Approach for Data Centre Performance Measurement Solution

By Petrus Francois de Kock

Submitted in partial fulfilment of the requirements for the degree of

BACHELORS OF INDUSTRIAL ENGINEERING

in the

FACULTY OF ENGINEERING, BUILT ENVIRONMENT AND

INFORMATION

TECHNOLOGY

UNIVERSITY OF PRETORIA

OCTOBER 2011

Executive summary

iScript is a software developer and system integrator that has been charged with supplying their client with a system that has the ability to integrate all its clients current systems into one centralised point. The client seeks a solution that has the capacity to provide as far as possible real time data to be used in business decision making and other calculations required by legislation.

The deliverables include a data dictionary that contains all current systems used, their unit of measurement and a suggested standardised format to be used for comparison and quality management purposes. The scope of this project includes a literature study and theoretical conceptual design, but excludes final implementation.

The project plan is broken down as follows:

- 1. Define measurements
- 2. Set benchmarks
- 3. Obtain measurements
- 4. Define scorecard
- 5. Present measurements

The use of SiSense's Prism ElastiCube Manager will simplify the decision making process by putting information that is currently housed in separate, nonstandard information systems in one convenient location. It will also simplify calculations such as total running cost of a specific data centre and carbon footprint produced by placing data on dashboards displaying the necessary KPI's.

Table of Contents

Exe	ecutive	summary	2	
List	of figu	res	5	
1.	Introd	luction	6	
2.	Problem Statement and Project Goals			
	2.1	Background	6	
	2.2	Project Aim	7	
	2.3	Project Scope	7	
	2.4	Project Deliverables	8	
	2.5	Project Constraints	8	
3.	Litera	Literature Review		
	3.1	Business Intelligence	9	
	3.2	Data Warehousing	12	
	3.2.1	Building a data Warehouse: Getting started	13	
	3.3	Pitfalls to avoid	17	
	3.4	Enterprise Architecture	18	
4.	Data Analysis			
	4.1 2007)	Integrated Information Infrastructure Reference Model (III-RM) (The Open G 20	roup,	
	4.1.1	Background	20	
	4.1.2	4.1.2 Components of the model		
	4.1.3	Key Business Drivers	20	
	4.1.4	4.1.4 Domains		
	4.2	III-RM Scope	23	
	4.3	Integrated Information Infrastructure Reference Model	24	
	4.3.1	Business View	24	
	4.3.1.	1 Location Type Model:	24	
	4.3.1.	2 Location Model:	24	

	4.3.	.2	Application View	26
	4.3.	.2.1	Application Model:	26
	4.3.	.3	Data View	27
	4.3.	.3.1	Logical Data Model:	27
	4.3.	.3.2	Performance Data Measurements:	28
	4.3.	.4	Technology View	28
	4.3.	.4.1	Infrastructure Class Model:	28
	4.3.	.4.2	Information Infrastructure Reference Model View	29
5.	Conceptual Design			
6.	Proposed Approach for Complete Solution			
7.	Conclusion			
Bib	liogra	aphy		36
Арј	pendi	ix		37
	a) Performance Data Measurements			
	b)	Exc	cel Spreadsheet with Captured Data	44
	c)	Dat	a Cube Modelling	46
	d)	Das	shboards	47

List of figures

Figure 1: Proposed Solution	8
Figure 2: Example of a Data dictionary created	10
Figure 3: The closed-loop in BPM approach (Matteo Golfarelli)	11
Figure 4: A complete architecture for BPM (Matteo Golfarelli)	12
Figure 5: The data model (Inmon, 2000)	13
Figure 6: Means by which information requirements can be collected (Inmon, 2000)	14
Figure 7: Selecting the correct sub area to investigate for the first iteration is	important
(Inmon, 2000)	15
Figure 8: Physical attributes (Inmon, 2000)	15
Figure 9: Defining the system of records (Inmon, 2000)	16
Figure 10: Moving transformation information to Meta data store (Inmon, 2000)	16
Figure 11: The transformation is turned into code (Inmon, 2000)	17
Figure 12: All levels of Architecture interact with each other	18
Figure 13: Enterprise Architecture as a cross layer view of aggregate artefacts (Fis	cher, May
2007)	19
Figure 14: Enterprise Portal Approach in III-RM	21
Figure 15: Based on the client's Value Chain Reference Model policy document	23
Figure 16: Four architecture models	24
Figure 17: III-RM View	29
Figure 18: Proposed Solution	30
Figure 19: Proposed Approach as discussed in the project proposal	32
Figure 20: Information to be included in the data dictionary	32
Figure 21: Data file example	33
Figure 22: Example of a score card	33
Figure 23: Example of a Dashboard to illustrate KPI's	34

1. Introduction

In the ever changing modern business environment having the correct processed data at the right time at the right place at the fingertips of the right people is becoming ever more important on both an everyday operational decision making level, as well as a strategic level.

The client operates a number of database software in a number of diversely different locations. These systems produce a vast amount of non-standard data every day.

The client would like to have all this unprocessed data turned into usable information located in a convenient locality where it can be used by those who need it.

2. Problem Statement and Project Goals

The client contacted iScrip to help identify and implement a feasible solution to the data management problems they are having.

2.1 Background

The client operates a number of data centres, switches and remote hubs that are distributed throughout South Africa. Each one of these facilities hosts an assortment of hardware essential to the successful operation of each location. Hardware ranges from network equipment to electric generators, climate control - and security systems. Each centre also hosts a number of monitoring and control systems to govern the functionality of all operating systems located at the various sites. These systems generate a vast amount of measurements and because not all monitoring equipment is sourced from the same manufacturer they often produce non-standard data. One small example of this non-standard data is the temperature readings taken inside the data centres. Different equipment produce different temperature scales (Celsius, Fahrenheit, and Kelvin). Maintaining the correct temperature is essential to the longevity of the electronic equipment housed inside the data centres. Currently a number of stumbling blocks are preventing the customer from generating reports essential to its operational requirements

• Different data centres are equipped with different types and makes of equipment, therefore making it difficult to compare the measurements of sites with each other.

- Equipment produces measurements in different formats, for example, some equipment will measure temperature in Degree Celsius, others in Fahrenheit and others in Kelvin.
- A number of software applications are used to produce measurements or extract measurements from equipment, often these applications performs the same function, thereby producing duplicate measurements.
- At the moment no benchmarks/ score cards are set for any of the measurements, making it difficult to indicate if a particular measurement is acceptable or out of the norm.

The goal is to provide a Boundaryless Information Flow (BIF) system that will give the client a competitive advantage over its rivals. BIF is a desired state for an enterprise's infrastructure and is specific to the business needs of the organisation. An infrastructure that provides BIF can be characterized with a set of standard components that provide services in a customer's extended enterprise that, combine multiple sources of information, deliver information to the places where that information is needed, and in the right context for the people or systems using that information. To achieve BIF an organisation needs to put in place infrastructure services that bring information together and provides that information to those users and applications that need it.

2.2 Project Aim

The aim of this project is to identify all shareholders, systems currently used by the client, and the data generated and collected by these systems. After the discovery process is completed the information that was gathered can be used to suggest and implement a suitable data warehouse system that will help store all the relevant information in one place. The largest problem the client is experiencing is that the raw data being produced is not standardised. The second aim of this project is thus to suggest a standard format for all relevant data. Once the data is standardised all relevant information will be presented on dashboards using Cube modelling software. This standard data can then be compared with sites of a similar specification. Another advantage of having standardised data is that it can be used for derived calculations such as the Carbon Footprint produced across all sites.

