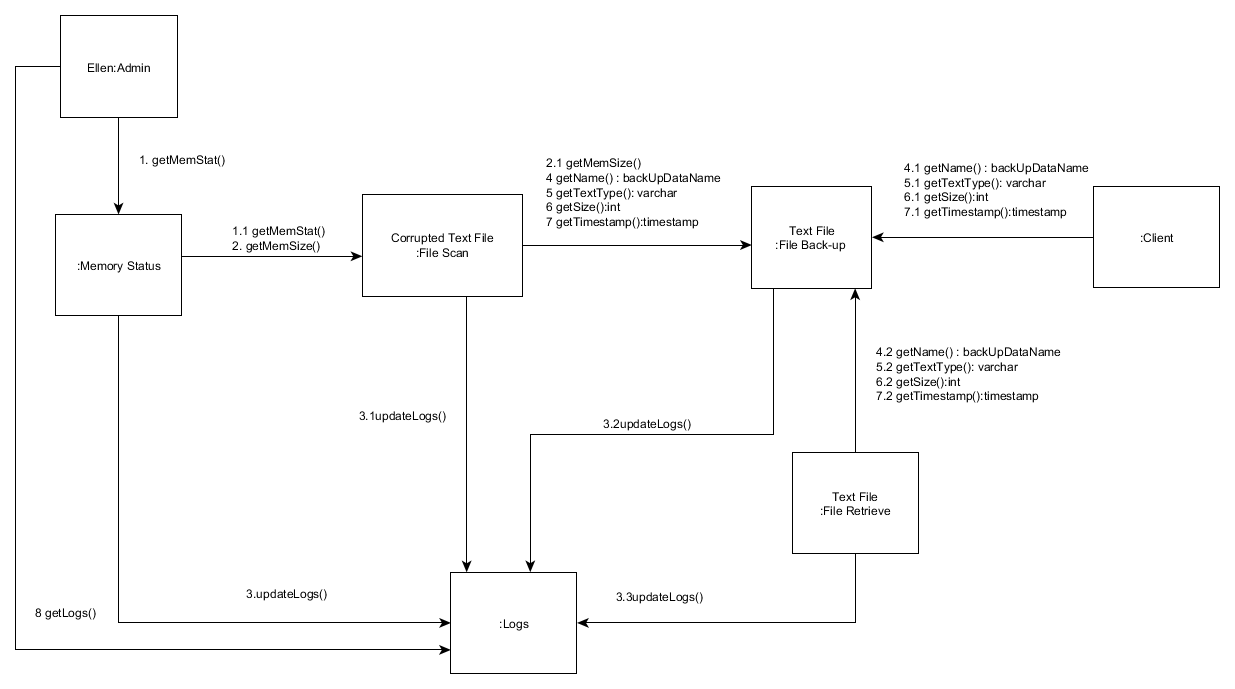
Download Logs



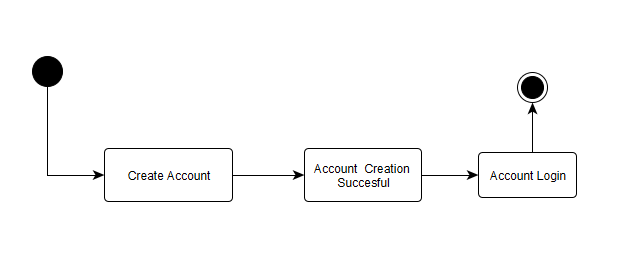
### Sequence Diagram

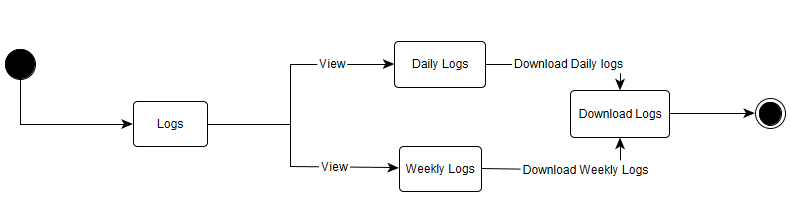


