UNIVERSITY OF THE WITWATERSRAND

Research Report IN-MEMORY Business Intelligence

A WITS Context

August 2014

Course Code	: ELEN7000 (Research Project)
Name	: Harshila Ravjee Dulabh
Student Number	: 8414034
Supervised by	: Prof Barry Dwolatzky Kevin Mc Loughlin

A research report submitted to the Faculty of Engineering and the Built Environment, School of Electrical Engineering, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree for Master of Science in Engineering.

Johannesburg 2014

DECLARATION

This project is being submitted for the Degree of Master of Science in Engineering in the University of the Witwatelll,,rsrand, Johannesburg.

I declare that this research report is my own unaided work. It is being submitted for the degree of Master of Science in Engineering (50% research report) in at University of the Witwatersrand, Johannesburg. It has not been previously submitted for any degree or examination in any other university.

____

(Signature of candidate)

19 day of AUGUST 2014

ABSTRACT

The organisational demand for real-time, flexible and cheaper approaches to Business Intelligence is impacting the Business Intelligence ecosystem. In-memory databases, in-memory analytics, the availability of 64 bit computing power, as well as the reduced costs of memory, are enabling technologies to meet this demand. This research report examines whether these technologies will have an evolutionary or a revolutionary impact on traditional Business Intelligence implementations. An in-memory analytic solution was developed for University of the Witwatersrand Procurement Office, to evaluate the benefits claimed for the in-memory approach for Business intelligence, in the development, reporting and analysis processes. A survey was used to collect data on the users' experience when using an in-memory solution. The results indicate that the in-memory solution offers a fast, flexible and visually rich user experience. However, there are certain key steps of the traditional BI approach that cannot be omitted. The conclusion reached is that the in-memory approach to Business Intelligence can co-exist with the traditional Business Intelligence approach, so that the merits of both approaches can be leveraged to enhance value for an organisation.

ACKNOWLEDGEMENTS

I would like to extend a special thank you to those who contributed to this paper:

Prof Barry Dwolatzky and Kevin McLoughlin, my supervisors, for their valuable comments and sharing their knowledge and providing guidance.

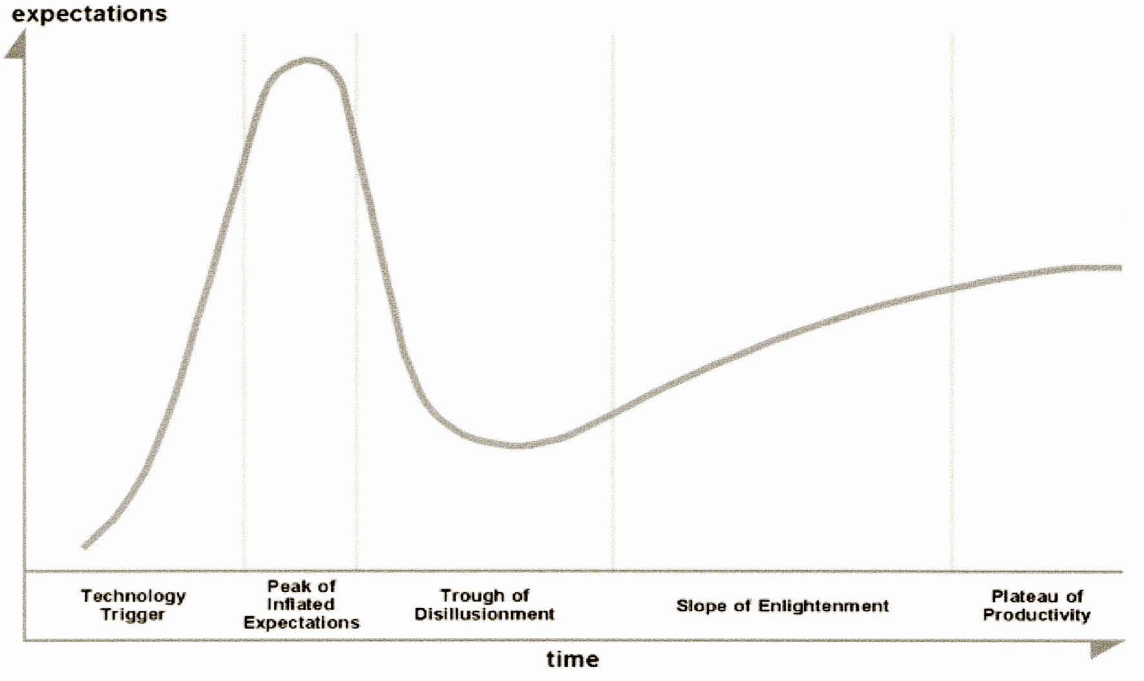
Emmanuel Mashandudze, who assisted with the server and licensing setup.

Zarina Hassim (Procurement Manager), Procurement category managers and buyers, for participating in the evaluation of the software.

Tony Bell, from Business Intelligent, for providing the QlikView temporary licenses.

TABLE OF CONTENTS

LIST OF FIGURES


expectations
Technology Trigger
Peak of Inflated Expectations
Trough of Disillusionment
Slope of Enlightenment
Plateau of Productivity
time

LIST OF ACRONYMS

Acronym	Meaning
ODI	Oracle Data Integrator
OWB	Oracle Warehouse Builder
OBIEE	Oracle Business Intelligence Enterprise Edition
OLAP	Online Analytic Processing
BI	Business Intelligence
ERP	Enterprise Resource Planning
ROLAP	Relational Online Analytic Processing
ETL	Extract, transform and load data
IT	Information Technology
ODS	Operational Data Store
IMC	In-memory computing
IMDB	In-memory databases
ERD	Entity Relationship Diagram
DW	Data Warehouse

1. INTRODUCTION

The Business Intelligence (BI) landscape is changing due to a higher demand for real-time analytic information to be made readily available to facilitate better decision making across all levels of an organisation, so that businesses can leverage their data to gain an insightful competitive advantage.

There have been transformative inroads in the way analytics are perceived within organisations and the way in which analytics are being used. The organisations that have adopted data-driven analytics to inform competitive strategies, are reaping the benefits. A diverse set of organisations, such as Amazon with online retailing, UPS that tracks the movement of packages and Barclays that has information-based customer management, have demonstrated the gain in competitive advantage, by using analytics pervasively within their organisations [1].

Thomas Davenport in his article 'Competing on Analytics' says that "Organisations are competing on analytics not just because they can – business today is awash in data and data crunchers – but because they should" [1]. This indicates that that it is imperative for businesses to adopt the use of analytics to maintain or gain a competitive advantage.

Advances in both software and hardware technology enabled organisations to achieve success in using data analytics. Business Intelligence has evolved from being focused on historical trends in order to project the future, to making decisions based on the current status [2]. Analytics are evolving and the requirement has moved towards data becoming more transparent and in-context, as well as being embedded in real-time applications. Predictive analytic results must be made available to users at the point of action or be inserted into the natural flow of processes [3].

This report evaluates the role of in-memory databases and in-memory analytics technology as enablers in delivering fast, flexible and near real-time Business Intelligence. The report also evaluates the advantages and the drawbacks of the technology, the adoption rate of the in-memory technology and the changes and investment that companies need to make in order to adopt the technology.

A further evaluation was done regarding the appropriateness of using an associative in-memory dashboard technology, within the Wits University context, which has a mature enterprise data warehouse and Business Intelligence infrastructure. An in-memory prototype of the current Wits procurement dashboards was built using the QlikView software. The initial dashboards were developed using a traditional BI approach. A survey was conducted on users of both old and new technologies to compare the user perceptions of the tools.

This research was conducted to determine whether the two types of Business Intelligence approaches can co-exist in an organisation such as University of the Witwatersrand (Wits), which has a mature BI infrastructure, or whether they are mutually exclusive technologies based on an organisational requirements and culture.

In particular the research was conducted to determine the role that in-memory databases and in-memory analytics may play in an environment that has an established and mature traditional BI infrastructure.

The questions investigated include:

- Are these two approaches to BI mutually exclusive or is there sufficient justification for them to co-exist within the same BI ecosystem?
- What are the business requirements that would require both technologies to be used?
- Does the in-memory associative approach shorten the development life cycle of a subject specific dashboard and deliver improved value, while at the same time being a more cost effective and faster approach than the traditional methodology used for BI?
- What are impacts and requirements of a real-time or near real-time offering on the business processes and procedures of an organisation?

2. HISTORY OF BUSINESS INTELLIGENCE

The term “Business Intelligence” (BI) was first used by Hans Peter Luhn in 1958, an IBM researcher, who described BI as "the ability to apprehend the interrelationships of presented facts in such a way as to guide action towards a desired goal." [4]

The term “Business Intelligence” evolved over the next 3 decades from Decision Support Systems to Executive Information Systems. In 1989 Howard Dresner described Business Intelligence as “concepts and methods to improve business decision, by using fact based support systems” [5].

A lot of the effort in the Business Intelligence ecosystem, since 1989, has been focused on developing and standardising BI architecture and the associated processes, as well as rationalising the hardware and software requirements, for data collection storage and retrieval. The terms Data Warehouse (DW), Data Marts (DM), Extract Transform and Load (ETL), Online Analytic Processing (OLAP) and Relational Online Analytic Processing (ROLAP) are pervasive and commonly adopted within the BI ecosystem. Harvey Koeppel refers to this period as the ‘the beginnings of transforming data into information and the use of information to help drive decision making”. [2]

There are two different views on what actually constitutes BI. The broader definition is "Business Intelligence is a set of methodologies, processes, architectures, and technologies that transform raw data into meaningful and useful information used to enable more effective strategic, tactical, and operational insights and decision-making." [5]. The second definition of BI is the use of the data collected in the data warehouse and data marts for analysis. That is, the data preparation and data usage are viewed as two different parts of BI but are closely linked. Forrester[1] refers to Business Intelligence as the reporting, querying, analysis and mining of data [5].

For the purposes of this report, the definition of BI is based on the first definition, where the entire BI architectural stack is considered to be BI.

The traditional BI approach of delivering management information for a selected few power users, required data to be extracted from Enterprise Resource Planning (ERP) systems and loaded into the data warehouse after going through lengthy extract, transform and load processes. This process was followed by the writing of custom or parameter driven reports to answer specific questions. This approach limited the user in the use of the information in terms of data discovery and analysis aspects, as the next question from the user would require another custom report to be developed. Traditionally, most BI implementations were controlled by IT professionals who were responsible for the full spectrum of activities, from requirements gathering, to presenting the data in an understandable way to the different stakeholders.

[1] **Forrester Research** is an independent technology and market research company that provides advice on existing and potential impact of technology, to its clients and the public

In newer BI initiatives, the focus has shifted to automating the procedures and processes and empowering the users to search and analyse the data through the use of different tools.

This approach allows the users to gain a greater understanding of the data and can make informed decisions on an unlimited number of queries with much less dependency on the IT professionals. [4]

Recent developments in both hardware and software have been enablers in delivering real-time, search based, interactive and visually rich BI to various types of stakeholders and at varying levels within an organisation.

In-memory computing and in-memory databases have played a major role, as enablers to delivering fast, actionable and relevant BI. The Gartner[2] report on 'Business Intelligence and Analytics platforms' [6] indicates that data discovery tools have become part of the mainstream BI architecture, facilitated by in-memory and columnar database technologies. These technologies are responsive to the increasing demand for decentralised BI and user empowerment through BI and analytics [4].

2.1. TRADITIONAL BI AND ITS ROLE

The original vision for BI implementations was to improve organisational management, by providing quality data which is understandable and easily accessible to various stakeholders within an organisation. The underlying ideology is that of achieving 'a single version of the truth'.

Initially organisations did not have any enterprise wide strategy or governance models with respect to BI. This gap resulted in departments implementing their own solutions to meet their immediate requirements. The outcome of this fragmented approach was that there were many versions of the same information being reported within an organisation. There were numerous disadvantages to this approach which included excessive costs in terms of decentralised investments in technology, information duplication and redundancies and the lack of stakeholder confidence in the integrity of the data. The time and effort spent in trying to reconcile the information and fix errors required skilled IT professionals. Further, the lack of common business rules and definitions being applied across the organisation, resulted in information 'chaos'. [7]

To address these issues, the notion of creating a single version of the truth arose. A high level BI architecture to support this requirement is represented in Figure 1.

[2] Gartner is an information technology research and advisory company providing technology related insight.

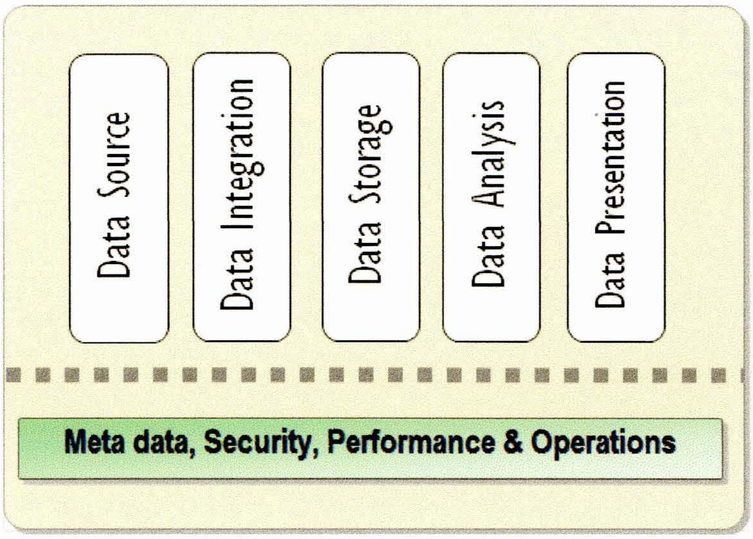


Figure 1 : Conceptual system architecture [8]

Within each of the 5 pillars of the conceptual architecture are several further components that may be required, for an integrated BI implementation. As shown below in Figure 2, the BI architecture can be extensive, with numerous processes and steps required within each of the pillars of the conceptual BI architecture, for a fully integrated solution. An implementation that covers the 5 pillars of the conceptual architecture is required to meet the requirements of data consistency, integration, completeness, accessibility, and cost of ownership, security and common metadata.

BI Conceptual Architecture

- **Data Source** – the system from which the data originates or the system that feeds data into the data warehouse. The data source for a BI implementation can be from heterogeneous sources e.g. ERP system, excel, text files, third party sources such as surveys and census data and even unstructured data. Data from the numerous sources form the data source layer.

- **Data integration** – refers to the data assets, processes, methodologies, tools and philosophies of the organisation by which fragmented data in multiple disparate systems is integrated to support business goals.

  An important step in the data integration is data cleansing or data preparation which involves detecting and removing of errors an inconsistencies [9]. The data collected from multiple sources can have incomplete information, misspelt, and improperly formatted, duplicated or invalid data. Data profiling is a step in the data cleansing process which collect statistics and information about the data being integrated for the data warehouse. This provides information about the quality of data being extracted for the data warehouse.

The ETL process can used to highlight dependency violations, rule violations and duplication in the data by storing the erroneous records in error tables. The data cleansing process is essential in data warehousing as it helps maintain the integrity of the data.

- **Data Storage** – refers to the database layers/schemas where the detailed and summarised data is stored for retrieval. These layer are the :-
  - Staging layer
  - Operational data store layer
  - Data warehousing layer

The transformed and cleansed data is stored in the data storage layers. The staging layers is where the source data is stored with minimal transformations and cleansing. Having the data in a common place allows for easier integration and processing of the data. The operational data store is where the more transactional data is stored after being through the extract, transform and load process where the data is cleansed and rules are applied so that the data can be used for analysis. The summarised data is stored in the data warehousing layer in subject specific data marts. Depending on the organisational needs one or all of these layers may be present in the data warehouse.

The data warehouse is defined as the storage architecture that holds the source systems data which is specifically structured for fast and easy query and analysis, for abstracted subject areas. It holds time-variant version of the same record and is either in detail or summary. The structures that support fast query and analysis are data marts which are designed using multidimensional design. The data mart is usually focused on a specific subject area within an organisation, for example, procurement. The multidimensional design has central fact table and one or more dimensions often referred to as a star schema or cube (See appendix A for the procurement star schema). The fact table has the facts or measures for the specific subject area which can be analysed across one or more dimensions. Dimensions are categories of attributes organised for ease of data visualisation. Dimension may have one or more hierarchies which allows for drilling up or down the cube, which allows for analysing data at different levels (i.e. summarised of in detail). Multiple, integrated data marts are referred to as an Integrated Data Warehouse.

- **Data Analysis** – refers to the process of changing information into knowledge for an organisation so that informed decisions can be made. The data can be used in predictive modelling, data mining and forecasting and presented to the user via the reporting tool.

- **Data Presentation** – refers to the process of exposing the data in the data warehouse to users through reports, graphs, dashboards, alerts, ad-hoc querying, emailed reports or alerts that notifies users of exceptions. The metadata layers holds the information about the data the data warehouse. For example, the logical model of the data mart and the associated business rules. A reporting tools is used that to make the data easily accessible to the user and presented in a form that the user can understand.

These five pillars must be further supported by enterprise data governance strategies, metadata management strategies, data quality and information security strategies.

Processes, procedures and standards need to be implemented in order for a BI implementation to be successful. Aside from the technology, the organisational culture needs to change to be more receptive to making informed decisions supplied by a single data source.

A complete implementation requires huge investment in infrastructure and personnel and the time to delivery can range from months to several years. Although the user is shielded from much of the complexity, many BI initiatives have not been successful. Failure is due to long development times and the high level of dependency on IT departments. Figure 3 shows that the value of BI is only realised after significant investment and protracted development time. Data latency and time to load pre-calculated aggregated OLAP cubes, are further examples of the shortcomings of the traditional BI approach in trying to meet the increasing demand for data discovery and real time analytics.