2.3 Project Scope

The focus of this project will be to present a data warehouse system, standardise format for all relevant data, and a dashboard solution to place the standardised data at the finger tips

of the client's business analysts. Due to time restraints on completing this project only a prototype of all the above mentioned systems and documentation will be presented.

2.4 Project Deliverables

The following will be included as project deliverables:

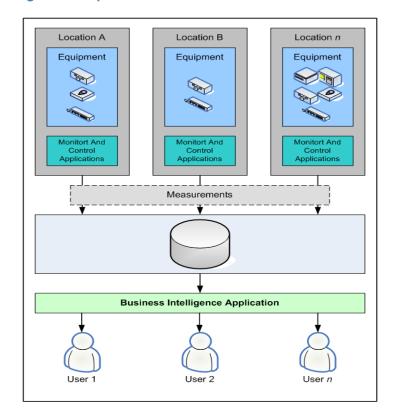


Figure 1: Proposed Solution

- Templates for standardised data
- Prototype Business Intelligence Application

2.5 Project Constraints

The project was subjected to the following constraints:

- Time was a major constraint. Not only did the student have other obligations in terms of commitments towards his studies at the University of Pretoria, but also his project sponsor's fork was filled with a considerable amount of hay.
- The technical experience of the student was another constraint.
- The client was very reluctant to grant access to an outsider to its system.

3. Literature Review

During the completion of this project I will have to deal with a new concepts such as, Data Warehousing and all it entails. This include OLAP cube modelling, Business Performance Management, and Enterprise Architecture. Therefore the literature review is quite cumbersome in an attempt to touch on all of these points.

3.1 Business Intelligence

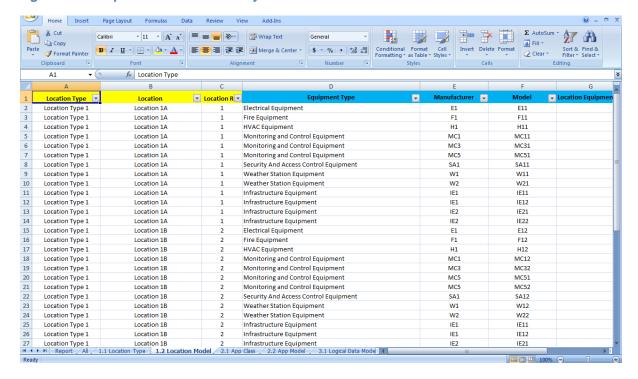
"Enterprise executives understand that timely, accurate knowledge can mean improved business performance." (W.F. Cody, 2002)

According to (W.F. Cody, 2002) through the usage of business intelligence modern database management can be expanded to build even larger data warehouses, and to employ data mining techniques to extort business advantage from large amount of data presented. They go on to claim that in the modern business environment it is becoming more and more important "to take advantage of all available information." However the challenge is compounded by enterprises producing increasing amount of raw data and the need for more people to have access to standardise data. Business have been investing in business intelligence and then further in data warehousing as a tool to gain a competitive edge.

Data warehousing is a procedural approach to collecting data in a single point. Here it can be standardised and tested against preset norms to allow for analysis and business decisions to be based upon. The origins of raw data are referred to as operational data stores (ODS). After the raw data has gone through a process called data cleansing, to detect and resolve any irregularities, the processed records are modelled on an OLAP cube. The OLAP cube is an online analytical processing multidimensional model. Tools are available from various vendors and during this project we will have to select one that will best integrate with the client's current information system.

In another article (Matteo Golfarelli) claims that merely making use of bottom-up extraction of data while making use of a data warehousing process is not enough to enforce the company's strategy. A new approach called Business Performance Management in emerging. Although it includes data warehousing, it also requires "a reactive component capable of monitoring the time-critical operational processes to allow tactical and operational decision-makers to tune their actions according to the company's strategy." Business Intelligence can typically be defined as the process of turning "data into information and then into knowledge." (Matteo Golfarelli)

Figure 2: Example of a Data dictionary created



BI was developed in the 1990's for the needs of modern managers who demanded more data on a more frequent basis. After academics to up the challenge the field has been greatly refined and today we are at the point where the introduction of a BI system is much more focussed on the implementation of sound business processes and adopting an end-toend strategy to align the internal strategy from supplier to client. Managers are thus forced to measure themselves continuously against the KPI's and scorecards that are benchmarked by the new strategy. The success of the strategy depends on it's the ability for goal sharing and measurements of KPI's to be available across all levels of the business. Allowing the strategy to be visible throughout the enterprise makes it clear to employees what managers expect of them. "The DW process essentially helps managers to understand their companies by supporting bottom-up extraction of information from data." (Matteo Golfarelli). Unfortunately the means that the DW process fails to enforce the top-down strategy of the company. The turning point in BI came when it stopped being a technique for data mining, but an active and tangible system for effective business management. This new representation of BI is referred to as Business Performance Management (BPM). This approach helps business manage all its resources, such as people, machines, time and money, in an optimised manner.

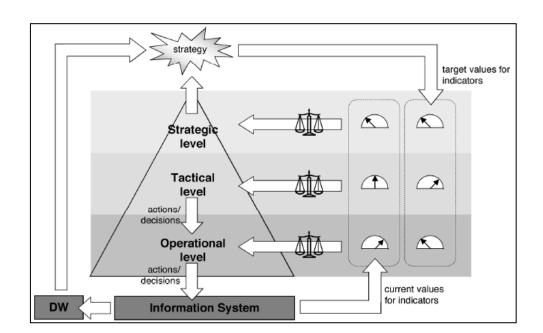


Figure 3: The closed-loop in BPM approach (Matteo Golfarelli)

Currently the BPM solution set forth by IT vendors is a classic multi dimensional OLAP coupled with a data integration system.

The complete BPM approach is graphically illustrated in figure 4. It can be seen that the user inputs is processed by the BPM process, all the different information systems is brought together inside a Data Warehouse which has as its output an OLAP model. The KPI's is tracked by both reports generated and a dashboard that displays the information in a graphical format. This will allow the users of this data to test it against the pre-set standards described in the scorecards produced.

In conclusion Matteo Golfarelli has the following to say:" DW systems led to quantify business information, to make it promptly available and certified. On the other hand, the role of BPM is to quantify the enterprise strategy and targets, in order to decentralize decision making. DW is not enough to this end since its technology is neither suitable for the grain nor for the freshness of the collected information, that should quickly flow throughout the different levels of the company."