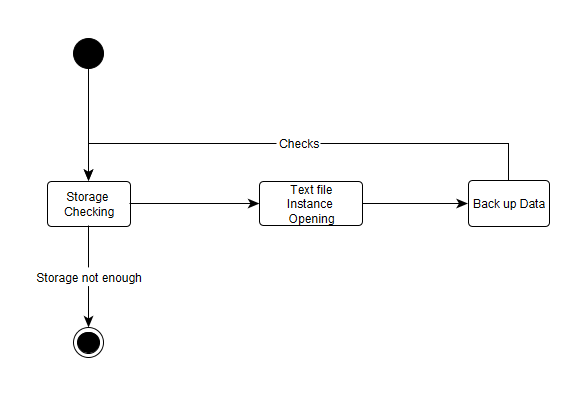
### Communication Diagram

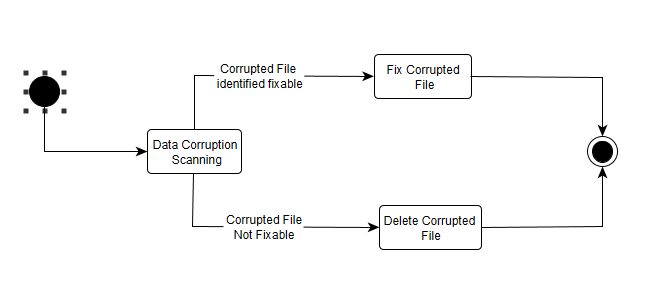


### State Diagram

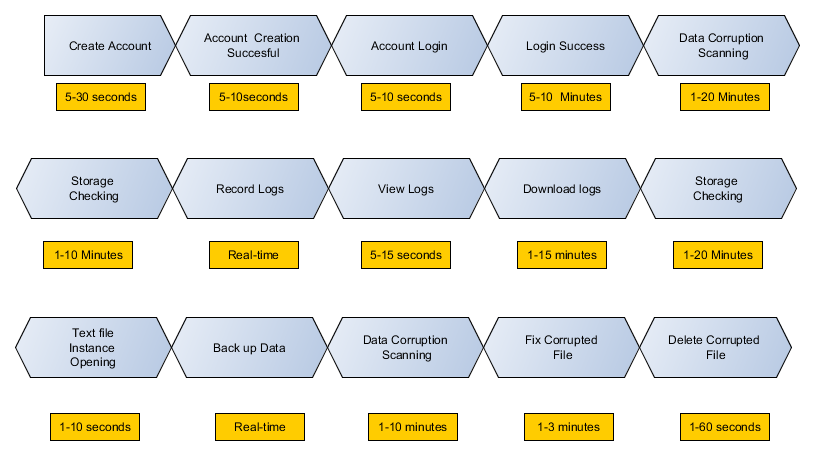