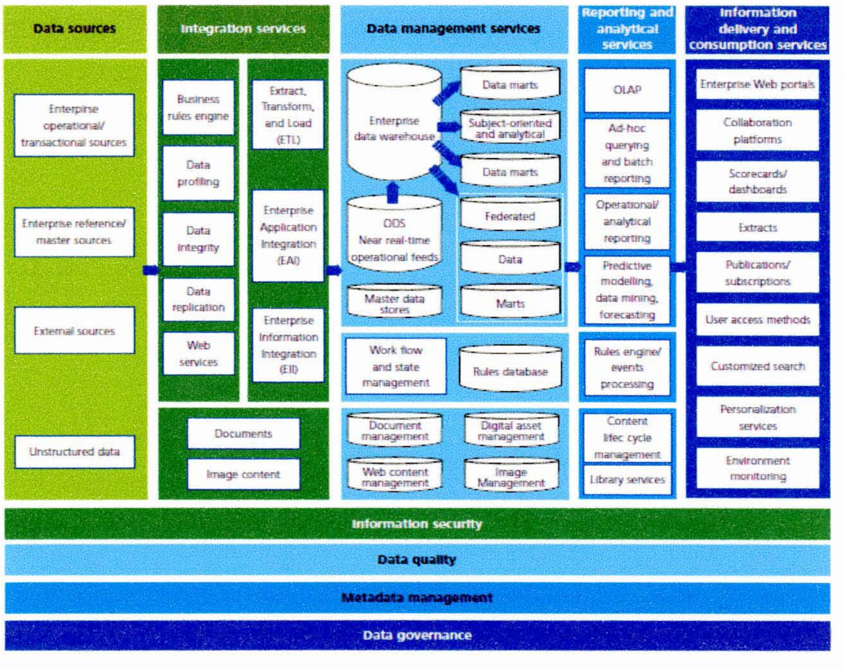


Figure 2 : Comprehensive Business Intelligence Architecture [8]

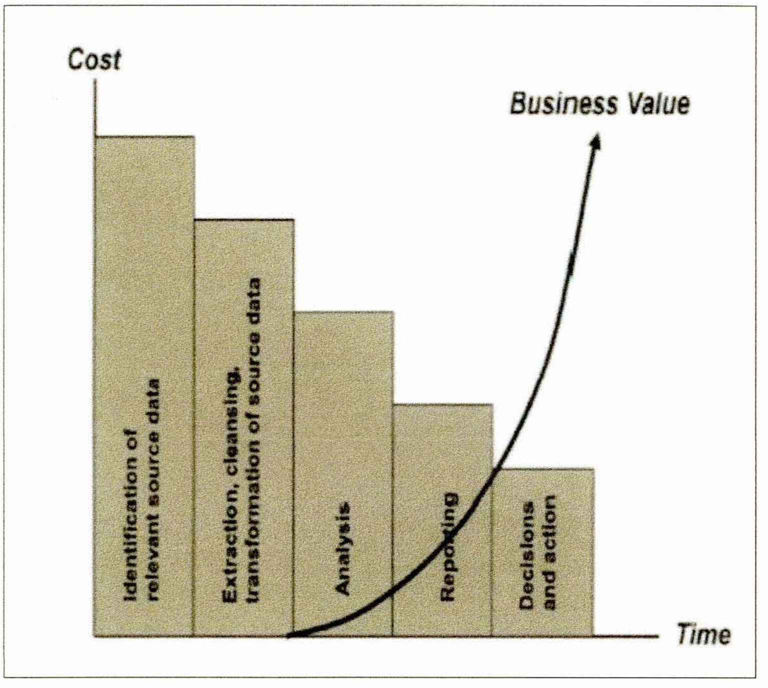


Figure 3 : BI value chain against cost and time (Source: Adopted from Platon A/S (2009)) [10]

2.2. WITS' BI Infrastructure

The University of the Witwatersrand (Wits) has been an early adopter of Business Intelligence. Although BI was not initially part of the high level strategy, there was sufficient executive support to build the BI infrastructure that supports the reporting and analytic requirements of a diverse user base, through an integrated data warehouse.

The first online analytic processing system (OLAP) called the Executive Information System (EIS), was built in 1995. This system was aimed primarily at the senior executive team of the University, to track patterns of student enrolment, human resources capacity and high level financial metrics. The models built were highly summarised, with no drill down facility to the underlying detailed information.

Frustration with access to the summary level information only signalled the need for reporting at a lower operational level. A financial data mart was built, with data presented to users via the Business Objects reporting tool. This enhancement allowed for seamless detailed and summarised reporting. However, there was no integration between the three core subject areas (students, human resources, finance) resulting in BIS dealing with regular data integrity issues. Reasons for data integrity issues included that data was drawn from other bespoke systems or directly from transactional systems. Additionally, departments were maintaining their own records in Excel, which often did not reconcile back to the operational system.

The opportunity to develop an integrated BI infrastructure arose when the university decided to move to an ERP system. The Enterprise Data Warehouse (EDW) was built in 2004 for the three core business areas. The Wits BI architecture is represented in Figure 4. The BI implementation at Wits has matured over the years and various other business areas have been included for reporting and analysis. The Wits EDW has also evolved over time to include further functionality, such as dashboards for each of the strategic priorities at Wits, alerts, pattern checking and error reporting.

One of the key pillars for BI implementation success is executive support [11]. The challenge is for BI centres to quantify the benefits of an integrated BI implementation, for an educational institution with limited resources. These benefits are not always tangible or easily quantifiable.

Due to executive level support, the Wits EDW is currently a stable environment, which satisfies the reporting requirements of various types of stakeholders (approximately 700 users) across the university, as well as various external stakeholders. The EDW facilitates timely access to integrated data and audited data for all the core business areas.

As the chronology of changes indicates (See Table 1), the BI implementation is not a static, once off implementation. It has to evolve with changing business requirements and respond to the environmental changes.

The current challenges to be addressed at Wits are:

- The need for real-time information in certain areas and processes;
- More user empowerment in terms of data discovery (ad-hoc querying) and data analysis;
- The need to consolidate the toolsets used for the population of the warehouse and the toolset used for reporting. The consolidation is necessary in order to streamline resource requirements (technology and human) which are costly.

It is in response to these challenges that alternative approaches to delivering BI need to be considered at Wits.

Chronology of WITS BI Implementation							
Period	Systems Supported	ETL Tool	OLAP/ ROLAP Technology	Presentation Layer	Stakeholders	Level	Architecture
1995 - 1997	Student / GL Finance / HR	Custom written	Essbase	Execu- view	Senior Executive	Summary	Multiple sources into standalone data marts
	Student / HR		Essbase	Execu- view	Senior Executive	Summary	
1997- 2003	GL Finance	Custom written	SQL server	Business Objects	Middle management / Administrators	Detail	
2004- 2012	Students / HR		Oracle Relational database	[illegible]	Senior Executive / Middle management	Summary and Detail	Integrated Data Warehouse
	Students	OWB		Discoverer	Administrators	Detail	
2004- 2013	GL Finance / Project / Procurement	OWB	Oracle Relational database	Business Objects	All levels of stakeholders	Summary and Detail	
2010	Student Enrolment Tracking (near real time)	OWB	Oracle Relational database	[illegible]	Middle management / Administrators	Summary and Detail	
2011	Short courses	OWB	Oracle Relational database	Discoverer	Middle management / Administrators	Summary and Detail	
2012	Investment	[illegible]	Oracle Relational database	[illegible]	Middle management / Administrators	Summary and Detail	
2013	Students / HR	[illegible]	Oracle Relational database	[illegible]	Senior Executive / Middle management	Summary and Detail	
Nov-13	GL Finance / Project / Procurement	[illegible]	Oracle Relational database	Business Objects	All levels of stakeholders	Summary and Detail	
2014	Students at Risk	[illegible]	Oracle Relational database	[illegible]	Middle management / Administrators	Summary and Detail	
2014	Procurement/Students at Risk	Not required	[illegible]	[illegible]	Middle management / Administrators	Summary and Detail	
2014	Real Time Enrolment Tracking	[illegible]	Oracle Relational database	[illegible]	Middle management / Administrators	Summary and Detail	

Table 1: Chronology of Wits BI Implementation

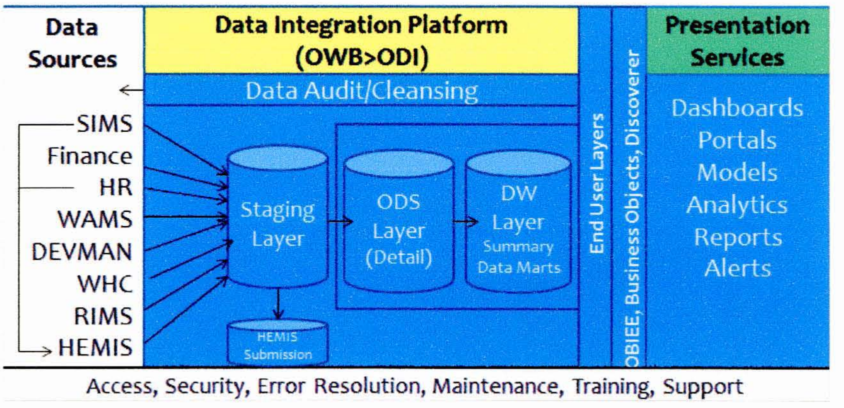


Figure 4 : WITS BI Architecture [12]

3. IN-MEMORY COMPUTING

In-memory computing (IMC), in-memory databases and in-memory data grids, in-memory appliances and in-memory analytics are the latest buzz words in the BI industry. In the Gartner IMC Hype Cycle report, it is noted that IMC "opens unprecedented and unexplored opportunities for business innovation" [13]. Refer to Appendix E for details of the hype cycle.

The IMC is defined as "a computing style in which the primary locus of the data for applications is the central memory of the computing environment (on single or networked computers) running these applications" [13]. There are several advances in hardware and software technologies that have enabled the IMC approach, in their application design. These technologies are at various stages in the Gartner Hype cycle shown in Figure 5.

The two technologies that will be considered for this research report are analytic in-memory Database Management Systems (DBMS) and in-memory analytics. Analytic in-memory DBMS is currently at the 'peak of inflated expectations' and in-memory analytics are going through the 'trough of disillusionment' in the hype cycle. However, it is expected that both these technologies will reach the plateau of productivity within the next 2 to 5 years [13].

The number of organisations that are considering IMC as an affordable technology is increasing, due to the decreasing costs of direct random access memory (DRAM) and non-volatile flash memory (NAND). Another contributing factor is the availability of multi-core 64-bit microprocessors that can directly access the large main memories of a computer [13]. The decreasing cost has made it possible for small to medium businesses to adopt IMC technologies, such as desktop in-memory analytic tools for data visualisation.

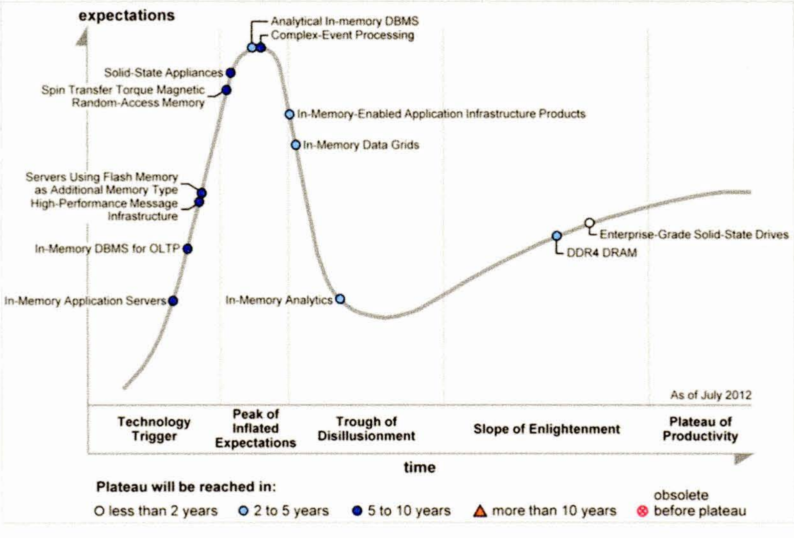


Figure 5 : Hype Cycle for In-Memory Computing Source: Gartner (July 2012) [13]

3.1. In-Memory Databases

An In-Memory Database Management system (IM DBMS) is a database where the entire database structure is stored in memory, which can be accessed without the use of input and output instructions. The high speed of in-memory can be leveraged in analytic IM DBMS. The technologies that use analytic IMDBMS are in-memory column store databases and in-memory massively parallel processing row based DBMSs.

The adoption of IM DBMSs has been due to the low latency data loading capabilities, which allow for calculations to be done in real-time and on the lowest level of detail of the data. The IM DBMS allows for rapid changing business requirements (where business cannot wait for the long development time required for the traditional BI) as well as the high speed of querying and calculation.

In-memory databases have been used previously as embedded systems, specifically to meet resource and performance requirements. In-memory databases eliminate the disk input and output (I/O) and are less complex than traditional DBMS that are fully deployed in RAM [14].

The 3 differences between IMDBMS and traditional DBMS in memory are:

- Caching

  IMDBs do not have caching therefore they eliminate the complexity and performance overhead of cache lookup and cache synchronisation
- Data transfer overhead

  IMDBs have little or no data transfer. The application accessing the data receives a pointer to the data, which enables the application to work with the data directly. The design is simpler and more reliable, as removing multiple copies of data reduces memory consumption and eliminates the multiple data transfers, thus streamlining the processing.

  Figure 6 below, shows the data flows in traditional DBMS when modifying a piece of data and then writing it back to the database. Additional data transfers and copies are made for the transaction logging [15].

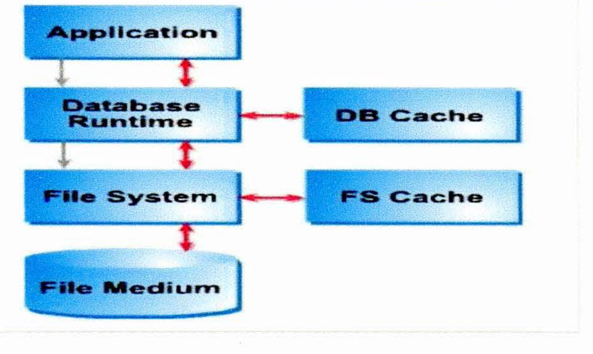


Figure 6 : Data flow in a traditional DBMS.
(Red arrows represent data transfer. Gray arrows represent message path.) [15]

- Transaction processing

  In-memory databases maintain a 'before' and 'after' image which eliminates the need for transaction log files being maintained, which is a complex, memory intensive activity of traditional DBMSs [15].

3.2. In-Memory Analytics

In-memory analytics is an alternative way of delivering BI, as opposed to the traditional BI approach.

The factors driving the adoption of in-memory analytics are [16]:-

- Speed
  The detailed data is loaded into memory, where the speed of calculations and querying against larger set of data is very fast. This approach does not require any pre-calculated OLAP cubes or aggregates, which not only makes the loading of the data into the in-memory database very fast, but also facilitates very fast ad-hoc querying of the data.
- Costs
  The adoption of in-memory analytics has been facilitated by the decrease in memory prices and organisations moving to 64 bit computing.
- Data Volumes
  Structured transactional data in organisations, as well as unstructured data, such as emails, graphics, and videos, is growing rapidly. For organisations that need to analyse large volumes data from legacy systems, in-memory analytics offers the opportunity to rapidly develop a database to perform the required analysis, without having to build complex data warehouse structures to facilitate the analysis.
- Real-time analytics
  In-memory databases have the ability to bulk load large volumes of data, in near real-time, which allows organisations to analyse and report on their data in near real-time. [16] Due to globalisation, organisations require their high volume, fast changing data, to be available 24/7. With in-memory analytics and in-memory databases organisations can meet the requirement.

The benefits of in-memory analytics are [16]:

- Performance Improvement
  Querying data or interacting with data with in-memory databases, is in the order of 3 times faster than querying data from disk;

- Cost Effective Alternative to BI

  In-memory databases are most cost effective for organisations that don't have the skilled IT staff and the financial resources required for large data warehouse implementations;
- Data discovery

  In-memory analytics offers self-service access to information, with rapid query execution, which allows users to develop quick insights into the business without having to wait for IT departments.

3.3. Merits and Drawbacks of Traditional BI and In-Memory BI

The technology to support BI implementations should provide the fundamental requirements of speed, flexibility, scalability and high availability, at minimum cost [16].

While any one solution will not provide all the requirements, each has its merits and drawbacks. Organisations will need to evaluate the trade-offs in order to decide which approach best suits their requirements.

Both in-memory computing and actionable analytics have been listed in the Gartner Top 10 Strategic technology trends for 2013 [3]. As organisations begin to consider these technologies, consideration needs to be given to several other factors, such as costs, skill requirements, changes to business process in order to support the technology, etc.