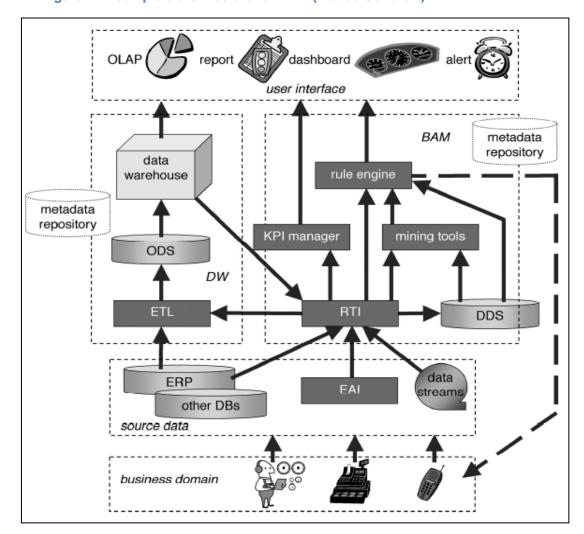


Figure 4: A complete architecture for BPM (Matteo Golfarelli)

3.2 Data Warehousing

The data warehouse has a few requirements it must adhere to

- The data warehouse must make an organisation's information easily accessible
- It must present the organisation's information consistently
- It must be adaptive and resilient to change
- The data warehouse must be a secure bastion that protects our information assets
- It must serve as the foundation for improved decision making
- The business community must accept the data warehouse if it is to be successful (Ralph Kimball, 2002)

3.2.1 Building a data Warehouse: Getting started

When building a data warehouse it is not possible to know the entire set off requirement before the data warehouse is built and information is available for analysis. "For this reason, the data warehouse is built in iteratively, in small, fast bursts of development". (Inmon, 2000)

The first small burst cannot be done by embarking on a massive requirement discovery mission. The developer has to move through the steps of design and development quickly. This in radically deferent from the way classical operational environment has traditionally been build.

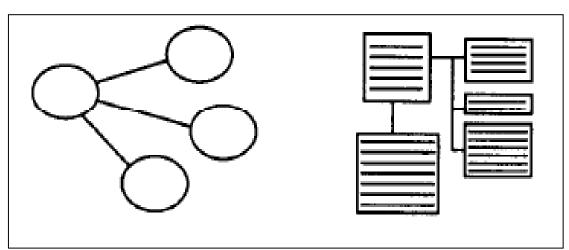


Figure 5: The data model (Inmon, 2000)

The first step is the development of the data model. This serves as the roadmap for development. The corporate data model is combined with the new know information requirements of the client. A time box method is commonly used. This is where a deadline is set and only requirements known by this point in time are added to the model. This is done to ensure that the discovery process does not continue indefinitely and the first iteration gets build.

The data model has to consist of at least a high level - and a mid level model. The high level model will contain all the major enterprise entities and their relationships. The mid level model illustrates all attributes and how they are grouped together.

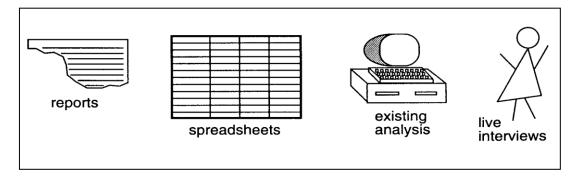
Part of the prerequisites for the development process is to select both the hardware and software that will be used. This depends on a few factors: (Inmon, 2000)

- The volume of data to be accommodated
- The speed with which data is needed
- The history of the organisation
- Which level of data is being built
- How many users there will be
- What kind of analysis is to be performed
- The cost involved

Both hardware and software must be selected. The assumption is made that the client, due to its relative size in the market, will have the mainframe capacity to process the amount of data. The challenge will be to identify the optimal software to meet all demands and restrictions.

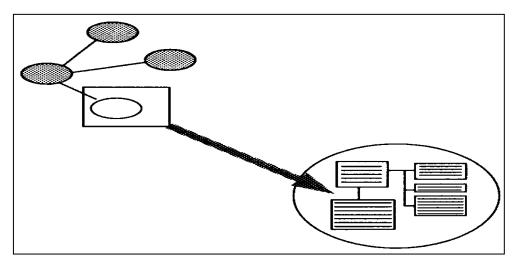
Next a process of information discovery needs to be embarked upon. This can be done by reviewing reports, spreadsheets, existing analysis and by interviewing the employees that will be using the data generated by the warehouse.

Figure 6: Means by which information requirements can be collected (Inmon, 2000)



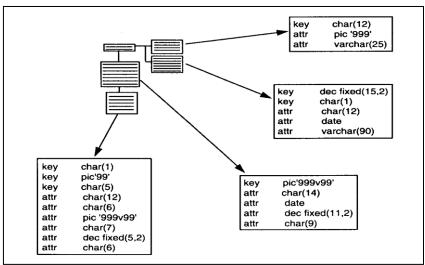
The first iteration should then be build. During this step it is important to load enough data to be meaningful but also little enough so that it can be done quickly. We also have to make sure that the right functional area is targeted to generate meaningful knowledge for the following iterations.

Figure 7: Selecting the correct sub area to investigate for the first iteration is important (Inmon, 2000)



After selecting the meaningful functional areas the next step is to add the physical attributes.

Figure 8: Physical attributes (Inmon, 2000)



Here we need to move back into physical data base design mode with the first step being specifying the physical attributes from the data model as shown in Figure 8. In this step it is also required to define the unit of time that exists in each unit of data warehouse data. It has to be kept in mind that the unit of time selected determines the occurrence of data.

The system of records can also be defined as the source data. In the case of our client these would be all the traditional data basis' that is already in place. It must be noted that these data sources must be complete, accurate, and timely.

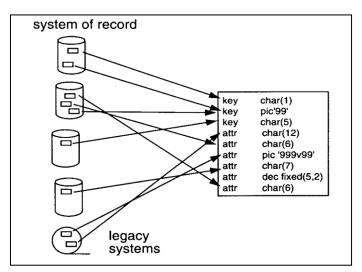


Figure 9: Defining the system of records (Inmon, 2000)

The transformation needs to be specified next. After specification the transformation is transferred to the metadata infrastructure that sits above the data warehouse. The metadata houses information regarding the origins of data and transformations it has undergone. The regularity of transformation depends on the amount of data, the complexity of the data, and the business need for the data.

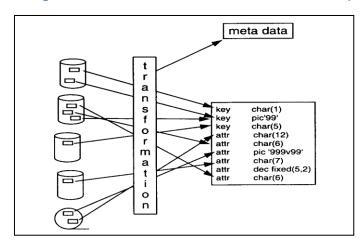


Figure 10: Moving transformation information to Meta data store (Inmon, 2000)

The transformation can now be turned into code. With the use of an automatic code generator this is an easy task. Without one of these software applications the coding needs to be done manually. This is not a straightforward exercise. If the data is automatically captured from the legacy system into the metadata it decreases the likelihood of errors occurring.

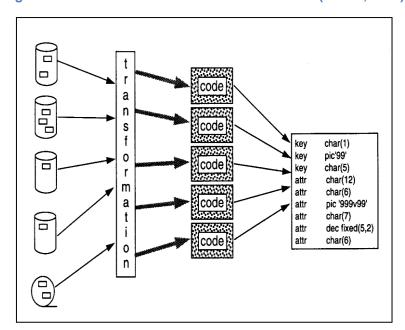


Figure 11: The transformation is turned into code (Inmon, 2000)

3.3 Pitfalls to avoid

In his book, The data warehouse toolkit, Ralph Kimball describes a number of pitfalls to be avoided when creating a data warehouse.

- Falling in love with the data and technology rather than paying attention to the business's needs and goals.
- Fail to recruit the correct project sponsor within the client company.
- Attempting to embark upon a massive multi-year project, instead of breaking it down into manageable iterative spurts.
- Spending time and money on building a normalised database, and then running out
 of money before being able to construct a workable presentation area based on
 dimensional models.