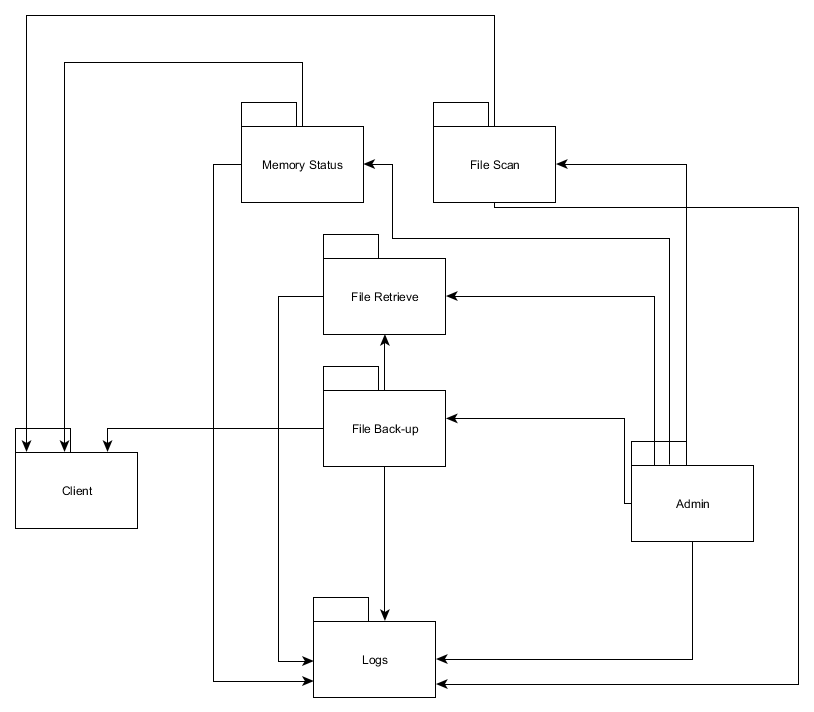




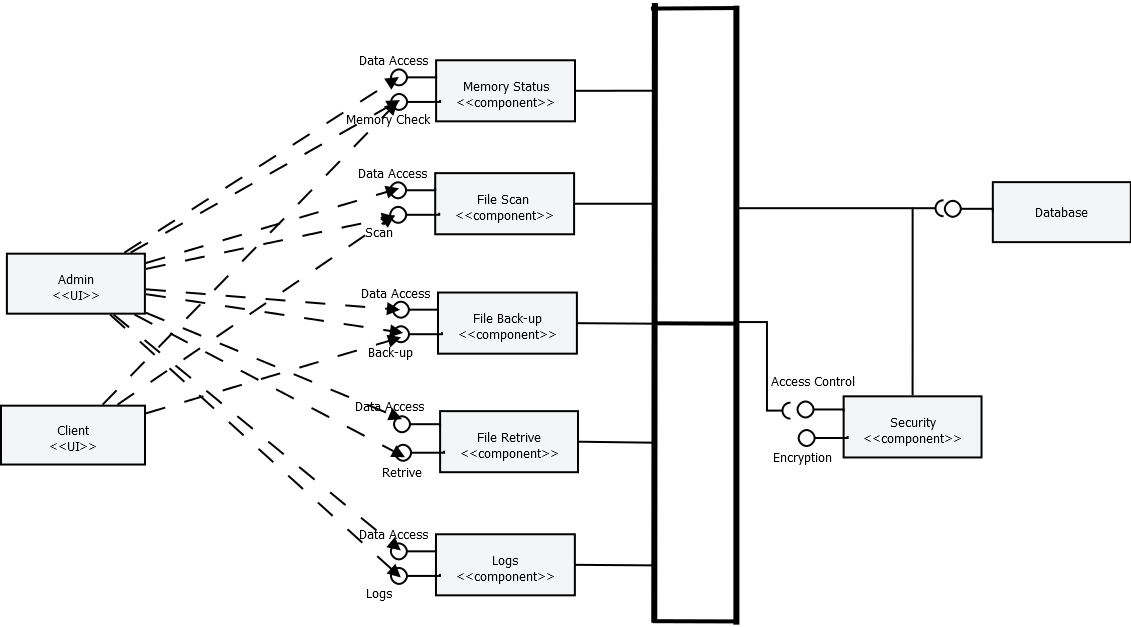
### Timing Diagram



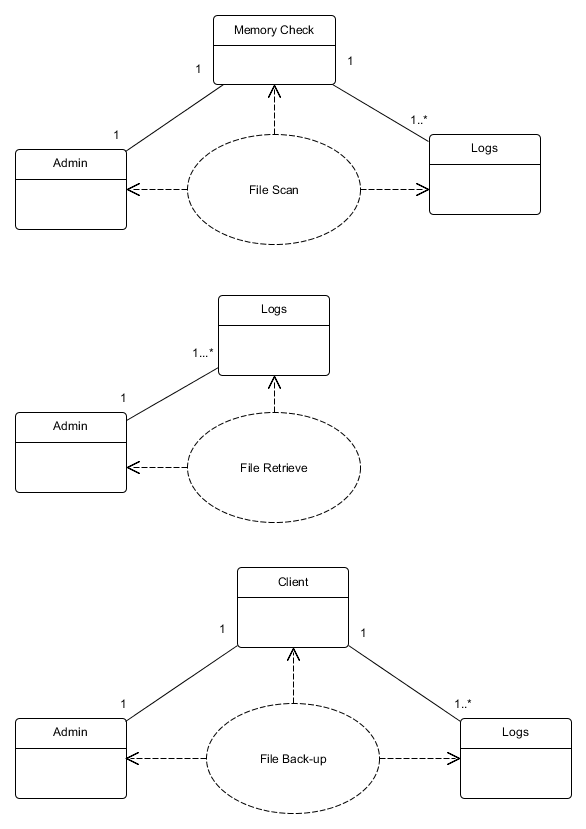