3.3.1. Traditional BI

3.3.1.1. Merits of using the traditional BI approach include [11], [16] *:*

- Integrated data from multiple sources
- Consistent data across organisational systems
- Cleansed current and historical data
- Data analysed across dimensions and hierarchies
- Fast delivery of analysis
- Data integrity and data security
- Metadata management and data lineage
- Monitoring and predictive capabilities.
-

3.3.1.2. Drawbacks of using the traditional BI approach include [11], [16] *:*

- Heavy reliance on IT
- Complex disk-based performance layers (e.g. pre-calculated OLAP cubes with fixed measures and dimensions ,relational-based aggregates)
- Limited data scalability, analytic scope and drill through capability
- Inflexible
- Difficult to build data marts that will satisfy requirements of all the users
- Average implementation time of 17 months

- Reporting requirements have a large impact on I/O time and network resources when analysing large volumes of data
- ETL processes may take several hours to complete which involves index creation and sorting
- Limited user adoption
- Data latency

3.3.2. In-Memory BI

3.3.2.1. Merits of using the in-memory approach include [16], [17]:

- High performance – faster because of the minimum disk I/O which improves query performance by a factor of 10
- Flexible analysis / seamless navigation
- Speed and ease of in-memory analysis development
- User transparency
- Near- real time reporting
- Ease of use / less user training
- Affordability
- Plug in-and out system architecture / disposable analytics
  No recalculation or aggregation is required, no indexes are required

3.3.2.2. Drawbacks of using the in-memory approach include :

- Without 64 bit technology there is a significant limit to the amount of data held in memory
- Data in memory is not the data in the data store – "always" out of date i.e. data must be moved from the data store to main memory therefore the most recent data is not in-memory until the main memory is refreshed
- Fragmented silos of data which occur when organisations have multiple departmental implementations of in-memory databases and the data is not refreshed simultaneously
- Loading data into memory may be time consuming

3.3.2.3. Impact on current business processes in delivering real time analytics

Although IMC has been listed in the Top 10 strategic technologies by Gartner [3], the impact on the organisation must be considered before adoption. The Gartner projection is that by 2016, at least 35% of midsized and large organisations will adopt IMC [18].

The organisational requirements have to be assessed before adoption. The current adoption of IMC has mainly been in organisations with an initiative aimed at gaining or retaining a competitive advantage, rather than by organisations investigating ways to reduce IT costs [18].

Some of the challenges that organisations must consider are:

- Changes to the organisational application architecture and design s;
- Stakeholders/IT personnel not familiar with IMC technologies may not have confidence in storing an organisational asset in-memory, which could be lost if there are system crashes or power outages;
- IMC technology is still in its infancy stage and as a result there are no commonly agreed standards amongst vendors;
- The best practices for IMC have not yet been finalised within BI ecosystem;
- Skills required to develop, support and maintain the IMC technology are currently hard to find and therefore expensive;
- Designing in-memory based applications are complex and therefore organisations will face challenges in terms of consistency, integrity and debugging of data;
- Investment in additional hardware, software, IT reskilling and user training;
- Multiple data silos being created that will need to be carefully managed in order to uphold the single version of the truth philosophy;
- Ability to audit the integrity of in-memory data.

3.3.3. Co-Existence of Traditional BI and In-Memory BI

The traditional approach to BI and the newer in-memory approach, both have their merits and drawbacks. The question to be addressed is whether the in-memory approach is an incremental evolutionary change to current BI architecture, or whether it is a game changing revolutionary approach.

Research conducted by Winter, Bischoff and Wortmann [19] indicate that the two technologies complement each other, rather than in-memory approach signalling the end of the traditional approach to BI.

Depending on the business requirements and/or the data volume, in-memory technology can be strategically adopted to meet a particular need.

The 4 patterns identified are [19]:-

- Low/moderate data volume, no need for integration;
- Huge data volume, no need for integration;
- Low/moderate data volume, need for integration;
- Huge data volume, need for integration.

The focus of traditional warehouses is to create a consistent snapshot of the data, where data from heterogeneous sources is integrated, after going through a process of extraction, cleansing and harmonising. Figure 7 indicates that in-memory technology can be adopted in an integrated data warehouse, for analysis and reporting, where several pre-calculated fixed data marts are replaced by a flexible in-memory database. Alternatively, in-memory databases could be used in an OLTP environment, where there is demand for context aware analysis. The results can be fed back to the operational environment to inform the next business transaction.

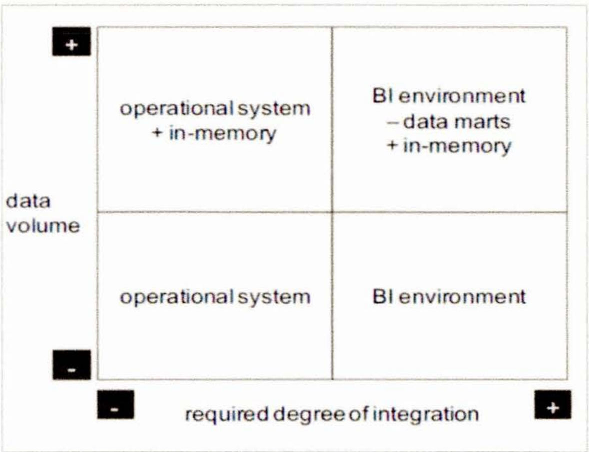


Figure 7 : Four system architecture stereotypes [19]

BI solutions have been inaccessible for small and middle sized (SMEs) organisations due to high costs, high investments in hardware, complexity for the users, irrelevant functionality and low flexibility and niche business requirements [20]. The tool set used for larger organisations may not be relevant for smaller businesses. SMEs require lightweight, cheap, flexible, simple and efficient solutions [20]. With in-memory databases and in-memory analytics, even small and middle sized organisations can leverage the benefit of BI. Increasing data volume necessitates the need to perform data analysis, even for smaller organisations.

3.4. In-Memory Architectures

Currently there are several approaches in the architecture used to improve performance of in-memory databases and in-memory analytics. Organisations choose a particular technology based on the organisational requirements in terms of cost and functionality.

3.4.1. Columnar Databases

For analytic applications, the ad-hoc queries or reports query the database to analyse selected attributes of a larger number of records. When using a row-orientated database the entire row must be read to access the required columns resulting in more data being read than what is required for the query. In columnar databases each column is stored separately and only the requested column values are retrieved reducing the I/O. This approach enables rapid access to the data and significantly improves the query performance.

The benefits of column orientated databases include [34]:

- Engineered for analytic performance
  Indexes are used to store data rather than a reference pointing to data as in row-orientated databases so that only required columns are fetched. This I/O can be done in parallel as the columns can be stored on different disks.

- Rapid joins and aggregation

  Increase performance when doing aggregations as the data does not have to be stored together which allows for parallel access and aggregation.

- Smaller storage footprint
  The data is stored within the index thereby eliminating the additional storage overhead of different types of indexes in row-orientated databases.

- Suitability for compression
  Columnar databases can reduce the storage requirements and maintain performance through the use of different types of compression techniques. For example, compression of actual data values where certain columns of data have repeated values. These values can be replaced by tokens, one per value and the token length shorter than the actual value in the column. This type of compression is better suited to columnar databases than row orientated database.

- Rapid data loading
  The columns in a column orientated database can be stored separately which allows for parallel loading of the columns using multiple threads which improves performance.

The columnar approach is used by several leading vendors of in-memory BI platforms e.g. SAP HANA (High performance analytics).

3.4.2. Associative Approach

The associative approach is discussed in Section 4. This approach is unique in that it used associations between entities to return results rather than using the query based approach to return the required data. This approach does not have the standard database management system. Requested data is returned based on the associations between certain tables. QlikView in-memory database uses the associative approach for in-memory BI.

3.4.3. Software and Hardware Engineered Solutions

Leading vendor Oracle has introduced BI solution (Exalytics) that is engineered for high performance, by using the combination of hardware and software designed specifically to overcome the bottlenecks in data processing. The availability of high-speed networking, improved processor throughput, increased memory capacity and persistence memory storage, like Flash, are used in combination with optimised software to deliver the 'speed of thought' analytics. The software optimisation includes the use of query optimisation, parallel computing and different types of storage management.

Oracle's Exalytics solution in Figure 8 aims to overcome the challenges companies are facing in delivering fast actionable analytics [21]. Survey results indicate that one of the leading challenges for organisations is accessing and consolidating the right information from disparate data sources in order to present a single version of the truth for analysis. This is due to the fact that organisations have to integrate data from a number of different types of hardware, software, networking and data storages in order to deliver a complete analytic solution. The Exalytics architecture offers a single optimised server with a single set of software that reduces the complexity of integrating data, duplication of data and resources as well as maintenance and administrative costs. [21]

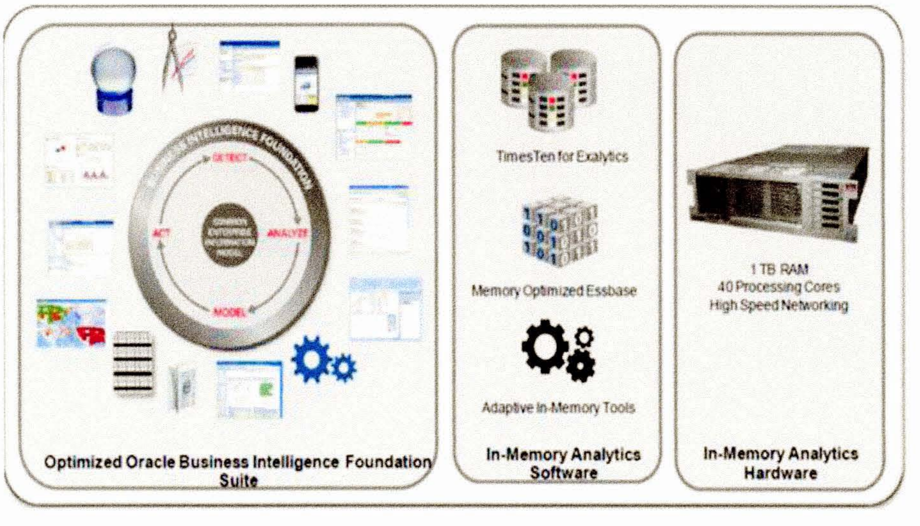


Figure 8 : Exalytics Hardware and Software Architecture

The costs of the Exalytics engineered hardware and software solution may not render it as a viable option for the small to medium organisations however the Exalytics solution is considered a better option when organisations have large data volumes.

Another leading vendor in the in-memory analytics space is SAP with SAP HANA Enterprise 1.0. Similar to the Exalytics it is an in-memory computing appliance that combines the SAP database with pre-tuned server, storage and networking hardware as displayed in Figure 16 [22]. The appliance supports both analytics processing as well as transactional processing. The data replication into the in-memory database allows for near real time access from both analytic and transactional applications.

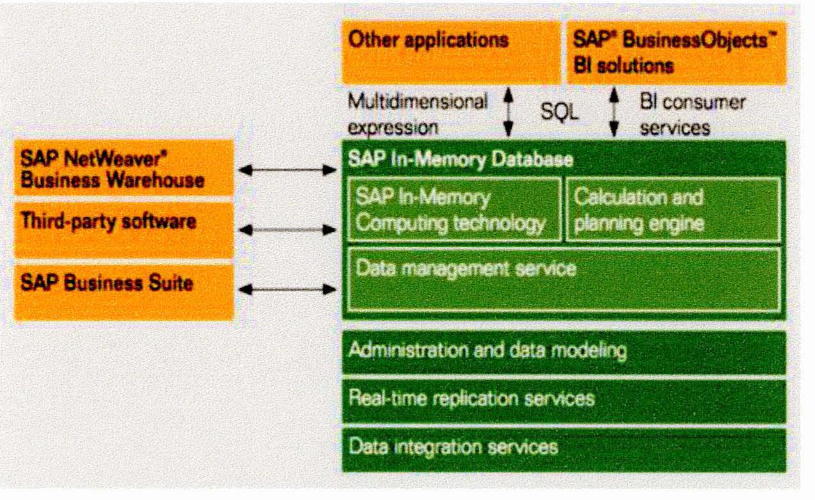


Figure 9 - SAP HANA Architectural Design Source: SAP 2011

Large volumes of data can be integrated from disparate sources into in-memory database which a massively parallel processing data store that combines row-based, column-based and object-based storage techniques [22].

The SAP solution allows for different vendors to be used for the components that make up the appliance thereby not enforcing vendor lock in. The costs of the combination of hardware and software requirements may not feasible for small to medium sized organisations however it provides a flexible and economical analytic option for the larger organisations.

4. RESEARCH METHODOLOGY

The research undertaken was to rebuild the Wits Procurement dashboards using in-memory technology, in order to evaluate the appropriateness of using an associative in-memory dashboard technology within the Wits context, which has a mature enterprise data warehouse and Business Intelligence infrastructure. The current procurement dashboards and reports were built for both managerial and operational staff within the procurement business function, using the traditional BI approach.

The in-memory database technology used is QlikView by Qliktech; which is in the leaders' quadrant of Gartner Magic Quadrant for Business Intelligence and Analytic Platforms (See Figure 11). QlikView is an in-memory database, which uses an associative technology where every data point in every field is associated with every other data point, through the one key field that joins the tables together. Based on the user selection, all other fields will be filtered and aggregations will be done in real-time without any queries being initiated [23]. QlikView is an integrated BI platform that offers features that include dashboards, interactive visualization, mobile BI, search-based BI, scorecards, ad hoc query, Microsoft Office integration, OLAP, and development tools. The high level QlikView architecture is represented in Figure 10.

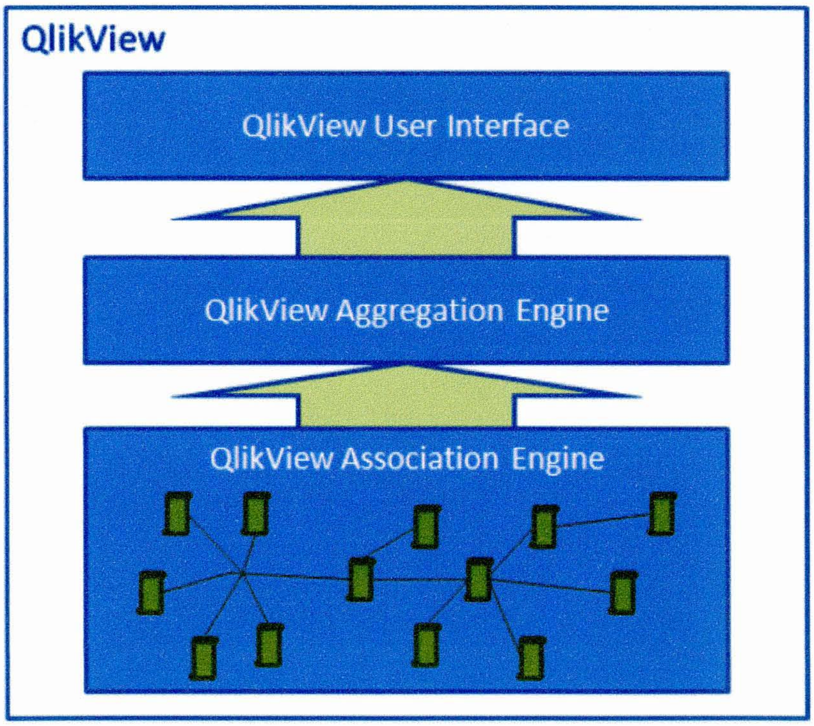


Figure 10 : QlikView Architecture [17]

The selection of this tool for the development of an in-memory prototype was based on the functionality available, free access to the community edition of the QlikView software and the suitability of the software for delivering in-memory departmental level BI.

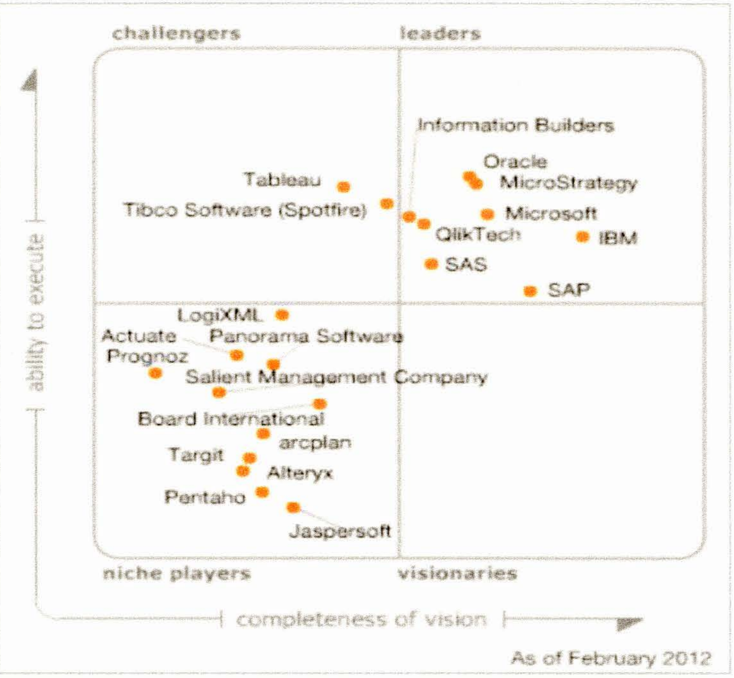


Figure 11: Magic Quadrant for Business Intelligence Platforms Source Gartner (2012) [6]

4.1. Research Purpose

This research was conducted to determine the role that in-memory databases and in-memory analytics may play in an environment that has an established and mature traditional BI infrastructure.

The questions investigated include:

- Are these two approaches to BI mutually exclusive or is there sufficient justification for them to co-exist within the same BI ecosystem?
- What are the business requirements that would require both technologies to be used?
- Does the in-memory associative approach shorten the development life cycle of a subject specific dashboard and deliver improved value, while at the same time being a more cost effective and faster approach than the traditional methodology used for BI?
- What are impacts and requirements of a real-time or near real-time offering on the business processes and procedures of an organisation?