- Being more concerned with the ease of programming than with the performance of query handling.
- Make the data presentation platform excessively complicated and not used friendly
- Populate dimensional models without consideration for data architecture that binds them with shared, conformed dimensions
- Load only summarised data into the presentation platform
- Presume that the business and all its dimensions are static with regards to time
- Keep the end user in mind. All project should have as its main goal the clients satisfaction

3.4 Enterprise Architecture

According to ANSI/IEEE Std 14712000, architecture is defined as the "fundamental organisation of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution" (IEEE 2000). Enterprise architecture (EA) therefore is understood as (1) the fundamental organisation of a government agency or a corporation, either as a whole, or together with partners, suppliers and / or customers ("extended enterprise"), or in part (e.g. a division, a department, etc.) as well as (2) the principles governing its design and evolution (Opengroup 2003) . (Fischer, May 2007)

In the past EA has focussed only on the IT related artifacts, but it is crucial that in the modern business environment that EA should in addition investigate business related artifacts such as organisational goals, markets, KPl's, business process and the product a company produces. Once business artifacts is covered by EA it can be used to for management activities.

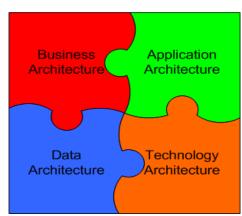


Figure 12: All levels of Architecture interact with each other

Enterprise architecture comprises of four layers. (Fischer, May 2007)

Business architecture

 Represents the elementary structure of the business. Artifacts most commonly found on this level include, relationships with customers, market segments, services presented, organisational goal etc. Design and evaluation can be done according to Porter's value chain.

Process architecture

 Is concerned with the services offered by a business and how they are structured inside the business. Typical artifacts include responsibilities, KPI's, business process and flow of information.

Integration architecture

Technology Architecture

- Essential organisation of BI systems elements. Includes data flow,
 Application groupings and Enterprise services.
- Technology (or infrastructure) and Software architecture
 - o Represents communication networks. Also includes computer services

Business
Architecture

Process
Architecture

Integration
Architecture

Software
Architecture

Enterprise

Figure 13: Enterprise Architecture as a cross layer view of aggregate artefacts (Fischer, May 2007)

4. Data Analysis

4.1 Integrated Information Infrastructure Reference Model (III-RM) (The Open Group, 2007)

4.1.1 Background

With the shift of business into a cyber environment the primary return on investment for most organisations have moved from the Application Platform space to the Application Software space. TOGAF Technical Reference Model (TRM) focuses on the Application Platform Space. Therefore we need to look at a reference model that has as its goal the Application Software platform. Hence Integrated Information Infrastructure Reference Model (III-RM).

4.1.2 Components of the model

III-RM has two components that assumes that a computing and network platform exist and thus do not cover these in the model as in TRM.

- A taxonomy, which defines terminology, and provides a coherent description of the components and conceptual structure of an integrated information infrastructure.
- An associated III-RM graphic, which provides a visual representation of the taxonomy, and the inter-relationship of the components, as an aid to understanding.

4.1.3 Key Business Drivers

"The Boundaryless Information Flow problem space is one that is shared by many customer members of The Open Group, and by many similar organizations worldwide. It is essentially the problem of getting information to the right people at the right time in a secure, reliable manner, in order to support the operations that are core to the extended enterprise." (The Open Group, 2007)

For years the approach to maximise a large business' efficiency was to optimise each department individually. This approached ensured that specialised skills developed in employees, who applied their skills only in specialised areas of a business' activities. In the current climate it is speed, flexibility, an quick reactions to ever changing markets that make the difference between failure and success.

The solution is to provide staff with two things (The Open Group, 2007):

- Integrated information so that different and potentially conflicting pieces of information are not distributed throughout different systems.
- Integrated access to that information so that staff can access all the information they need and have a right to, through one convenient interface.

One way of doing this is to create "enterprise portals" that provides access to information via a web-based interface. (Coloured section at end of cylinder)

Sell Space Customer Support Selling Internal Space Manufacturing Legal Finance Assembling Online Buy Space Systems Procuring Design Systems ERP Systems Requirements Systems Systems Procurement Systems

Figure 14: Enterprise Portal Approach in III-RM

4.1.4 Domains

As with EA there are four domains and their component parts are:

Application Architecture:

Model	Purpose Objects	Example
Application Class	To provide a conceptual classification of	Business System,
Model	the various types of applications required	Infrastructure Control
Model	to support business.	System
	To provide a list of the physical business	
Application Model	and infrastructure applications used within	
	the organisation.	

Business Architecture:

Model	Purpose Objects	Example
Location Type Model	To provide a grouping of locations based on	Data Centre, Remote
Location Type Model	common functionality or purpose.	Hub
Location Model	To provide a breakdown of the physical locations	14 Avenue, Rustenburg
	relevant to architecture.	14 Avenue, Rustenburg

Technology Architecture:

Model	Purpose Objects	Example
Infrastructure Class Model	To provide a conceptual classification of the various types of infrastructure devices required to support business.	HVAC, Security, Electricity
Infrastructure Device	To define the physical infrastructure devices that	Generator, Server, AC
Model	is implemented within the business.	Unit

Data Architecture:

Model	Purpose Objects	Example
	To describe the logical data structure of the	
Logical Data Model	organisation's data stores in terms of entities,	
	attributes and relationships.	

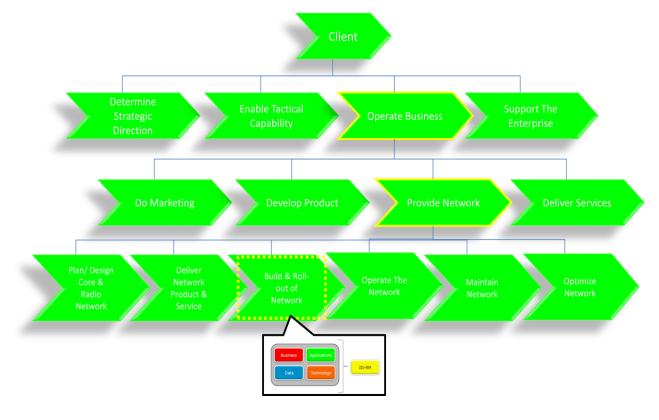
Reference Models (III-RM):

Composite Model	Purpose Objects
	To depict the logical design of IT solutions in terms of the
Logical Solution Design Model	logical data, application, environment and technology
	components.

4.2 III-RM Scope

The following figure illustrates the organisational scope of the project.