### Package Diagram



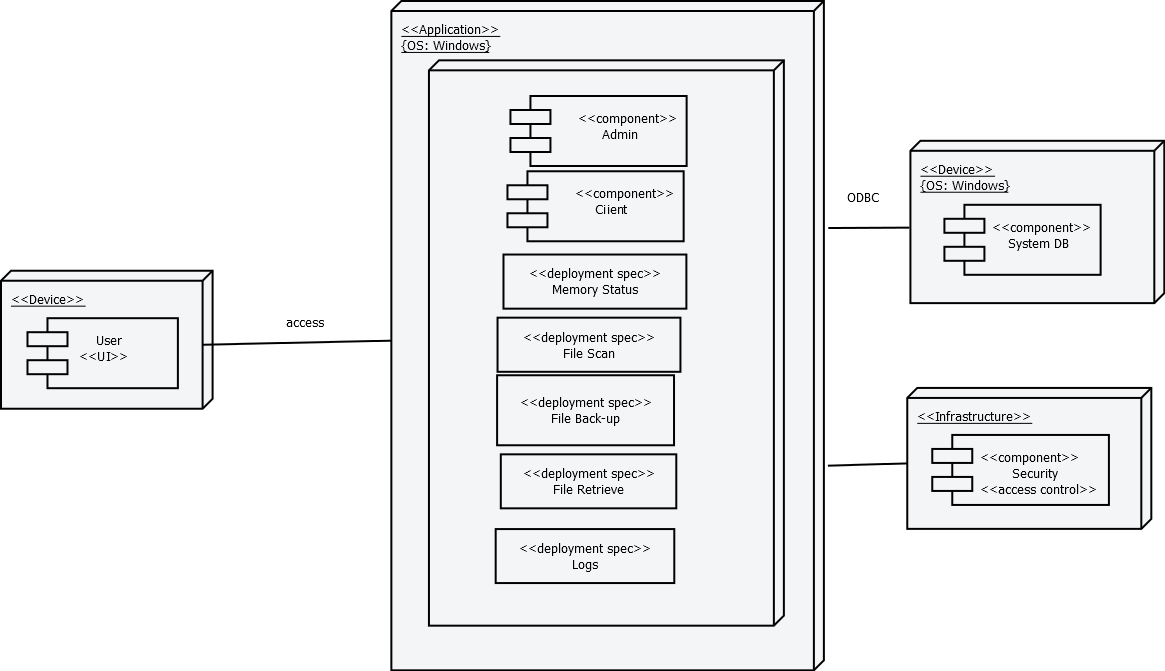
### Component Diagram



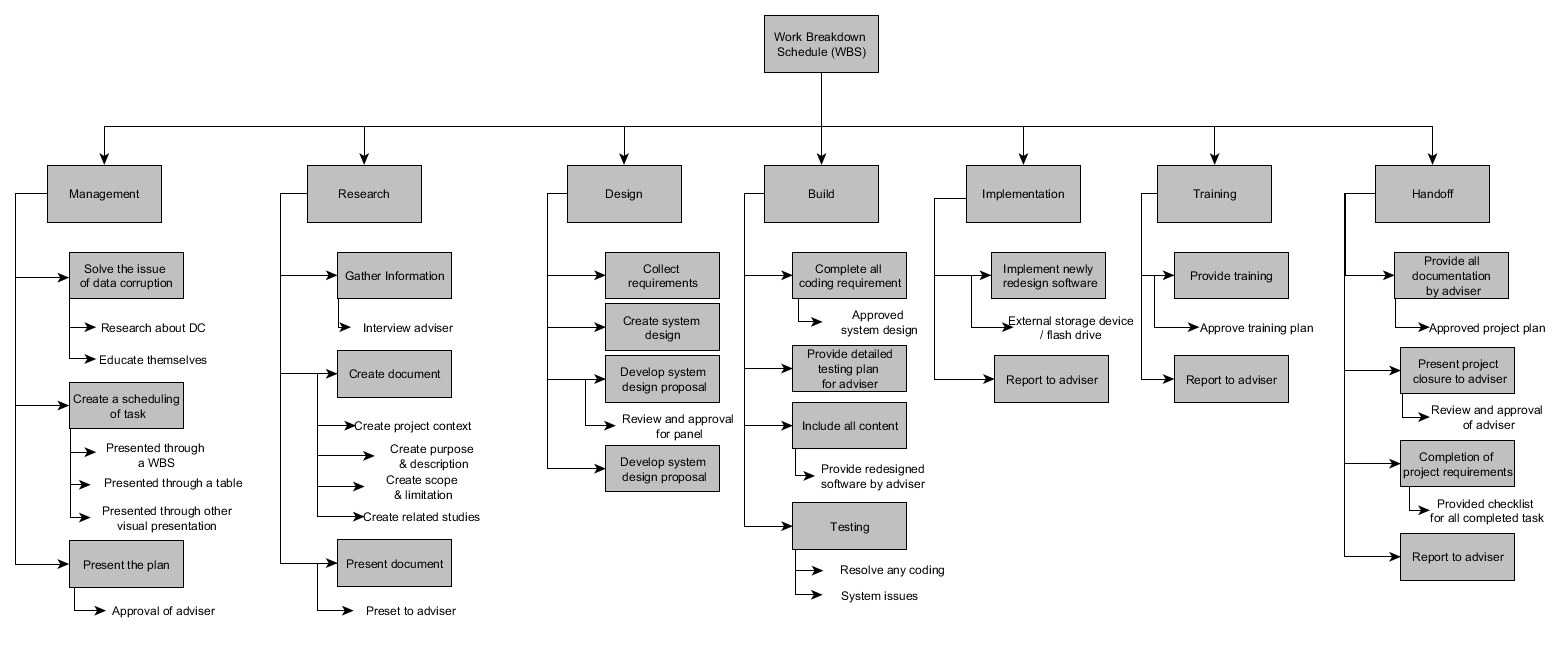
### Composite Structure