4.2. Research Approach

An in-memory database was built on the subject specific area of procurement, using Wits procurement data as the data source.

An evaluation was done comparing the traditional data warehouse approach to BI and the new in-memory approach, using the following criteria:-

- Speed of development
- Resources
- Skills requirements
- Volume of data
- Flexibility in accommodating business changes
- Flexibility of analysis
- Real-time capability
- Integration level
- Information security
- Reporting capabilities
- Data integrity
- Durability/Persistence of Data

Users of the existing traditional Wits procurement dashboards were exposed to the dashboards developed in QlikView, using the same source data. A user survey was conducted to understand user perceptions regarding the two technologies based on a subset of the criteria listed above. Results of the survey were analysed.

4.3. Research Strategy

Two approaches were used to develop the in-memory database and dashboards for the Wits procurement users:

- Use the procurement data in the data warehouse as the data source
- Use the ERP system as the data source

These approaches were used to determine the variance in time-to-delivery between:

1. extracting data from a data warehouse (when the source data has already been cleansed, enriched and aggregated etc.); and
2. extracting data directly from the operational transactional system

The second approach was also used to implement the near real-time scenario using QlikView to assess the challenges encountered with delivering near real-time data. The entity relation diagram (ERD) of the WITS Procurement data mart from the data warehouse, is represented in Appendix A.

The diagram showing QlikView's representation of relationships between the tables that were sourced directly from the ERP, is shown in Appendix B.

4.4. Data Collection Method

Users were given access to a web version of the QlikView tool. A training session was conducted with the users and they were given a month to use and evaluate QlikView. Users were requested to participate in an online survey, to collect data on user perceptions of the QlikView tool in comparison with the traditional BI, Oracle Business Intelligence Enterprise Edition (OBIEE) tool.

5. RESULTS/FINDINGS

5.1. Speed of development

5.1.1. *Learning QlikView*

The development of the procurement in-memory database and dashboards was done using QlikView Community edition software. QlikView is an intuitive development tool that was easy to learn using the online documentation and help available. Having prior knowledge of data warehousing was beneficial in terms of the time required to learn the tool and the QlikView architecture.

One of the reasons that the adoption rate of in-memory technology is increasing is due to the quick time-to-delivery development cycle. In-memory analytics makes it simpler to build applications since the need to build complex performance layers is eliminated [16]. That is, a faster implementation of BI is facilitated, as there is no need for data indexing, pre-calculated measures or aggregate tables. This approach is particularly suitable when departments within an organisation require very specific questions answered quickly from the transactional data, as IT departments cannot build specialised data marts to meet all user requirements [16].

Learning the tool and developing the first prototype was achieved within 1 week. This validates the assertion that application development times can be reduced using in-memory technology when comparing to the traditional approach. The time taken for the design of the Procurement data mart for the enterprise data warehouse plus the time for ETL routines to be written was approximately three months. Further time was required to build customised reports for various stakeholders. The report building is an on-going process.

Figure 12 displays results from a QlikView customer experience survey [24]

The results are based on the QlikView customer responses to the survey question:

> *How long did it take from the initial purchase of the software to complete implementation and achieve payback with QlikView?*

The time taken to prototype the WITS procurement dashboards in QlikView are in line with the survey results shown in Figure 12. Results indicate that 77% of the respondents had an implementation time for 3 months or less and 30% of the respondents achieved payback in three months on less. The results from the survey confirm that quick implementation time is a key success factor [24].

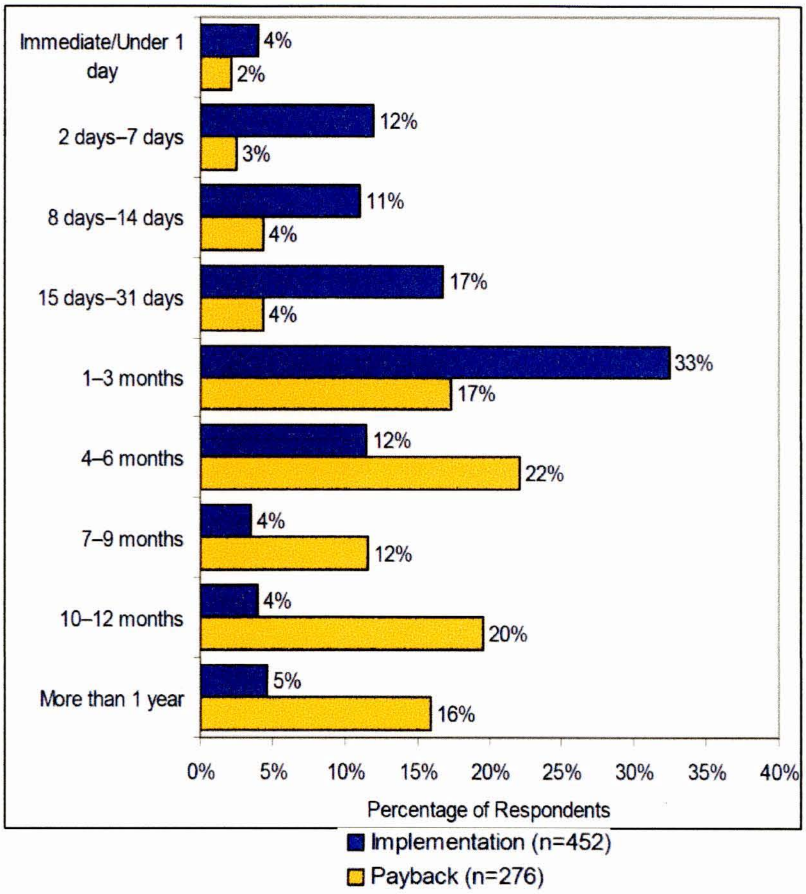


Figure 12 : Time to Value of QlikView. Source: IDC survey of QlikView customer base, Jan–Mar 2009 [24]

5.1.2. *Traditional Date warehouse as a data source*

By placing the in-memory database within the data warehouse infrastructure to facilitate the speed of analysis, allows one to leverage the investments already made in BI infrastructure and the data that has already been validated, cleansed and enriched.

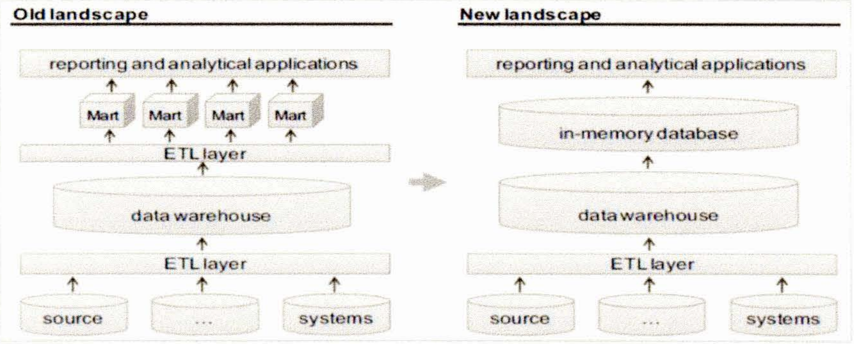


Figure 13: Enabling in-memory data warehousing [19].

The architecture suggested by Winter et. al. [19] in Figure 13 indicates that query response times can be decreased by implementing an in-memory database within a traditional BI infrastructure. This architecture eliminates the need for building a complex performance layer and facilitates flexible, as well as detailed querying.

The Wits Procurement dashboard, using QlikView, was built using the data in the data warehouse as a source. The development of this scenario took a developer, working on a part time basis, two weeks to complete. The advantage of this approach is that the data has been de-normalised into dimensions for reporting, enriched with descriptions and audited against the source. It gives the users the required flexibility for data discovery and fast response times against a validated data source.

In this scenario the in-memory database plays a complementary role with respect to the traditional data warehouse.

5.1.3. *ERP as Data Source*

Developing the Wits Procurement dashboards in QlikView, using the ERP as a source took a developer working part time (3 hours daily), 6 weeks to develop. The development time for this scenario was longer due to the time spent on data preparation.

In most BI implementations the major portion of the time is spent on data preparation. Firstly, finding the correct data from the myriad of ERP source tables is time consuming and requires specialised IT skills. The data has to be restructured and cleansed before any analysis can be performed on the data. The value a BI implementation can offer is be negated as soon as the quality of the data is questioned by users. Therefore, the vital step of data preparation must be done thoroughly.

When using the ERP as a direct source for data, the data has to be enriched in order to make analysis effective. That is, adding descriptions to key fields, defining the hierarchies required for reporting and defining the time dimension, are steps that cannot be skipped.

The time taken to identify and prepare data for the traditional BI implementation or an in-memory implementation is similar which is estimated to be between 60-80% of the project time [25].

The advantage the in-memory approach has, is that once the data has been identified it can be loaded frequently as required by business as the data does not need to be put through the 2 ETL processes as required by the traditional BI architecture. Further, no aggregations or complex OLAP cubes need to be re-calculated.

5.2. Resources

The investment required for infrastructure in traditional BI implementations can be very costly, depending on the organisational requirements.

Organisations need to make investments in various types of technologies, from multiple vendors, in order to deliver an integrated BI solution. These technologies include databases (relational and/or multidimensional), ETL tools and different types of reporting and dashboard tools. The yearly licencing cost for proprietary software is dependant of the vendors licencing structure which can vary from 10% to 30% of the original software cost.

Figure 14 illustrates the difference in types of resource requirements in a traditional BI implementation and in a QlikView implementation. An integrated in-memory offering, such as QlikView, makes BI accessible to small and medium organisations as the total costs are lower.

However, the amount of data that can be analysed in an in-memory implementation is limited by the amount of addressable RAM available. As the data volumes increase, the amount of RAM required also increases. For very large data volumes it may not be possible to buy sufficient memory to support the requirements or it may be simply too expensive to acquire more RAM.

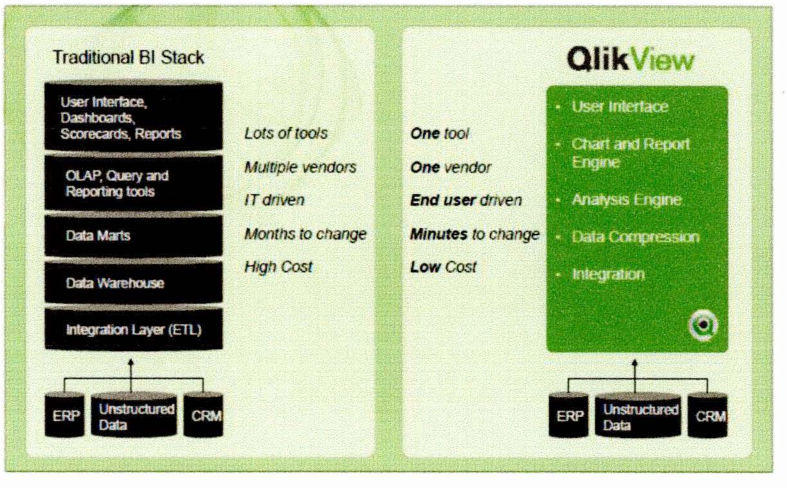


Figure 14: Providing as Integrated Approach (adapted from [26])

5.3. Skills requirements

The IT skills requirements will change with the adoption of in-memory technology.

The recommendation is three core skills should be considered when reskilling for in-memory computing [27]:

- The team should have development skills for designing in-memory applications;
- The team must have the required architectural skills for implementing in-memory applications – in-memory architecture is more complex than traditional architecture and will require a considerable amount of re-engineering to realise the full benefits;
- The team will need to develop the skills to evaluate the impact on IT operations because in-memory computing is different and will require operations staff to be re-skilled.

5.4. *Volume of Data*

The volume of data that can be handled by an in-memory database or traditional BI is dependent on the infrastructure in terms of hardware. Both environments can handle large volumes of data, however, in an in-memory environment as the volume of data increases or the number of users' increases, the amount of RAM required increases. This increase will affect hardware costs as disk-based memory is cheaper than RAM.

5.5. Flexibility in accommodating business changes

When business requirements change at a rapid rate, in-memory computing provides a more flexible approach. The number of steps required to get the data into the in-memory database are fewer. That is, there are no lengthy ETL processes or aggregation steps required. In-memory databases offer greater flexibility if the database is only a subset of the data where there are no integration considerations to be taken into account.

In-memory databases can be used as an effective agile BI tool to respond to the demands of agile business environment.

5.6. Flexibility of Analysis

Traditional BI provides limited and fixed scope of analysis based on the predefined dimensions for a particular data mart. The drill-through capabilities are also limited. In an in-memory database all the data is loaded into memory so that the entire dataset is available for fast query and analysis. The user has access to aggregated data as well as the lowest level of granular data.

The data in one query is intrinsically related to the data of another query when querying the data in memory, thereby providing a dynamic and flexible way of exploring data. This flexibility in data analysis reduces the business users' high dependency on specialised IT staff to answer specific questions.

5.7. Real-time capability

The in-memory databases offer a near real-time capability. That is, the large volumes of data can be loaded into memory in a short space of time. Since the data has to be extracted and loaded into memory it is not real-time data. However, the required data can be loaded frequently so that the available data is near real-time.

The data loaded into in-memory databases does not go through the lengthy ETL processes required to aggregate the data. Data can be loaded quickly into the in-memory database thereby providing near real-time data availability. However, the speed of loading can be impacted if the data volumes are too large.

5.8. Data Integration level

A fundamental approach of traditional BI is to integrate data from various heterogeneous data sources to provide an integrated and consistent platform for data analysis. Many current implementations of in-memory databases are aimed at departmental implementations, which result in fragmented silos of data. IT departments have spent years of effort in building data warehouses that eliminate decisions being made on individual silos of information. The small department in-memory databases have the potential to repeat history in terms of trying to find a single version of the truth. Due to the heterogeneous sources of data, the integration step will need to be performed prior to the data being loaded into the in-memory database [28]. That is, the vital integration step is not easily replaceable when using the in-memory approach.

5.9. Information security and data governance

The security concerns with in-memory databases relate to who has access to the data and where is the data stored. It is recommended that data be stored on a centralised server in order to avoid the risk of that data being stolen or leaving the organisation's premises [16]. The user access profiles must be managed the same way that access is managed on

transactional systems in order to adhere to the data security measures of an organisation [16].
Multiple implementations of data discovery tools like QlikView can result in disconnected information silos which result in organisations having to address governance, reusability and information sharing issues [18].

5.10. Reporting capabilities

Users of traditional BI systems are accustomed to reports being produced in a consistent formatted way. The reports are customised for user requirements and are re-useable. Users of BI have differing reporting requirements and there will still be some users that are not capable or empowered enough to 'discover data' for themselves to generate the required reports.
While in-memory data discovery tools like QlikView offer the user a great degree of flexibility in discovering data, it can lead to problems when users produce different sets of results, due to a difference in data selection.

5.11. Data integrity and data quality

To maintain the same level of data integrity and data quality, the same amount of effort has to be committed in either an in-memory or traditional BI environment.

The audit and control mechanisms used in traditional BI implementations will need to be adopted in in-memory implementations so that user confidence in the data is maintained.

- The administrative tasks on both traditional BI and in-memory BI implementations require the same level of organisational effort and resources. The Ventana Research study indicates 47% of the challenges facing BI are data quality related [8]. Figure 15 illustrates the 7 data quality steps adopted by WITS [12] to support the traditional BI approach.
  The steps are explained below:
    - Step 1 - Active data governance is designed to prevent data errors being entered (where possible) into the system by introducing validation rules and drop down lists on the ERP system. Active data also includes training on systems and processes for data capturers and relevant stakeholders;
    - Step 2 - Data profiling is conducted on an ad-hoc basis by Data Proprietors and Offices of Record to find and rectify data errors;
    - Step 3 – Reports are sent out at various times of the year by data stewards for departments to perform reality checks on the data captured, so that corrections can be made timeously;

- Step 4 – Alerts are triggered when new records are captured and all the required fields have not been completed. The missing data is emailed to the relevant manager;
- Step 5 – ETL process rejects any records that do not meet the integrity requirements for the data warehouse. The error records are sent to data stewards for correction;
- Step 6 – Audit processes validate that the data in the data warehouse is equivalent to the source system;
- Step 7 – Data is further validated against rules for the Higher Education Management Information System (HEMIS) which is the system used for statutory reporting to government

This indicates that irrespective of the type of BI approach being used by an organisation, assuring data quality weighs heavily on the success of the implementation. These steps need to be adopted and followed in both types of implementations.

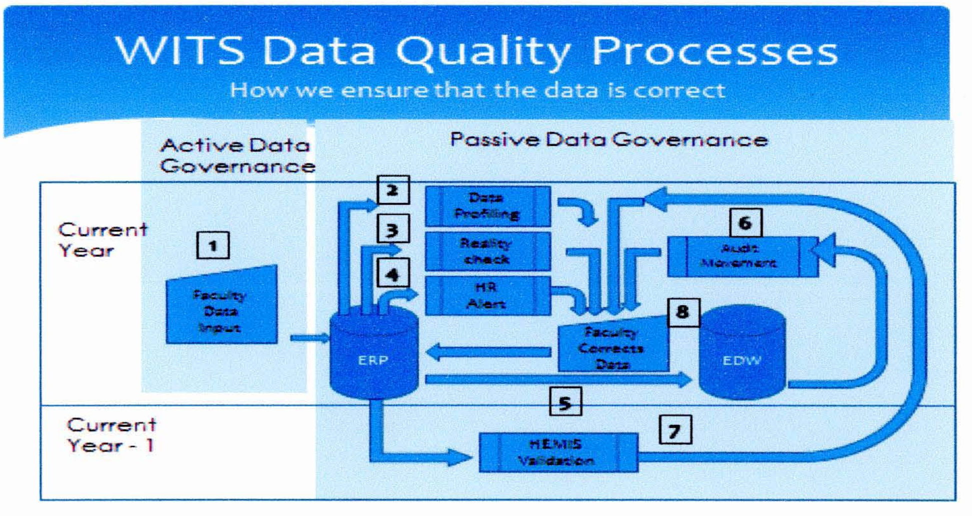


Figure 15 : Wits Data Quality Process [12]

5.12. Durability/Persistence of Data

In in-memory databases the data is held in volatile main memory and all data can be lost during a reset or power failure. That is, in-memory databases do not support the durability of the ACID (atomicity, consistency, isolation and durability) requirement. However there are several ways to overcome this issue. For example, using transaction logging, non-volatile RAM or performing online backups [14].