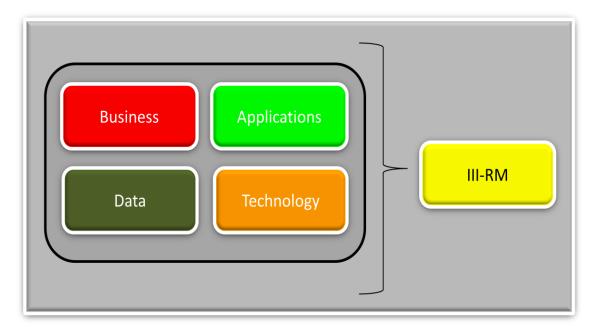
Figure 15: Based on the client's Value Chain Reference Model policy document



4.3 Integrated Information Infrastructure Reference Model (The Open Group, 2007)

III-RM consists of 4 architecture models:

Figure 16: Four architecture models



4.3.1 Business View

4.3.1.1 Location Type Model:

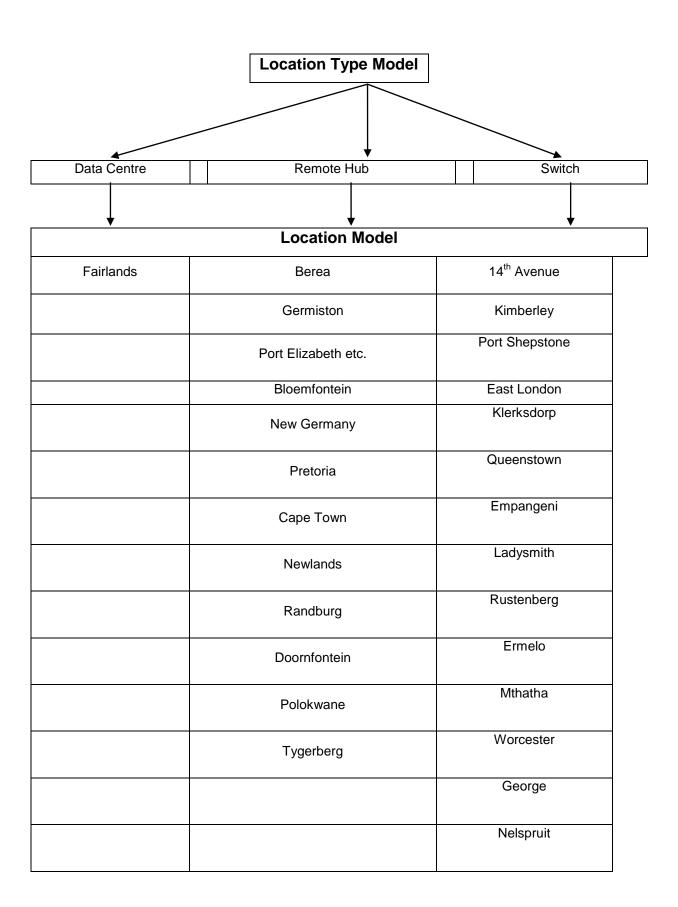
The purpose of this model is to provide a grouping of locations based on common functionality or purpose:

- Data Centre
- Remote Hub
- Switch

4.3.1.2 Location Model:

The purpose of this model is to provide a breakdown of the physical locations relevant to the enterprise architecture: Example:

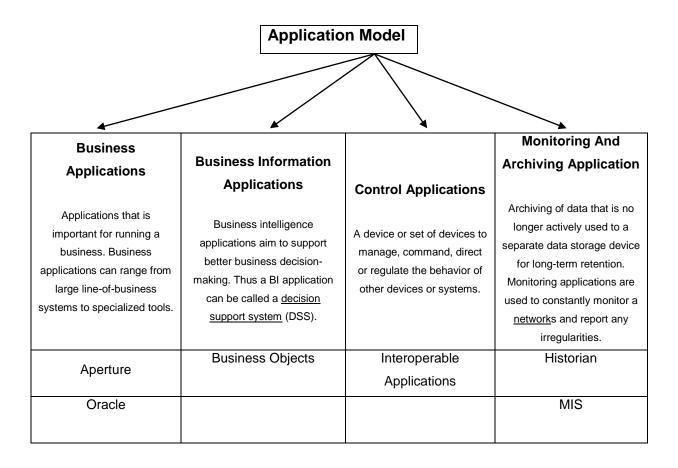
- Berea
- Germiston
- Port Elizabeth etc.



4.3.2 Application View

4.3.2.1 Application Model:

This model provides a list of the physical business and infrastructure applications used within the organisation.



4.3.3 Data View

4.3.3.1 Logical Data Model:

The purpose of the model is to describe the logical data structure of the organisation's data stores in terms of entities, attributes and relationships.

	Logical	Data Model	
Device Data:	Maintenance Data:	Application Data:	Performance Data:
Manufacturer	Maintenance Contract	Name	Measurement Name
Model	Maintenance Provider	Description	Measurement Description
Model Info	Last Service (Days)	Owner/ Administrator	Unit
Barcode Number	Service Date	Agreement Type	Frequency of Measurement
MTN Site	Service Type		
Floor	Service Status		
Room		ı	V
	_		Data Model
		_	

Performance Data				
Alarm	Flow	Speed	Wind Direction	
Atmospheric Pressure	Frequency	Runtime	Wind Speed	
Capacity	Humidity	Solar Radiation	Conductivity	
Phase Difference	Status	Consumption	Power	
Temperature	Current	Pressure	Torque	
Efficiency	Rain Fall	Voltage		

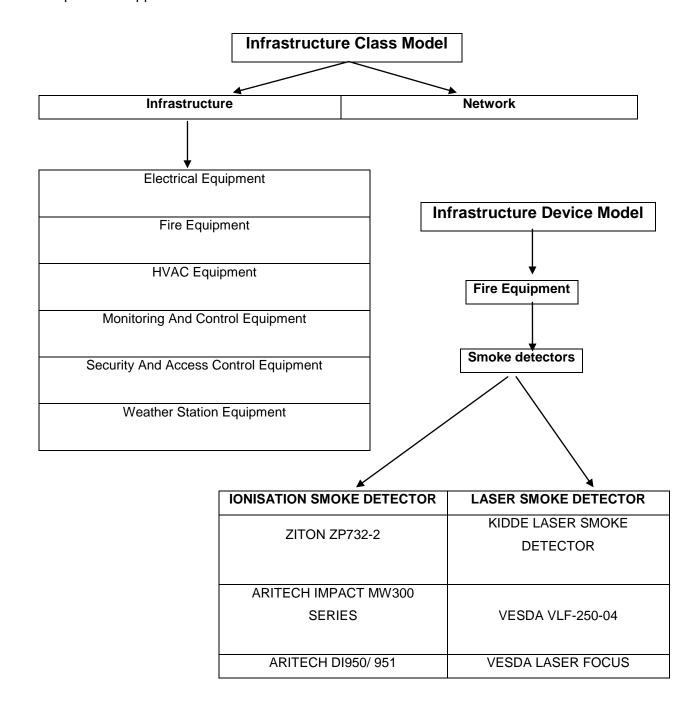
4.3.3.2 Performance Data Measurements:

Score cards were created to describe all the terms mentioned in the above Performance Data diagram. Please see Appendix for further detail.

4.3.4 Technology View

4.3.4.1 Infrastructure Class Model:

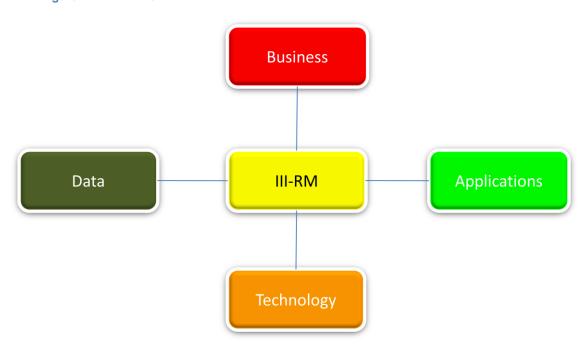
This model provides a conceptual classification of the various types of infrastructure devices required to support business.