Diagram



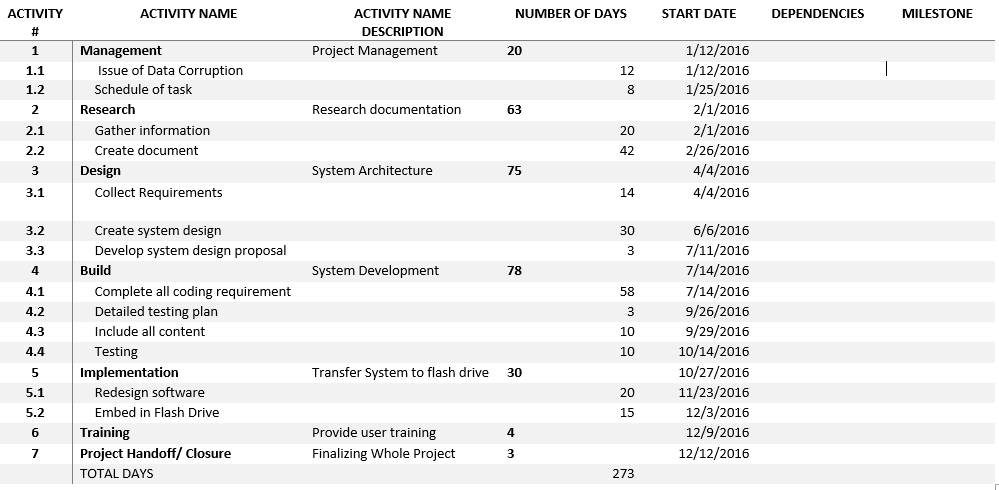
### Deployment Diagram



### Work Breakdown Structure (WBS)



### Activity List



### GanttChart1Gantt Chart