As in-memory database technology matures, the durability and persistence concerns will be addressed.

5.13. User Perceptions Results

The user survey required the users of the Wits Procurement dashboards to evaluate the traditional BI tool versus the in-memory tool with respect to the following criteria

a) Navigation;
b) Performance;
c) Visualisation;
d) Reporting.

The survey questions posed to the users, for each of the above criteria, are shown in Appendix F.

Users of the WITS Procurement dashboards, which were developed using the traditional BI approach, were required to rate their experience regarding OBIEE tool.

The same users were given access to the in-memory dashboards developed using QlikView. Users were required to rate their user experience with respect to the QlikView tool.

The same questions were posed to users for both tools. The rating scale used was

- 1 – Strongly disagree
- 2 – Disagree
- 3 - Neither Agree not Disagree
- 4 - Agree
- 5 - Strongly agree

An average of the user responses was calculated for each of the categories for each tool. The results are shown in Figure 16.

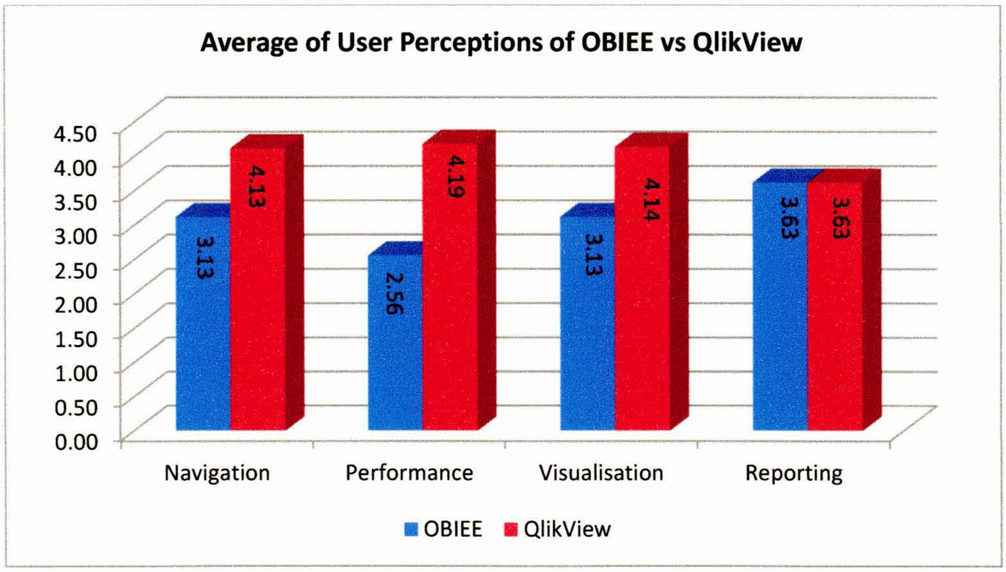


Figure 16 : User Perceptions of OBIEE versus QlikView

On average, the user experience using the QlikView tool exceeds that of the OBIEE tool in the navigation, performance and data visualisation categories. Appendix G shows the detailed analysis of each category.

The reporting capabilities of both tools were rated the same by the users.

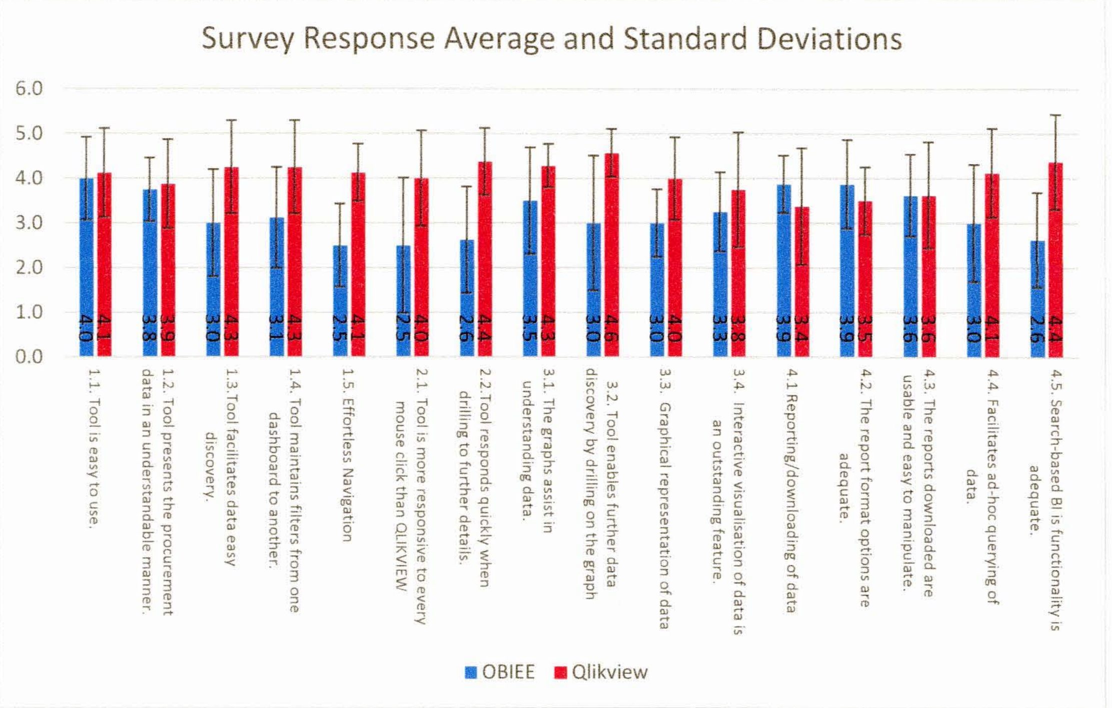


Figure 17 : Survey Responses showing the standard deviations

For each of the detailed questions the means and standard deviations represented in Figure 17 indicate that for OBIEE and QlikView tools; the difference between the two means is not statistically significant. That is, users perceived the Qlikview tool to be marginally better than the OBIEE tool.

The distributions of the scores are represented in Figure 18. For the navigation category the histogram is skewed right for QlikView and is centred for OBIEE indicating the user preferred the navigational functionality in QlikView. The same pattern is observed for the visualisation and reporting categories with the exception of performance where the distribution is for OBIEE performance has 2 peaks. This indicates the some users strongly disagreed and the same number of users neither agreed nor disagreed on the OBIEE tools performance. However, the performance of the QlikView tool was preferred over the performance of the OBIEE tool.

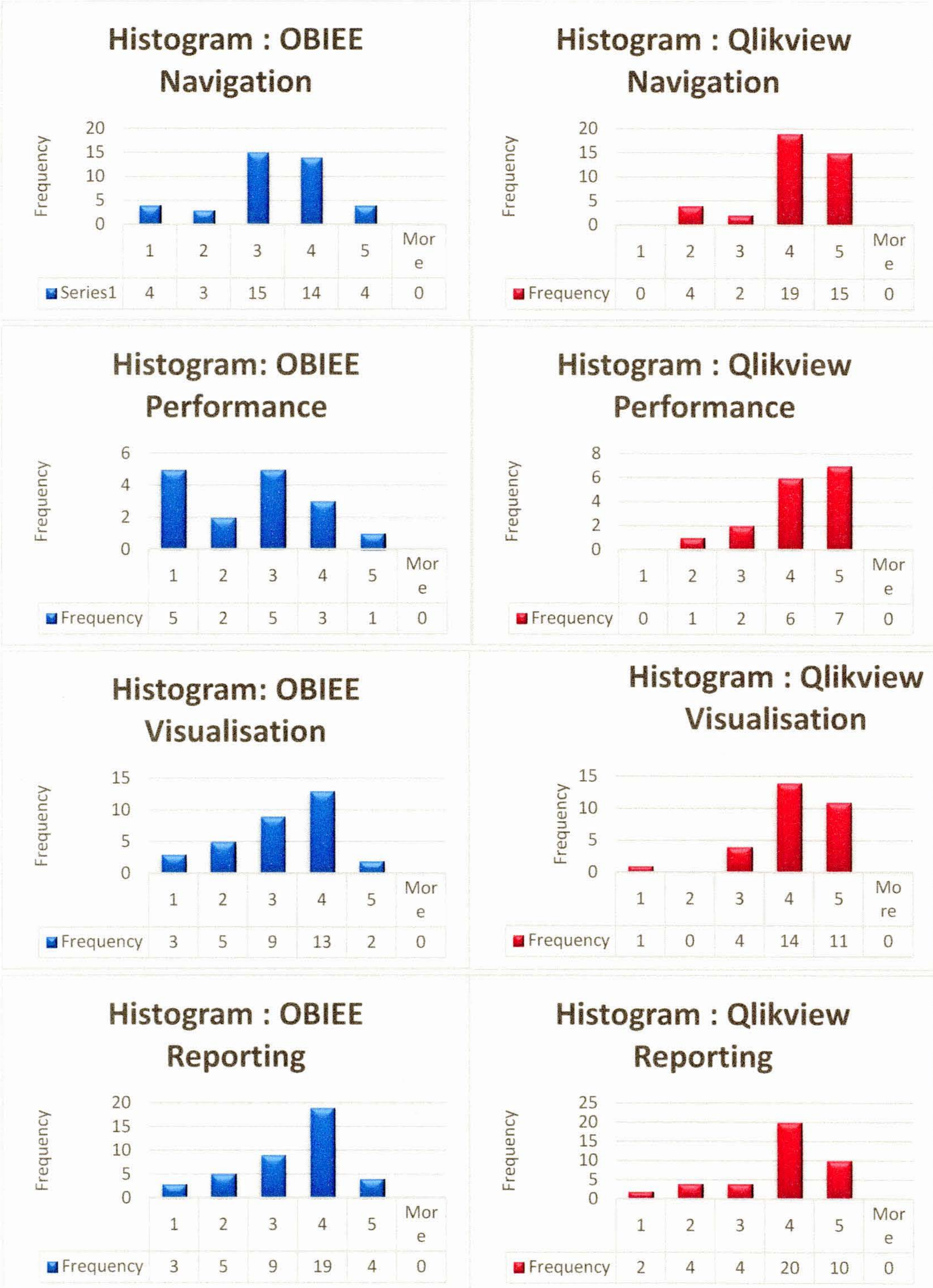


Figure 18: Survey Categories Histograms

The results confirm the benefits of the in-memory approach over the traditional BI approach for certain evaluation criteria, such as, ad-hoc querying and data discovery. As shown in Figure 19, queries requiring data to be fetched from disks take a longer time. The performance of in-memory databases to ad-hoc querying is significantly faster than accessing data which is stored on disk. The in-memory approach supports the 'speed of thought' analysis by enabling users to build on previous queries interactively, thereby facilitating a deeper analysis in a shorter period of time [29].

The faster performance allows for more interactive and visually rich dashboards to be created thereby improving on the user experience.

Action	Time
Main memory access	100 ns
Read 1 MB sequentially from memory	250,000 ns
Disk seek	5,000,000 ns
Read 1 MB sequentially from disk	30,000,000 ns

Figure 19 : Access and Read times from Disk and Main Memory [29]

The results from a survey conducted on the QlikView customer experience [24] on QlikView clients corroborate with results from the survey conducted on Wits procurement users.

The results presented in Figure 20 are based on the user responses to the survey question:

> *What was the increase in customer satisfaction after the QlikView implementation (where an increase was reported)?*

The number of respondents to this question were 244. The average increase in customer satisfaction after the QlikView system was 37%. The increase in customer satisfaction represents a positive indicator for business growth as the BI system can monitor the measurements of customer satisfaction as well as monitor the business performance and then make correlations to ensure that improvements in one is not offset by unexpected reductions in the other[24].

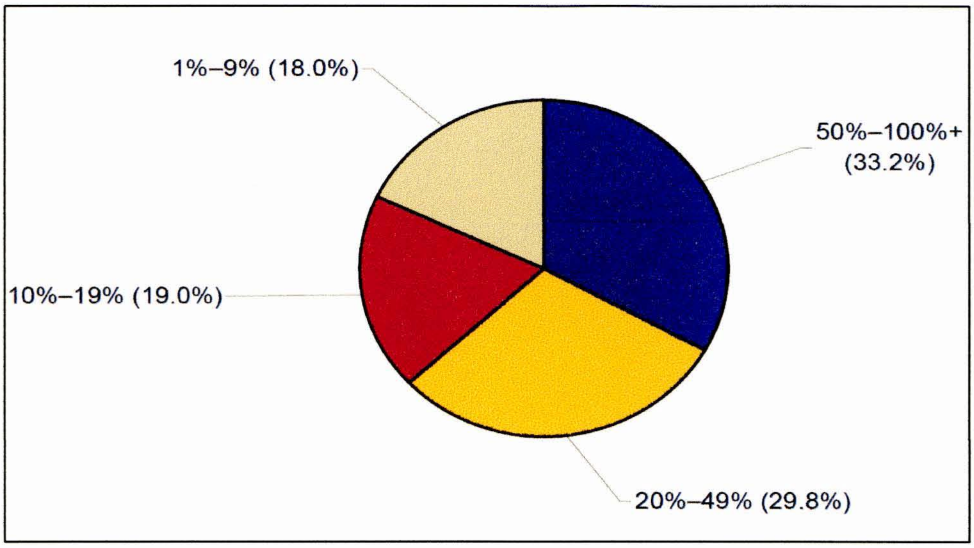


Figure 20 : Increase in Customer Satisfaction. Source: IDC survey of QlikView customer base, Jan–Mar 2009 [24]

The results presented in Figure 21 are based on the user responses to the survey question:

> *What was the increase in business agility after the QlikView implementation (where an increase was reported)?*

The number of respondents to this question were 306. The average increase in business agility achieved by QlikView customers was 39%. The "business agility" in the survey question referred to the ability of the business to change its operations in response to external changes and the ability to swiftly measure performance and monitor effects of any changes implemented.

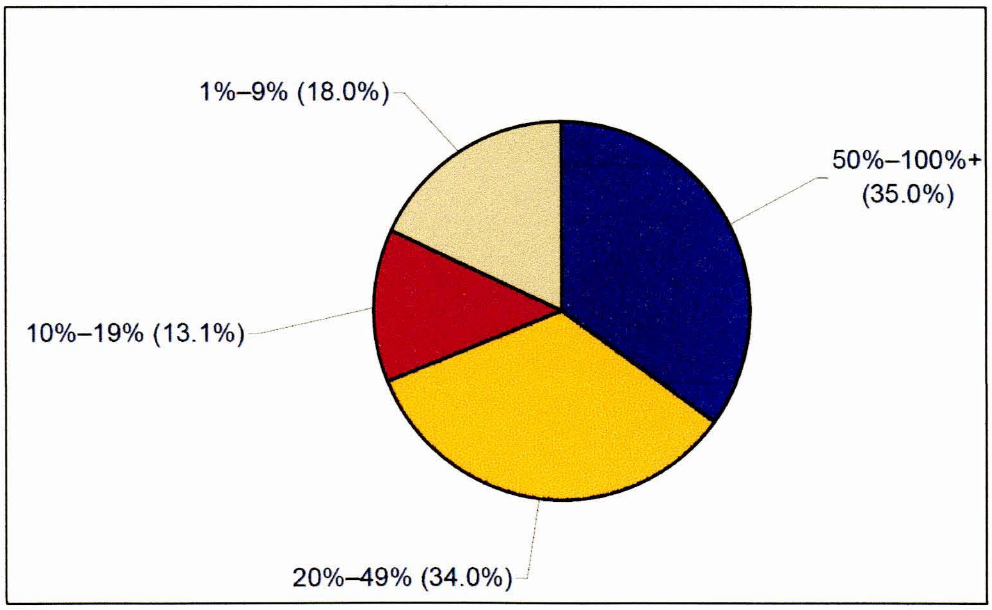


Figure 21 : Increased Business Agility after QlikView. Source: IDC survey of QlikView customer base, Jan–Mar 2009 [24]

5.14. Extract and Load Times

A comparison was performed on the extract and load times, for loading procurement data into the in-memory tool QlikView versus loading time for procurement data into the data warehouse.

The loading criteria for traditional data warehouse environment was an incremental load. That is, only new and changed records are extracted for update or insert into the data warehouse. The loading criteria for the in-memory environment was to store the history (years 2010 to 2013) into compressed files. Only the current year's data was extracted for every load. Both history and current year's data is loaded into memory.

The average extract and load time for a month in the traditional BI approach ranges from 156 minutes to 241 minutes. The average extract and load time for in-memory approach was 10 minutes. See Table 2 and 3 below.