4.3.4.2 Information Infrastructure Reference Model View

By inspecting the graphical representation of the III-RM view it can be seen that it describes the deployment of business, application, technology, and data elements as well as they're geographical location.

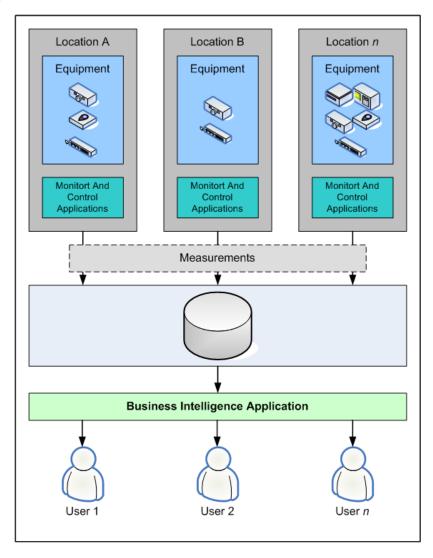
Figure 17: III-RM View



5. Conceptual Design

In order to meet the client's needs a system is proposed to enable the client to gather data from all sites and turn it into standardised information. This information will then be classified in such a manner as to enable different sites to compare related reports with each other. As the information will be stored in a centralised location the measurements could then be used for derived calculations and operational decision making.

Figure 18: Proposed Solution



During the analysis phase certain Performance Data Measurements were defined and score cards generated. The client could then use these score cards to define the frequency of which measurements need to be captured. The score cards also provide the client with the

opportunity to define benchmark values. This could be done using various standards. These could include but are not limited to.

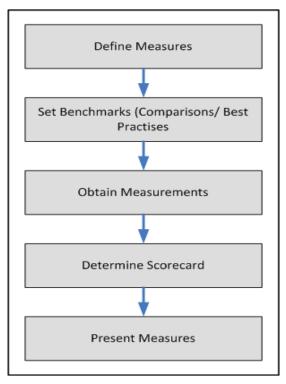
- Specifications provided by the manufacturer for operations under normal conditions
- Ideal operating conditions for maximised longevity (Temperature, humidity, etc inside a room which houses a large number of electronic equipment)
- Using a 3-sigma approach to determine a upper- and lower limit from historical data collected
- Defining true or false as the norm for systems that could only be in one of two states such as equipment that has alarms connected to it (smoke detector, security, etc.)

Apart from comparing standardised information with likewise data, all measurements can be weighed against the yardstick set out in the score cards. A process of data mining can then be started to determine the root cause of any measurement found not to fall within the parameters set. Let us imagine that the monthly fuel consumption across all sites is higher than the upper limit set out in the fuel consumption score card. The data mining can then reveal which site is using higher than normal amounts and further investigation could provide insight into a possible leak or maybe even theft. It could also give clues to electrical problems within the site or possibly just simply a higher demand.

One of the main wishes to client has is to determine its carbon footprint. This system will provide the basis for this calculation. If the client decides to make being environmentally friendly part of its business philosophy on a strategy level the system can be used to determine which of its vendors' products are best suited to minimise dependence on fossil fuels.

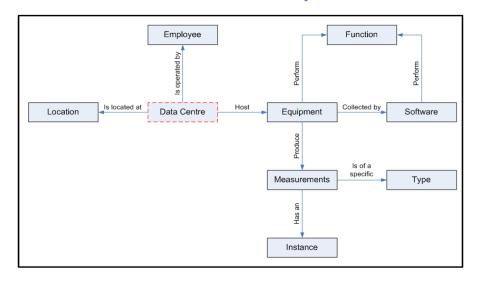
6. Proposed Approach for Complete Solution

Figure 19: Proposed Approach as discussed in the project proposal



The first step is to define all the measurements required. This is done by making use of a data dictionary approach.

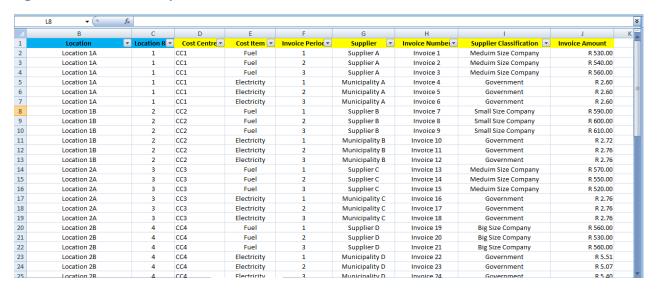
Figure 20: Information to be included in the data dictionary



The second step is to set the benchmarks and best practise comparison. As discussed in the conceptual design this step will be done after the first iteration was build. The first iteration is build only to determine further specifications as set forth by Inmon, 2000.

The third step is to obtain the measurements. The client opporates a number of BI software to capture this data and it will be uploaded into the final deliverd product from there. For the purpose if this I have populated a number of Excel spreadsheets that can be found in the Appendix.

Figure 21: Data file example



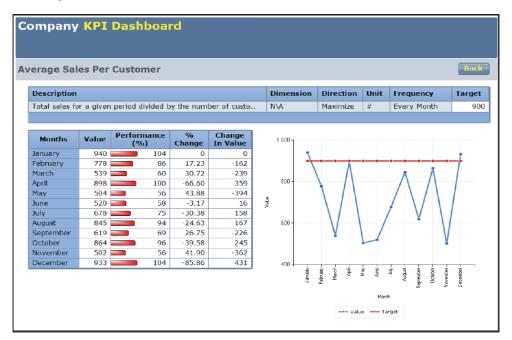
The fourth step is to determine the score cards. The score cards can be found in the Appendix.

Figure 22: Example of a score card

Temperature:				
Measurement:	Temperature			
Description:	Temperature is a physical property of matter that quantitatively expresses the common notions of hot and cold.			
Unit:	°C (Degrees Celsius)			
Frequency Of Measurement	Actual Values			
Capture	Daily Values			
	Monthly Values			

The fifth and final step is to present the data obtained in a dashboard. Dashboard in Appendix.

Figure 23: Example of a Dashboard to illustrate KPI's



7. Conclusion

Modern business has evolved into an organism that demands a large amount of information in order to keep a finger on the pulse of all its internal systems. By implementing the data warehouse system proposed and making use of the SiSense Prism Cube data modelling software the client will have a process that it can follow to standardise the masses of data it produces and then be able to store it in one centralise location. Once the data is standardise and centralised the client can use this information for both strategic and operational purposes.

By placing all its information in one place the client will save money by saving time and energy on data analysis.

In conclusion this project holds the following prospective advantages for the client:

- Quicker turn-around times on the interpretation of data as it will now be standardised.
- Interpreting data will require less employees. This will free up manpower to be utilised in other areas of the business which will improve productivity.
- Enhanced data mining will help to get to the bottom of anomalies in compared results faster.
- Having all the required information in one place will help the client reach its strategic goal of being more environmentally conscious, by making it easy for them to see which of its current hardware systems has the least adverse effect on the environment.

When all of these advantages is taken into account it comes down to one thing. The implementation of a data warehouse coupled with a BI system will result in a meaningful reduction in the number of employees needed to interpret data and perform data mining in order to discover the root cause of anomalies. This means less personnel is required for the business analysis process, and as these are usually people with a considerable level of training and industry experience it will make a significant difference in salaries paid.

On the enhanced data mining side of things it means less resources will be lost due to infrastructure using more of a specific resource than normal due to maintenance requirements. It will also curb losses brought on by a criminal element.