The time difference of the extract and load times for the two environments confirms that data can be loaded into an in-memory database much faster than the traditional BI approach. Data is accessible and useable in the in-memory environment in near real-time whereas due to the multiple layers and transformation steps required for the traditional approach, the same data can only be accessed approximately 3 and half hours later.

It should be noted that the hardware specifications for the two environments are not comparable as OBIEE is on a production server and the QlikView environment was developed on standard 64 bit workstation with additional memory. The time to have the data available for analysis using QlikView is considerably less even with hardware of lower specifications.

	QLikView
	Load time (Minutes)
Load 1	0:07:17
Load 2	0:07:11
Load 3	0:07:11
Load 4	0:06:56
Load 5	0:06:56
Load 6	0:07:39
Load 7	0:05:45
Load 8	0:05:39
Load 9	0:14:47
Load 10	0:24:40
Load 11	0:13:48
Load 12	0:13:55
Average	**0:10:09**

Table 2: QlikView Extract and Load Times

OBIEE	
	Average load times (Minutes)
Apr-13	156
May-13	183
Jun-13	217
Jul-13	241

Table 3: OBIEE Extract and Load Times

.

The results presented in Figure 22 are based on the user responses to the survey question:

> *How did QlikView change the time for end users to generate and access, and analyse information?*

The time taken to generate and access information refers to the time from loading data into QlikView to when the data is given to the user for analysis. The results indicate that there was an average reduction of 51% in the time taken to generate and access the information. The average change in time to analyse the data was reported a reduction of 48%. This could be attributed to the ease of use of the tool in terms of navigation and easier to understand the data due to the rich visualisation in the dashboard [24].

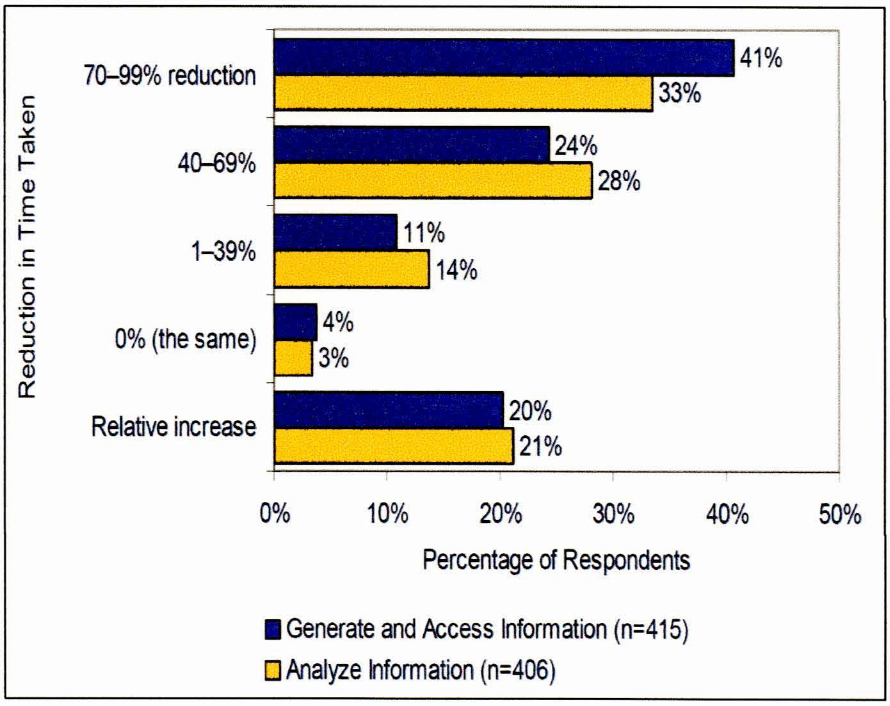


Figure 22 : Time taken to Generate, Access and Analyse Data Source: IDC survey of QlikView customer base, Jan–Mar 2009 [24]

6. DISCUSSION

6.1. Are these two approaches to BI mutually exclusive or is there sufficient justification for them to co-exist within the same BI ecosystem?

The in-memory approach to BI offers many benefits to users in terms of time to value, fast querying times, strong visualisation and intuitive and interactive data discovery interfaces. This approach meets the organisational demand for analysing large volumes of data in near time. Although the approach offers numerous benefits, the maturity level of the technology does not indicate that it can replace the traditional approach to BI. Research indicates [19] that at this point in time the traditional and in-memory approaches can both be leveraged to meet organisational demands. That is, both the technologies can co-exist within the current BI ecosystem to deliver business value.

6.2. What are impacts and requirements of a real-time or near real-time offering on the business processes and procedures of an organisation?

The in-memory approach can be used in organisations with a mature BI infrastructure where the business requirement indicates the need for near real-time analysis of data.

Organisations that require to move and analyse larger volumes of data in near real time can achieve this by using the in-memory approach like QlikView which is not dependent on a lengthy ETL process before data is loaded and available for analysis. Other approaches like Exalytics and SAP Hana can be used however these will require large investments by organisations in hardware and software.

Depending on the approach being used the organisation should be careful not to create information silos that could result in data integrity issues.

6.3. Does the in-memory associative approach shorten the development life cycle of a subject specific dashboard and deliver improved value, while at the same time being a more cost effective and faster approach than the traditional methodology used for BI?

The development times of building the Wits Procurement dashboards using the in-memory QlikView technology when compared to the development times of the same subject area using the traditional approach indicates that the development life cycle can be considerably shortened. Depending on technology used and the configuration, the in-memory approach can be a more cost effective approach. However, this depends on each organisations requirements and infrastructure strategies. The memory costs can escalate if the number of users increases or the amount of data increases as the amount RAM required will need to be increased accordingly.

The user perceptions survey results indicate the overall user experience was better on the in-memory approach when compared to the traditional approach for a specific subject area namely, procurement data. However, it should be noted that this approach will not meet the requirements of users requiring further analysis that extends beyond the procurement domain. This is, where the benefits of the traditional approach can be leveraged as it uses an integrated data warehouse, which has the capability of reporting across several related domains.

6.4. What are business requirements that will require both technologies to be used?

Organisations are being challenged daily to maintain a competitive edge in order to remain viable. The constant change in the landscape requires organisations to be flexible, responsive and agile. These business requirements are steering organisations to analysing large complex set of data in short time frames and respond accordingly based on the analysis.

At the same time organisation must remain accountable to various statutory bodies and be able to produce reliable and validated information. It is in these situations organisations will need to maintain both the traditional BI approach and the fast and flexible approach which is facilitated by in-memory analytics.

Organisations adopting the in-memory technology will need to carefully manage the resource requirements in terms of infrastructure and personnel. There are numerous challenges that an organisation will need to address before adopting an in-memory BI approach (refer to section 4.3.2.3). The justification for adopting this approach must be clear and the benefits should overweigh the drawbacks.

7. CONCLUSION

In-memory databases and in-memory analytics offers organisations an alternative approach to delivering BI, to meet the organisational demand for real-time, fast, flexible and cost effective BI.

In the current BI ecosystem, these technologies will play a complementary role with respect to the traditional approach to BI. While adding value in terms of performance and data discovery, in-memory databases and in-memory analytics have not yet reached the maturity level that addresses certain key concerns of data persistence, data security, scalability and impact of information silos.

However, in-memory databases and in-memory analytic tools provide organisations that have little or no investment in BI, with the ability to introduce BI using a phased or incremental approach. This will be particularly applicable to environments like small businesses or organisational departments, where data informed decision windows are small and the traditional BI approach is not meeting the demand timeously.

The users of BI systems are demanding that the tools are fast, visually rich, interactive and enable self-service BI. The user responses to the QlikView survey indicate that the overall user experience, when using the in-memory tools, exceeded that of the traditional BI tool, in terms of performance, navigation and data visualisation, thereby meeting the demand. The shortened development time, using the in-memory databases, is also an attractive feature for organisations that are changing constantly due to external drivers such as competition, cost and regulatory demands. These organisations require insights into rapidly changing, high volumes of data, within small turnaround times.

It is concluded that in-memory databases and in-memory analytics can complement and co-exist within the current BI ecosystem, as they address many of the challenges facing organisations today. As the technology matures, the adoption rate will increase as many organisational perspectives on analytics is moving from 'nice to have' to a 'must have' in order to gain a competitive advantage. A further driver for the adoption of this technology is the reducing cost of memory making the technology accessible to various types of organisations and the reduced time to value due to the quick implementation cycles.

Future Research

Future research on in-memory databases and in-memory analytics would be to evaluate their use on operational systems, impact on mobile computing and how the technology can be leveraged for big data analytics. Further investigation is required of the role of the in-memory databases in making analytics transparent to non-traditional users by embedding analytics into processes at the point of decisions or actions.

8. REFERENCES

[1] T. H. Davenport, "Competing on Analytics," *Harvard Business Review,* pp. 1-12, January 2006.

[2] H. Koeppel, "BI and Analytics : Are We There Yet?," February 2013. [Online]. Available: http://searchcio.bitpipe.com/data/document.do?. [Accessed 19 May 2013].

[3] D. W. Cearley and C. Claunch, "The Top 10 Strategic Technology Trends for 2013," Gartner, 2013.

[4] S. Drew, "The New Business Intelligence is Changing the Game," 22 May 2012. [Online]. Available: http://midsizeinsider.com/en-us/article/the-new-business-intelligence-is-changin. [Accessed 20 February 2013].

[5] "Business Intelligence," [Online]. Available: http://en.wikipedia.org/wiki/Business_intelligence. [Accessed 10 March 2013].

[6] K. Schlegel, R. L. Sallam, D. Yuen and J. Tapadinhas, "Magic Quadrant for Business Intelligence and Analytics Platforms," Gartner, 2013.

[7] C. S. Kadali H., "Developing a Single version of the Truth," 2010. [Online]. Available: http://www.pwc.com/us/en/supply-chain-management/pdfs/business-intelligence-practice-overview.pdf. [Accessed 21 May 2013].

[8] Pant, Prashant;, "Business Intelligence (BI) How to build a successfull BI strategy," Deloitte Development, 2009.

[9] E. Rahm and H. H. Do, "Data Cleaning: Problems and Current Approaches," [Online]. Available: http://dc-pubs.dbs.uni-leipzig.de/files/Rahm2000DataCleaningProblemsand.pdf. [Accessed 25 May 2014].

[10] N. Vodapalli, "Critical Sucess Factors of a Business Intelligence Implementation," 2009.

[11] K. P. Celina M. Olszak and Ewa Ziemba University of Economics, "Critical Success Factors for Implementing Business Intelligence Systems in Small and Medium Enterprises on the Example of Upper Silesia, Poland," *Interdisciplinary Journal of Information, Knowledge, and Management,* vol. Volume 7, pp. 130-150, 2012.

[12] Mc Loughlin K., "WITS Data Governance Framework," 2011.

[13] M. Pezzini, R. Edijali and D. Feinberg, "Hype Cycle for In-Memory Computing Technology 2012," Gartner, 2012.

[14] L. V. Barbagallo, "In-Memory Database : Competitive Landscape and Performance Analysis," 2008.

[15] S. Graves, "In-Memory Database Systems," 01 September 2002. [Online]. Available: http://www.linuxjournal.com/node/6133/print. [Accessed 27 May 2013].

[16] Yellowfin, "Making Business Intelligence easy," 2010.

[17] "In-Memory Business Intelligence," April 2009. [Online]. Available: http://www.ranwood.nl/attachments/061_Article%20In%20Memory%20BI%20solutions.pdf. [Accessed September 2012].

[18] M. Pezzini, D. Feinberg, A. Norwood, R. Edjlali, J. Unsworth and J. Rischardson, "Predicts 2013 : In-Memory computing : Growing Gains, but also Growing Pains," Gartner, 2012.

[19] R. Winter, S. Bischoff and F. Wortman, "Revolution or Evolution? Reflections on In-Memory Appliances from an enterprise Information Logistics Perspective," in *Innovative Enterprise Applications with In-memory data management*, Bonn, 2011.

[20] O. Grabova, J. Darmont, J.-H. Chauchat and I. Zolotaryova, "Business Intelligence for Small and Middle-Sized Entreprises," *SIGMOD Record,* vol. 39, no. 2, pp. 39-50, 2010.

[21] "ORACLE EXALYTICS IN-MEMORY MACHINE: A BRIEF INTRODUCTION," Oracle, January 2014. [Online]. Available: http://www.oracle.com/us/products/applications/exalytics-x24-1500259.pdf. [Accessed 7 May 2014].

[22] J. Kelly, "Primer on SAP HANA," Webi, 12 June 2013. [Online]. Available: http://wikibon.org/wiki/v/Primer_on_SAP_HANA. [Accessed 8 May 2014].

[23] D. English, "Whitepaper-Understanding_QlikViews_Associative_Architecture_v1f," 12 May 2010. [Online]. Available: http://Qlikview.com. [Accessed September 2012].

[24] A. Woodward; D. Vesset, "Time to Value and ROI from BI : The Qlikview Customer Experience," October 2009. [Online]. Available: http://www.plannersistemas.com.br/propaganda/WP-IDC3_time_to_value_BI-EN.pdf. [Accessed 11 May 2014].

[25] M. R.L., "12 Predictive Analytics Screw Ups," 24 June 2013. [Online]. Available: http://www.predictiveanalyticsworld.com/patimes/12-predictive-analytics-screw-ups/. [Accessed 24 August 2013].

[26] Verdiensen, Jan, "Qlikview : In-memory vs. Traditional BI," [Online]. Available: http://www.osp.ru/data/130/404/1225/In-Memory.pdf. [Accessed 20 June 2013].

[27] C. Saran, "CIOs must reskill to succeed with in-memory computing," 4 April 2013. [Online]. Available: http://www.computerweekly.com/news/2240180804/CIOs-must-reskill-to-succeed-with-in-memory-computing-says-Gartner. [Accessed 1 June 2013].

[28] P. Loos, J. Lechtenbörger, G. Vossen, A. Zeier, J. Müller, W. Lehner, D. Kossmann, B. Fabian B, O. Günther and R. Winter, "In-memory Databases in Business Information Systems," 5

November 2011. [Online]. Available: http://link.springer.com/article/10.1007/s12599-011-0188-y/fulltext.html. [Accessed 29 March 2013].

[29] H. Plattner and Z. A, “In-Memory Data Management,” Springer, Berlin, 2012.

[30] “In-Memory Database,” August 2012. [Online]. Available: http://whatis.techtarget.com/definition/in-memory-database. [Accessed 17 October 2012].

[31] “IT-Glosasary,” [Online]. Available: http://www.gartner.com/it-glossary/business-intelligence-bi/. [Accessed 2012 September 2012].

[32] “Data Visualisation,” [Online]. Available: http://searchbusinessanalytics.techtarget.com/definition/data-visualization. [Accessed Spetember 2012].

[33] “SearchBusinessAbnalytis,” [Online]. Available: http://searchdatamanagement.techtarget.com/definition/columnar-database. [Accessed September 2012].

[34] “Business intelligence,” 14 May 2013. [Online]. Available: http://en.wikipedia.org/wiki/Business_intelligence#cite_note-6. [Accessed 21 May 2012].

[35] Microsoft, “Business Intelligence Guidelines Conceptual Framework,” Microsoft, 2006.

[36] David Loshin, “Gaining the Performance Edge using a column-Orientated Database Management system,” [Online]. [Accessed 11 August 2013].

[37] Persistent Systems Ltd., “How to Enhance Traditional BI Architecture to Leverage Big Data,” Santa Clara, 2013.

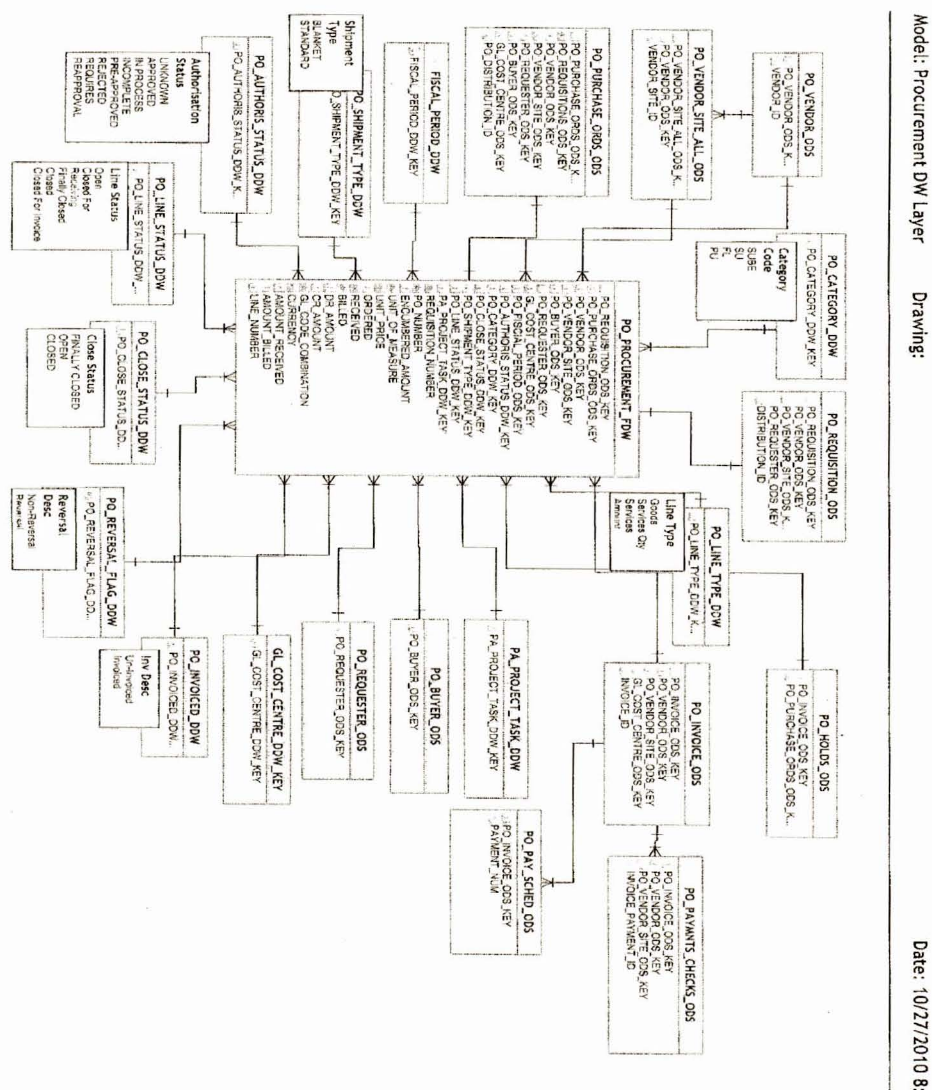
9. LIST OF RESOURCES

9.1.1. QlikView Evaluation Copy
9.1.2. Access to Procurement data
9.1.3. Access to Procurement BI dashboards
9.1.4. Access to 64 bit computer

10. APPENDICES

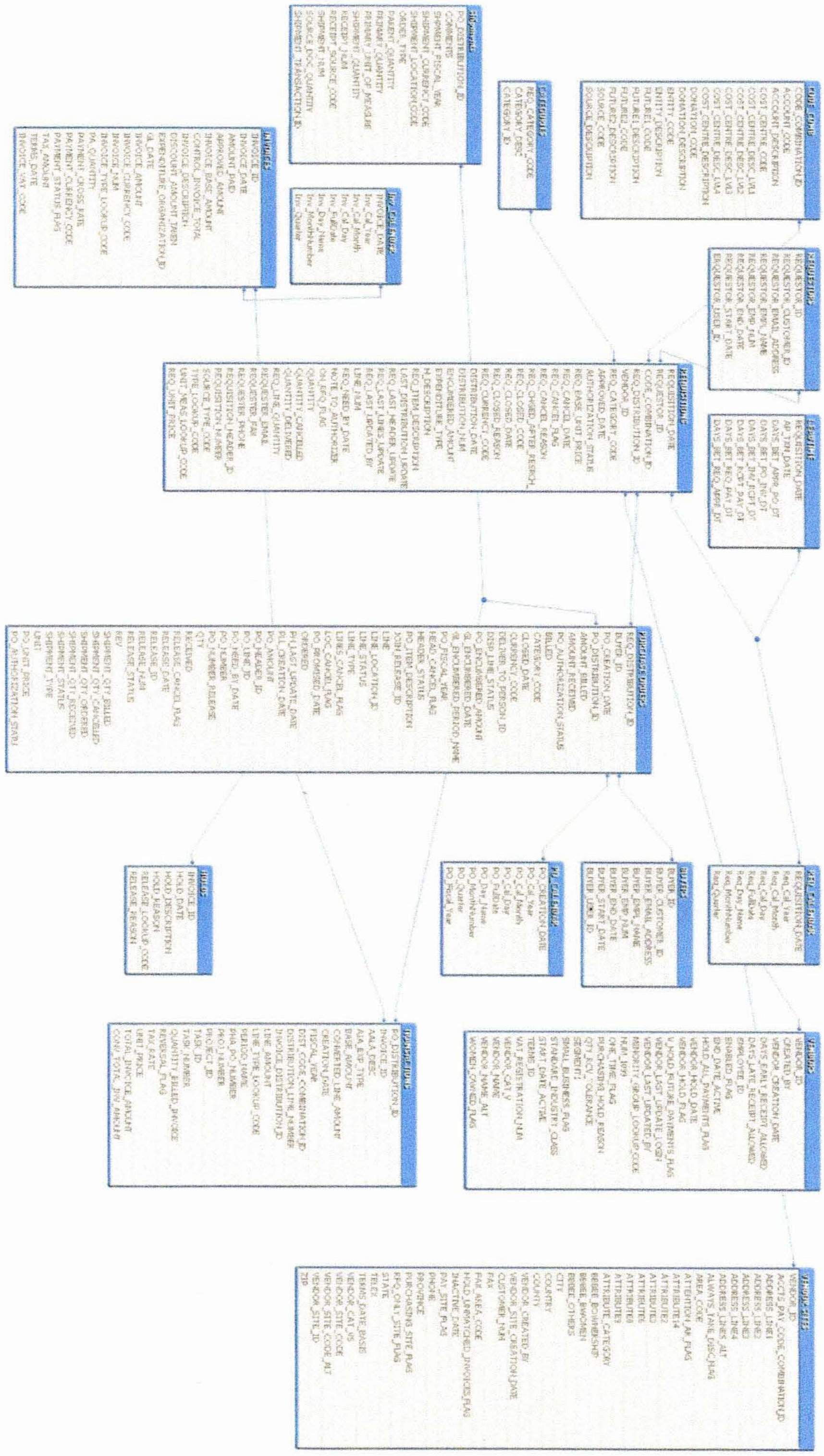
APPENDIX A

Wits Procurement ERD in the data warehousing (DW) layer



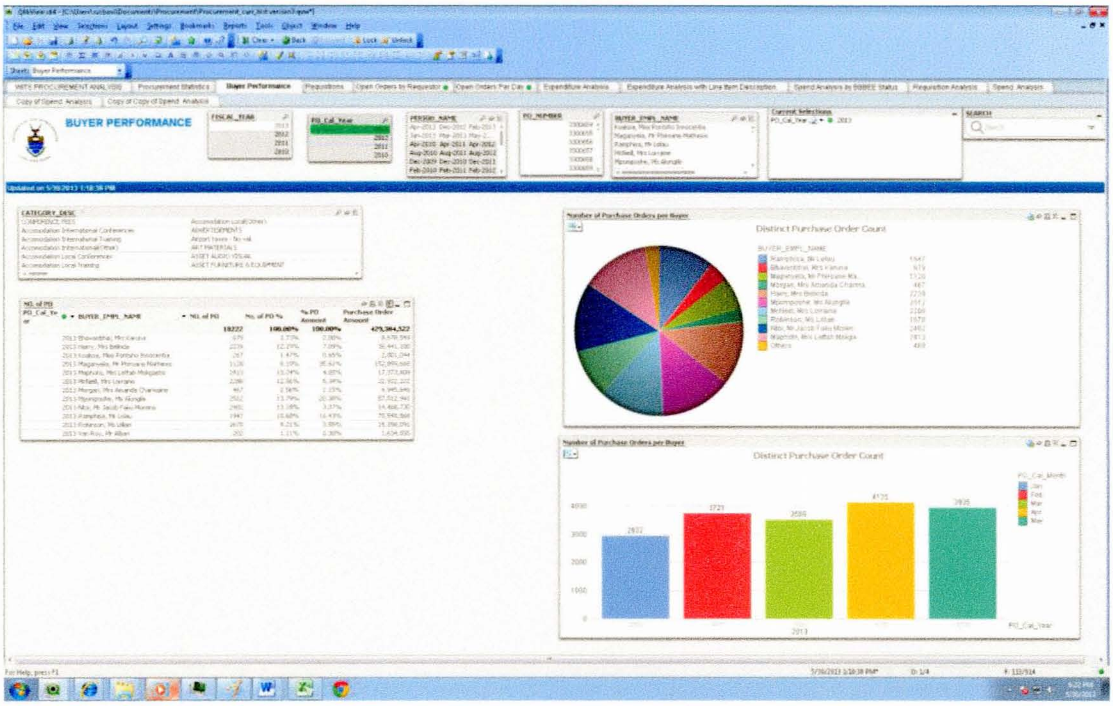
APPENDIX B

QlikView Table Relationship Structure

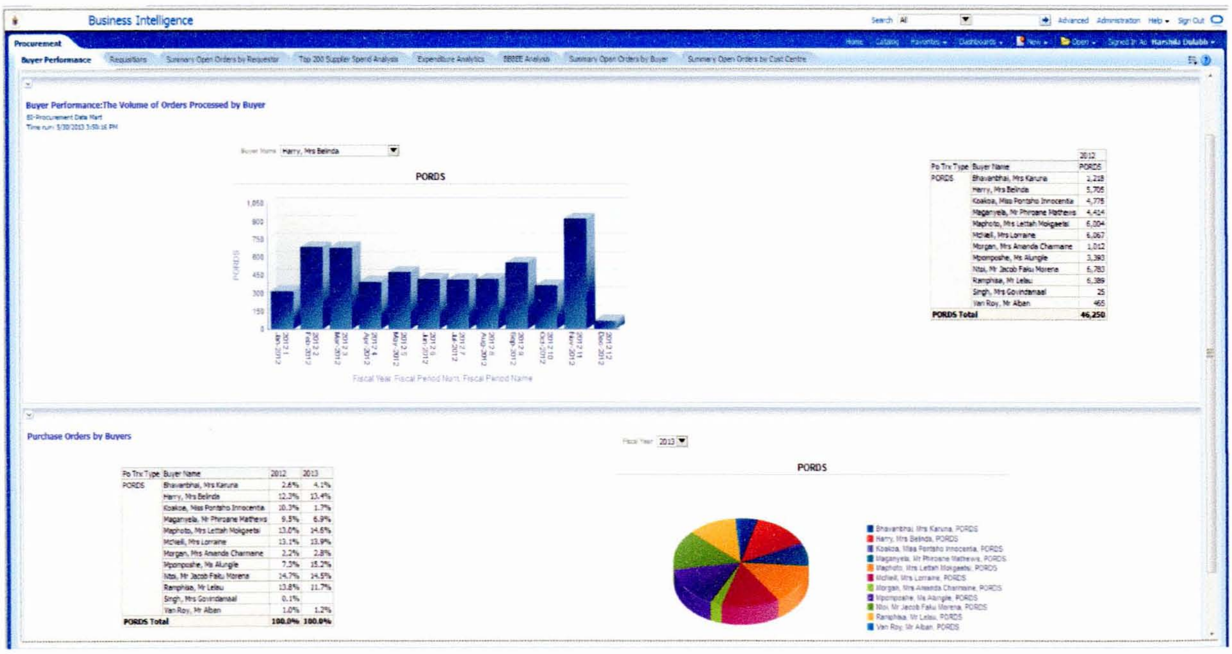


APPENDIX C

QlikView Data Visualisation Dashboard



OBIEE Data Visualisation Dashboard



APPENDIX D

DEFINITIONS OF TERMINOLOGY

a) In-Memory Database

In-memory database (IMDB) or main memory database (MMDB) is a database where the data is stored in the main memory in order to facilitate faster response times by streamlining the work involved in processing database queries. Data is manipulated and stored in the main memory thereby eliminating the need for disk access. Within the Business Intelligence context IMDB are generally read only databases that store historic and current data for querying and reporting [1].

b) Associative Architecture

Every field in a selected dataset is associated to every other field in the dataset through the key fields that link the tables together. When a data point is selected, all other fields are filtered based on the association and aggregates that are calculated in real time without a query being written [6].

c) Business Intelligence

Business intelligence (BI) is an umbrella term that includes the applications, infrastructure and tools, and best practices that enable access to and analysis of information, to improve and optimize decisions and performance [3].

d) Dashboards

Dashboards are a data visualisation mechanism that are used to display an organisations performance against key indicators by allowing users at all levels to gain insight into a performance area at a high level. Various visualisation techniques can be employed depending on the subject area, for example, charts, gauges etc. Dashboards provide the facility to drill down to more detailed levels for when further understanding is required. Dashboards can be easily rendered on various mobile devices giving users the ability to have access to time critical information. The purpose of dashboards is to enable business to improve decision making by having access to insightful information and in some cases real-time information [3].

e) Data visualisation

When data is represented graphically with gauges, heat maps, tree maps or with various types of charts that allow users to "see" data in a particular context and thereby allow them to have an easier way understand the information represented, it is called data visualization. This form of representing data allows the user to gain insight into data and discover patterns and trends over time by interacting and drilling down into the detail of the data. Data Visualization can also be used to alert users e.g. if certain thresholds have been passed [4].

f) Online Analytic processing (OLAP)

OLAP is computer processing that enables a user to easily and selectively extract and view data from different points of view. The data is stored in a multidimensional database where each attribute is a dimension.

g) Real Time Analytic Reporting

Real-time analytic reporting requires that data be visible to the end users with a minimal lag time from when it was captured to when it can be used to make decisions. This is achieved by using in-memory and columnar databases whose architecture supports fast retrieval of data which can be aggregated/manipulated in memory without the need for complex queries or aggregation rules.

h) Columnar Database

Data is stored in columns which facilitates efficient read and write from the hard disk storage and reduces the query response times. The data in a columnar database can be compressed and is self-indexing, therefore requires less space than relational databases. This form of storage also permits operations to be performed very quickly e.g. sum, max etc. [5]

i) Data latency

This is the time it takes to collect raw data, prepare it for analysis, and store it where it can be accessed and analysed. Important functionality here includes data profiling, extraction, validation, cleansing, transformation, integration, transformation, delivery, and loading.

j) NAND flash memory

NAND flash memory is a type of non-volatile storage technology that does not require power to retain data.

k) Context-awareness

The ability of an application to understand and use the current context of the users to adopt the application's operations, thereby delivering high quality and personalised user experience.

The ability of a system to understand and use the current context of the system to adapt the system's operations, thereby achieving high system performance

l) Big Data

Big Data is characterised by the "Four Vs": volume, velocity, variety and value.

Volume:

The large amount of data being generated by numerous systems and devices.

Velocity:

The rapid rate at which the data changes.

Variety:

The many types of data and sources, from databases to audio and video objects, unstructured mobile and social data.

Value:

Improvement in analysing Big Data yields more value for an organisation. For example, the ability of a system to understand and use the current context of the system to adapt the system's operations, thereby achieving high system performance

APPENDIX E

Gartner Hype Cycle

The following definitions have been extracted from 'Understanding Gartner's Hype Cycles, 2012'

Hype Cycles:

- Establish the expectation that most technologies will inevitably progress through the pattern of overenthusiasm and disillusionment, followed by eventual productivity;
- Provide a snapshot of the relative maturity of technologies within a certain segment, such as a technology area, horizontal or vertical business market, or a certain demographic audience;
- Show the speed at which each technology is progressing through the Hype Cycle by indicating how long it will take to reach the Plateau of Productivity and the start of mainstream adoption.

Hype Cycles help technology planners to decide when to invest in a technology. If a company launches its efforts too soon, it may suffer unnecessarily through the painful and expensive lessons associated with deploying an immature technology. If it delays action for too long, it runs the even-greater risk of being left behind by competitors that have succeeded in making the technology work to their advantage.

What Is the Hype Cycle?

The Hype Cycle is a graphical depiction of a common pattern that arises with each new technology or other innovation. Each year Gartner creates more than 90 Hype Cycles in various technology and application areas (such as social computing and ERP), information and IT services (cloud computing, big data) and industry (retail, life insurance) domains as a way for clients to track technology maturity and future potential.

Gartner's Hype Cycle, introduced in 1995, characterizes the typical progression of innovation, from overenthusiasm through a period of disillusionment to an eventual understanding of the innovation's relevance and role in a market or domain (see Figure 23).

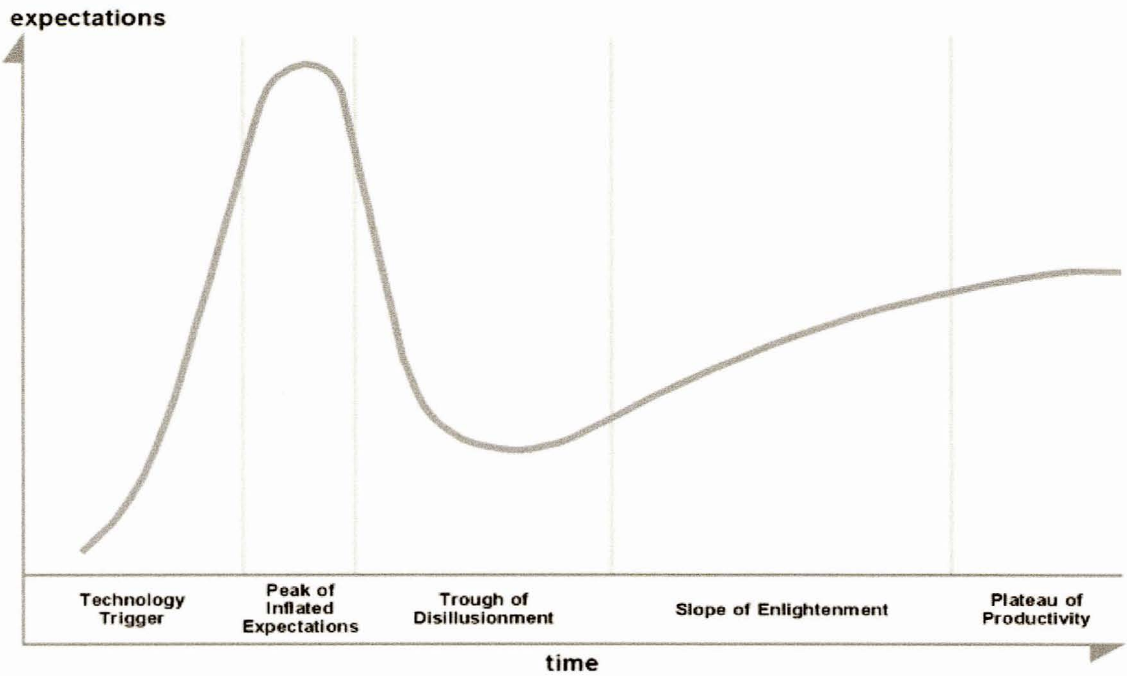


Figure 23. The Hype Cycle Source: Gartner (June 2012)

A technology (or related innovation) passes through several stages on its path to productivity:

- **Technology Trigger:** The Hype Cycle starts when a breakthrough, public demonstration, product launch, or some other event generates press and industry interest in a technology innovation;
- **Peak of Inflated Expectations:** A wave of "buzz" builds and the expectations for this new technology rise above the current reality of its capabilities. In some cases an investment bubble forms, as happened with the Web, social media and cloud computing;
- **Trough of Disillusionment:** Inevitably, impatience for results begins to replace the original excitement about potential value. Problems with performance, slower-than-expected adoption or a failure to deliver financial returns in the time anticipated all lead to missed expectations, and disillusionment sets in;
- **Slope of Enlightenment:** Some early adopters overcome the initial hurdles, begin to experience benefits and recommit efforts to move forward. Drawing on the experience of the early adopters, understanding grows about where and how the technology can be used to good effect and, just as importantly, where it brings little or no value;
- **Plateau of Productivity:** With the real-world benefits of the technology demonstrated and accepted, growing numbers of organizations feel comfortable with the now greatly reduced levels of risk. A sharp uptick ("hockey stick") in adoption begins, and penetration accelerates rapidly as a result of productive and useful value.