At the end of the day the proposed system will help the client focus on the functional areas of its business that contributes to its bottom line at the end of the financial period.

Bibliography

(n.d.). Retrieved from Wikipedia: www.wikipedia.org

Client. Client's strategic planning document.

Fischer, R. W. (May 2007). Essential Layers, Artifacts, and Dependencies. *Journal of Enterprise Architecture*, 1-12.

Inmon, W. (2000). Building the Data Warehouse: Getting Started. Billinmon.com, 1-19.

Matteo Golfarelli, S. R. (n.d.). Beyond Data Warehousing: What's next in Business Intelligence.

Ralph Kimball, M. R. (2002). The Data Warehouse Toolkit. Wiley Computer Publishing.

Robert Winter, R. F. (May 2007). Essential Layers, Artifacts, and Dependencies of Enterprise Architecture. *Journal or Enterprisse Architecture*, 1-12.

Roux, B. L. (2011). CI-IIIRM. Midrand: iScript.

The Open Group. (Version 8.1.1, Enterprise Edition). *The Open Group Architecture Framework (TOGAF)*. The Open Group.

The Open Group. (2007). TOGAF v8.1.1.

W.F. Cody, J. K. (2002). The integration of business intelligence and knowledge management. *IBM Systems Journal*, 697-713.

Zachman, J. (1987). A framework for information systems architecture. *IBM Systems Journal Vol 26 No.3*, 276-292.

Appendix

a) Performance Data Measurements

1. Score cards

Alarm:	
Measurement:	Alarm
Description:	An alarm gives an audible or visual warning about a problem or condition.
Unit:	(Positive/Negative) and (False Positives/False Negatives)
Standard Measurement	Off
Frequency Of Measurement	Actual Values
Capture	Daily Values
	Monthly Values

Flow:		
Measurement:	Flow	
Description:	The volume of gas/liquid travelling the period.	rough a tube during a specific time
Unit:	I/s (Liter per second)	
Standard	(Product Specific)	
Frequency Of Measurement Capture	Actual Values	
	Daily Values	
	Monthly Values	

Speed:		
Measurement:	Speed	
Description:	Revolutions per minute (RPM) is a measure of the frequency of a rotation. It annotates the number of full rotations completed in one minute around a fixed axis. It is used as a measure of rotational speed of a mechanical component.	
Unit:	RPM (Revolutions per minute)	
Standard	(Product Specific)	
Frequency Of Measurement Capture	Actual Values Daily Values Monthly Values	

Wind Direction:		
Measurement:	Wind Direction	
Description:	Wind direction is reported by the <u>direction</u> from which it originates	
Unit:	Cardinal directions or Zimuth degrees	
Standard:	Non	
Frequency Of Measurement Capture	Actual Values	
	Daily Values	
	Monthly Values	

Atmospheric Pressure:		
Measurement:	Atmospheric Pressure	
Description:	Atmospheric pressure is the force per unit area exerted against a surface by the weight of air above that surface in the Earth's atmosphere.	
Unit:	kPa (Kilo Pascal)	
Standard:	Non	
Frequency Of Measurement	Actual Values	
Capture	Daily Values	
	Monthly Values	

Frequency:		
Measurement:	Frequency	
Description:	Frequency is the number of occurrences of a repeating event per unit time . It is also referred to as temporal frequency.	
Unit:	hz (Hertz)	
Standard:	Product dependent	
Frequency Of Measurement	Actual Values	
Capture	Daily Values	
	Monthly Values	

Runtime:		
Measurement:	Runtime	
Description:	The period during which a machine operates without interruption.	
Unit:	hr (Hours)	
Standard	Product dependent	
Frequency Of Measurement	Actual Values	
Capture	Daily Values	
	Monthly Values	

Wind Speed :	
Measurement:	Wind Speed
Description:	Wind speed is the speed of wind, the movement of air or other gases in an atmosphere. It is a scalar quantity, the magnitude of the vector of motion.
Unit:	kn (Knots)
Standard:	Non
Frequency Of Measurement Capture	Actual Values
	Daily Values
	Monthly Values

Capacity:		
Measurement:	Capacity (Volume)	
Description:	Volume is the quantity of three-dimensional space enclosed by some closed boundary, for example, the space that a substance (solid, liquid, gas, or plasma) or shape occupies or contains.	
Unit:	I (Litre)	
Standard:	Product dependent	
Frequency Of Measurement	Actual Values	
Capture	Daily Values	
	Monthly Values	

Humidity:		
Measurement:	Humidity (Relative)	
Description:	Humidity is a term for the amount of water vapour in air, and can refer to any one of several measurements of humidity. Relative humidity is a measurement of the amount of water vapour in a mixture of air and water vapour. It is most commonly defined as the partial pressure of water vapour in the air-water mixture, given as a percentage of the saturated vapour pressure under those conditions	
Unit:	% (Percentage)	
Standard:	Centre dependent	
Frequency Of Measurement Capture	Actual Values Daily Values Monthly Values	

Solar Radiation:	
Measurement:	Solar Radiation
Description:	Solar radiation describes the visible and near-visible (ultraviolet and near-infrared) radiation emitted from the sun.
Unit:	(kW/m²) Kilowatts per square meter
Standard:	Non
Frequency Of Measurement Capture	Actual Values
	Daily Values
	Monthly Values

Conductivity:		
Measurement:	Conductivity	
Description:	The ability to which a specified material conducts electricity, heat or fluids calculated with the amount of heat/electricity/fluids that flows per unit time through a unit area.	
Unit:	Volume/Time	
Frequency Of Measurement	Actual Values	
Capture	Daily Values	
	Monthly Values	

Phase Difference:			
Measurement:	Phase Difference		
Description:	Phase difference is the difference, expressed in electrical degrees or		
	time, between two waves having the same frequency and referenced to		
	the same point in time.		
Unit:	° (Degrees)		
Frequency Of Measurement	Actual Values		
Capture	Daily Values		
	Monthly Values		

Status:				
Measurement	:		Status	
Description:			A word stating in what condition/state	a system or application is in.
Unit:	Jnit:		Status Word	
Frequency	Of	Measurement	Actual Values	
Capture			Daily Values	
			Monthly Values	

Consumption:				
Measurement:	Consumption (Electric Energy)	Fuel		
Description:	The form of <u>energy consumption</u> that uses electric energy.	The amount of fuel used to generate a specified amount of electric energy		
Unit:	kW·h (Kilo watt hour)	I/kwh (Litre per kilowatt hour)		
Standard:	Product dependent	Product dependent		
Frequency Of	Actual Values	Actual Values		
Measurement	Daily Values	Daily Values		
Capture	Monthly Values	Monthly Values		

Power:								
Measurement:	Power (Mechanical)		Power (Electrical) -		Power (Electrical) -		Power (E	lectrical)
			Apparent P	ower	Reactive Pov	wer	– Real Po	wer
Description:	Power (Mech	anical) is	Power in a	n electric	Reactive por	wer flow	Electric power is	
	the rate at w	nich work	circuit is th	e rate of	on the alt	ernating	the rate at which	
	is perform	ned or	flow of ene	ergy past	current trans	smission	electric e	nergy is
	energy is con	verted	a given poi	nt of the	system is ne	eded to	transferre	d by an
			circuit.	Apparent	support the transfer		electric ci	rcuit.
			Power the absolute		of real power over			
			value of	complex	the network.			
			power.					
Unit:	kW		kVA		kVAR		kW	
Standard:	Product dependent							
Frequency Of	Actual		Actual		Actual		Actual	
Measurement	Values		Values		Values		Values	
Capture	Daily		Daily		Daily		Daily	
	Values		Values		Values		Values	
	Monthly		Monthly		Monthly		Monthly	
	Values		Values		Values		Values	

Temperature:	
Measurement:	Temperature
Description:	Temperature is a physical property of matter that quantitatively expresses the common notions of hot and cold.
Unit:	°C (Degrees Celsius)
Standard:	Centre dependent
Frequency Of Measurement Capture	Actual Values
	Daily Values
	Monthly Values

Current :			
Measurement:	Electrical Current		
Description:	Electric current is a flow of electric charge through a medium.		
Unit:	A (Ampere)		
Standard:	Product dependent		
Frequency Of Measurement	Actual Values		
Capture	Daily Values		
	Monthly Values		

Pressure:				
Measurement: Pressure				
t			Pressure is the force per unit area applied in a direction perpendicular to the surface of an object (Pressure is an effect which occurs when a force is applied on a surface.	
Unit:			Bar	
Frequency Capture	Of	Measurement	Actual Values Daily Values	
			Monthly Values	

Torque:	
Measurement:	Torque
Description:	Torque is the tendency of a force to rotate an object about an axis,
	fulcrum, a torque can be thought of as a twist.
Unit:	Nm (Newton Meter)
Frequency Of Measurement Capture	Actual Values
	Daily Values
	Monthly Values

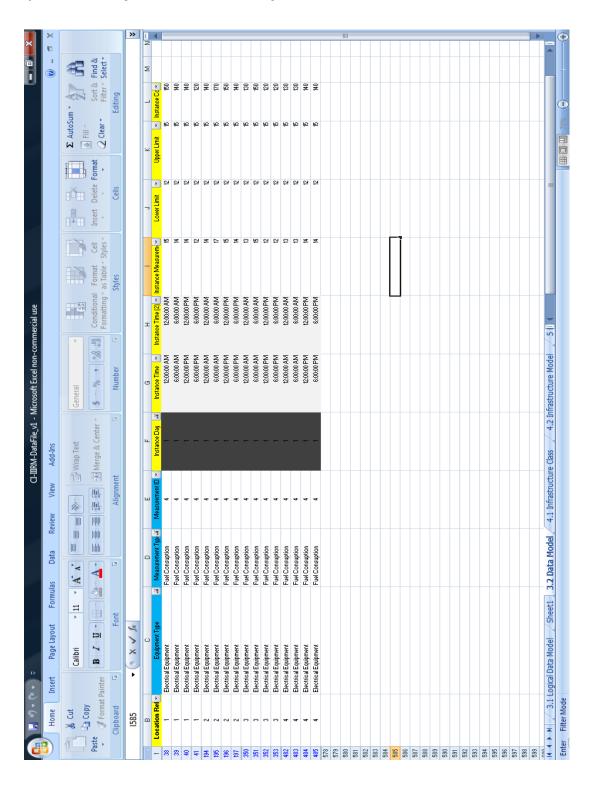
Efficiency:	
Measurement:	Efficiency
Description:	Efficiency in general describes the extent to which time or effort is well used for the intended task or purpose. Electrical efficiency reffers to useful power output per electrical power consumed.
Unit:	% (Percent)
Frequency Of Measurement Capture	Actual Values Daily Values
	Monthly Values

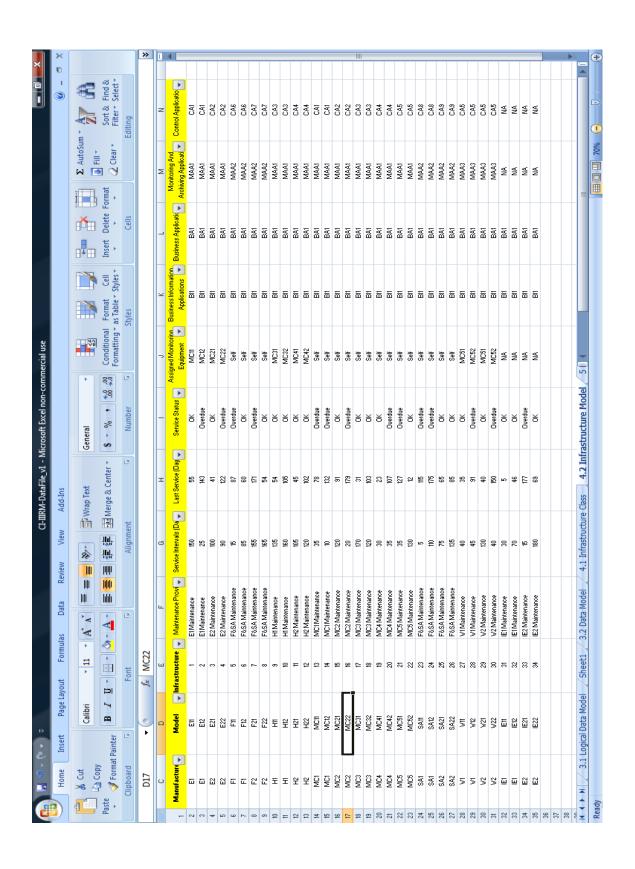
Rain Fall:		
Measurement:	Rain Fall	
Description:	Water condensed from atmospheric millimeter of rainfall is the equivaler meter	
Unit:	mm (Millimeter)	
Frequency Of Measurement Capture	Actual Values Daily Values	
	Monthly Values	

Voltage:			
Measurement:	Voltage		
Description:	Voltage is a measure of the energy of electricity, specifically, it is the energy per unit charge.		
Unit:	V (Volt)		
Standard:	Product dependent		
Frequency Of Measurement	Actual Values		
Capture	Daily Values		
	Monthly Values		

Definitions of measurements obtained from (Wikipedia).

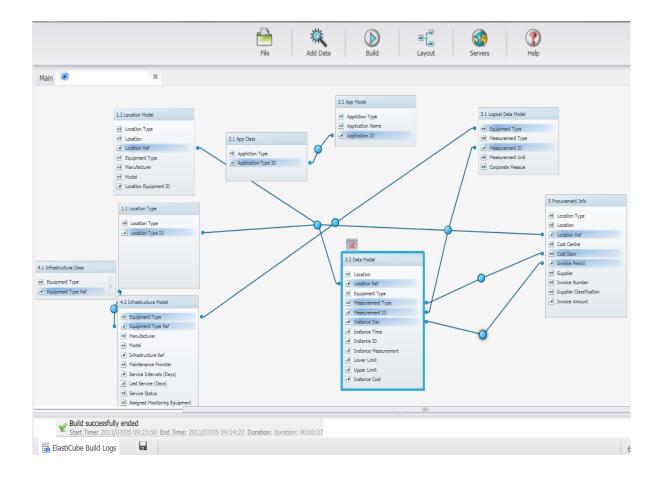
b) Excel Spreadsheet with Captured Data





Please note that in order to save paper, and thus trees, I have only inserted 2 screenshots, but as it can be seen there are many more. These will shown in the final project deliverable.

c) Data Cube Modelling



d) Dashboards