APPENDIX F

QlikView-OBIEE User Survey

A Comparison of Procurement Dashboards Using OBIEE versus QLIKVIEW
Background Information
As organisations are becoming more data centric there is higher demand for real-time analytic information to be readily available in order to facilitate better decision making across all levels of an organisation, so that businesses can leverage their data to gain insightful competitive advantage as well as manage the organisation more effectively.
There are several approaches/tools that enable the delivery of information to users so that they can make informed decisions or take informed actions. The traditional approach is the use of reporting tools on data that is stored in a data warehouse. A more recent approach is one that facilitates real-time data analysis using an in-memory database
The WITS procurement users have access to procurement dashboards via Oracle Business Intelligence Enterprise Edition (OBIEE) reporting tool which uses the data stored in the data warehouse. QLIKVIEW is an in-memory database and data visualisation tool that allows users to analyse data through various graphical representations and reports. The Wits procurement data has been extracted from the Oracle ERP system and loaded into QLIKVIEW. Various dashboards have been created in QLIKVIEW to facilitate data discovery and analysis.
The purpose of this survey is to compare and assess the user experience on the OBIEE and QLIKVIEW tools and each tools contribution to faster and easier decision making.
Please note that your response is very important and there are no right or wrong answers.
The focus in this survey focus is on evaluation of the tool rather than the data/data quality.
Please indicate the extent of your agreement with each statement, by selecting from the list of options. Detailed instructions are provided with the questions. The entire survey should take between 10 to 15 minutes to complete.
Thank you for participating in this survey.

This questionnaire is divided into 2 sections. Please answer all the questions which follow to the best of your ability. There are no right or wrong answers.

Have you used the BIS OBIEE Procurement dashboards in the past to obtain information and subsequently made a decision using that information? * If you reply no, you are not required to proceed. I would like to nevertheless thank you for having considered participating. *

- ( ) Yes
- ( ) No

Section A

For each of the following statements, please rate your level of agreement regard to the Procurement Dashboard on OBIEE, on a scale from 1 to 5.

1 = "Strongly Disagree"; 2 = "Disagree"; 3 = "Neither Disagree nor Agree"; 4 ="Agree" and 5 ="Strongly Agree".

1 OBIEE Navigation

1.1. The OBIEE tool is easy to use. *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.2. The OBIEE tool presents the procurement data in an understandable manner. *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.3. OBIEE tool facilitates data easy discovery. *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.4. OBIEE tool maintains filters from one dashboard to another. *

Note : Filters are the selections made for which data you would like to see

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.5. The procurement dashboard allows you to move from one subject area to another effortlessly in order to find information you looking for in OBIEE. *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

2. OBIEE Performance

2.1. The OBIEE tool is more responsive to every mouse click than QLIKVIEW

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

2.2. OBIEE responds quickly when drilling to further details.

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

3. OBIEE Data Visualisation

3.1. The graphs in OBIEE help in the further understanding the procurement data.

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

3.2. The OBIEE tool enables further data discovery by drilling on the graph

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

3.3. The different types of graphical representation of data in OBIEE is of higher standard than QLIKVIEW

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

3.4. In OBIEE the interactive visualisation of data is an outstanding feature.

1 2 3 4 5

( ) ( ) ( ) ( ) ( )

4. OBIEE Reporting

4.1 Reporting/downloading of data from OBIEE meets my requirements.

1 2 3 4 5

	( )	( )	( )	( )	( )

4.2. The report format options on OBIEE are adequate.

1 2 3 4 5

	( )	( )	( )	( )	( )

4.3. The reports downloaded for OBIEE are usable and easy to manipulate.

1 2 3 4 5

	( )	( )	( )	( )	( )

4.4. OBIEE facilitates ad-hoc querying of data.

1 2 3 4 5

	( )	( )	( )	( )	( )

4.5. OBIEEs search-based BI is functionality is adequate.

1 2 3 4 5

	( )	( )	( )	( )	( )

Section B

For each of the following statements, please rate your level of agreement regard to the Procurement Dashboard on QLIKVIEW, on a scale from 1 to 5.

1 QLIKVIEW Navigation

1.1. The QLIKVIEW tool is easy to use. *

1 2 3 4 5

Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.2. The QLIKVIEW tool presents the procurement data in an understandable manner. *

1 2 3 4 5

Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.3. QlikView tool facilitates data easy discovery. *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.4. QLIKVIEW tool maintains filters from one dashboard to another. *

Note : Filters are the selections made for which data you would like to see

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

1.5. The procurement dashboard allows you to move from one subject area to another effortlessly in order to find information you looking for in QLIKVIEW *

	1	2	3	4	5	
Strongly Disagree	( )	( )	( )	( )	( )	Strongly Agree

2. QLIKVIEW Performance

2.1. The QLIKVIEW tool is more responsive to every mouse click than OBIEE.

	1	2	3	4	5
	( )	( )	( )	( )	( )

2.2. QLIKVIEW responds quickly when drilling to further details.

	1	2	3	4	5
	( )	( )	( )	( )	( )

3. QLIKVIEW Data Visualisation

3.1. The graphs in QLIKVIEW help in the further understanding the procurement data.

1 2 3 4 5

	()	()	()	()	()

3.2. The QLIKVIEW tool enables further data discovery by drilling on the graph

1 2 3 4 5

	()	()	()	()	()

3.3. The different types of graphical representation of data in QLIKVIEW is of higher standard than OBIEE

1 2 3 4 5

	()	()	()	()	()

3.4. In QLIKVIEW the interactive visualisation of data is an outstanding feature.

1 2 3 4 5

	()	()	()	()	()

4. QLIKVIEW Reporting

4.1 Reporting/downloading of data from QLIKVIEW meets my requirements.

1 2 3 4 5

	()	()	()	()	()

4.2. The report format options on QLIKVIEW are adequate.

1 2 3 4 5

	()	()	()	()	()

4.3. The reports downloaded for QLIKVIEW are usable and easy to manipulate.

1 2 3 4 5

	()	()	()	()	()

4.4. QLIKVIEW facilitates ad-hoc querying of data

1 2 3 4 5

() () () () ()

4.5. QLIKVIEWs search-based BI is functionality is adequate.

1 2 3 4 5

() () () () ()

4.6. QLIKVIEW's integration to Microsoft office is satisfactory.

1 2 3 4 5

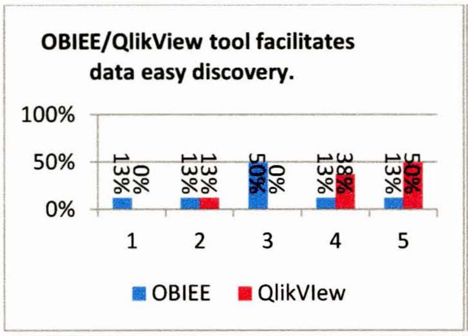
() () () () ()

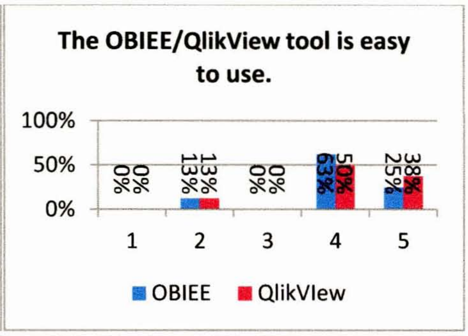
[Submit]

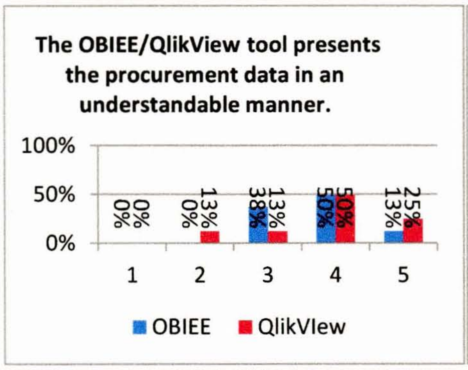
APPENDIX G

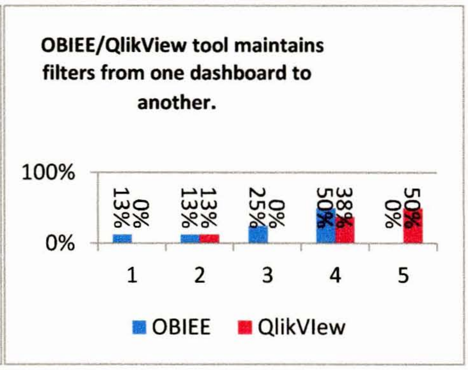
The graphs below represent the results of the survey questions in appendix F for each of the categories (navigational, performance, data visualisation, reporting). The charts display the percentage of respondents in the five categories that range from strongly disagree (1) to strongly agree (5). A summary is presented for each category for the results on the respondents that agreed and strongly agreed, comparing the results of QlikView and OBIEE.

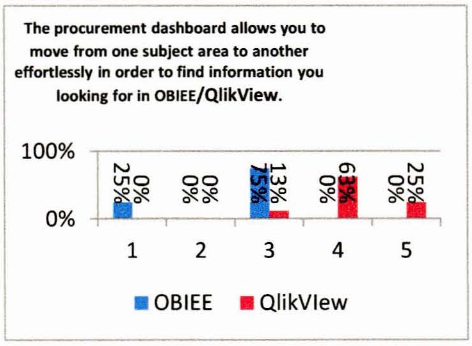
NAVIGATION ANALYSIS









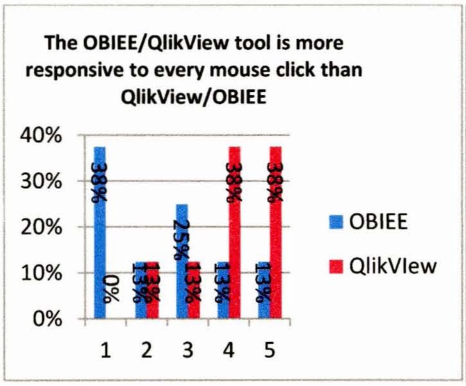


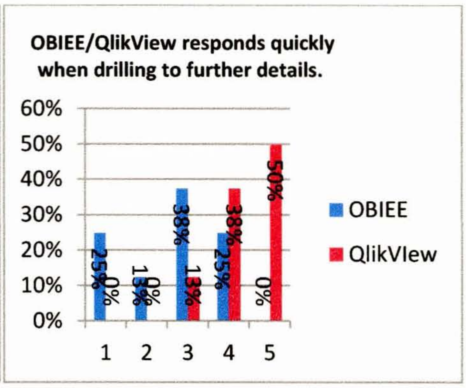
The percentage of users that 'agreed' or 'strongly agreed' on the navigation criteria are summarised below:

	OBIEE	QlikView
Data Discovery	88%	88%
Ease of User	63%	75%
Data Presentation	26%	88%
Maintaining Filters	50%	88%
Ease of navigation	0	88%

The results indicate that although users experience on the OBIEE tool is good; a higher percentage of users preferred the QlikView tool.

PERFORMANCE ANALYSIS



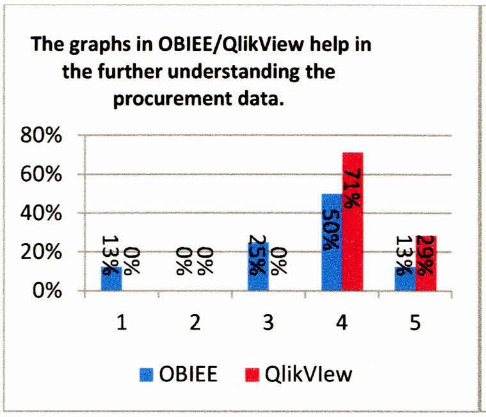


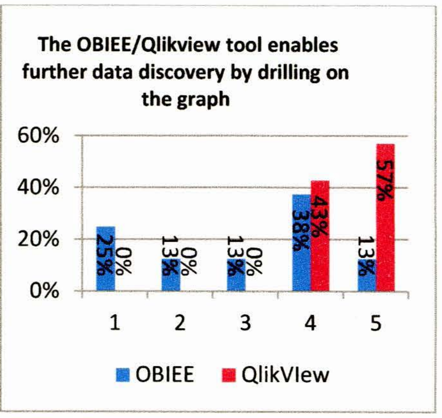
The percentage of users that 'agreed' or 'strongly agreed' on the performance criteria are summarised below:

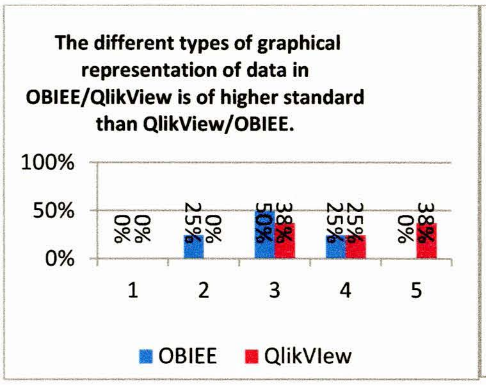
	OBIEE	QlikView
Responsiveness	26%	76%
Detailed Data	25%	88%

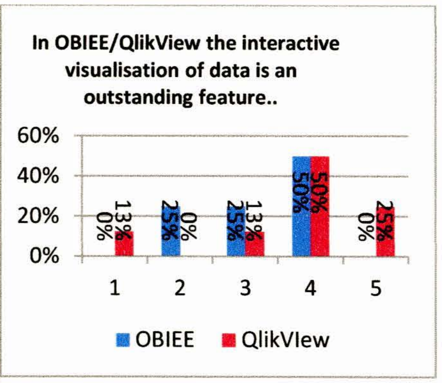
The results indicate that in terms of performance the user experience using Qlikview exceeds that of the OBIEE significantly.

VISUALISATION ANALYSIS







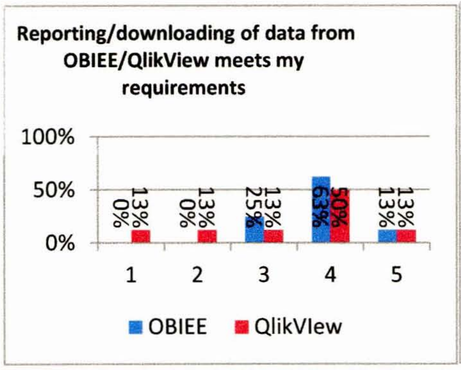


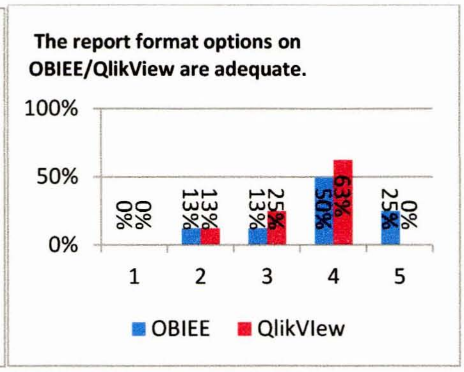
The percentage of users that 'agreed' or 'strongly agreed' on the visualisation criteria are summarised below:

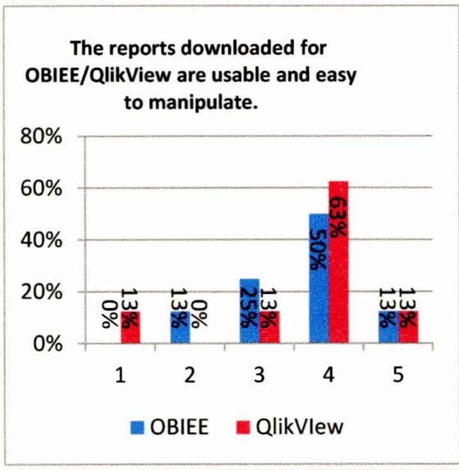
	OBIEE	**QlikView**
Better Understanding	**63%**	**100%**
Drilling to Detail	**51%**	**100%**
Graphical representation	**25%**	**63%**
Interactive Charts	**50%**	**55%**

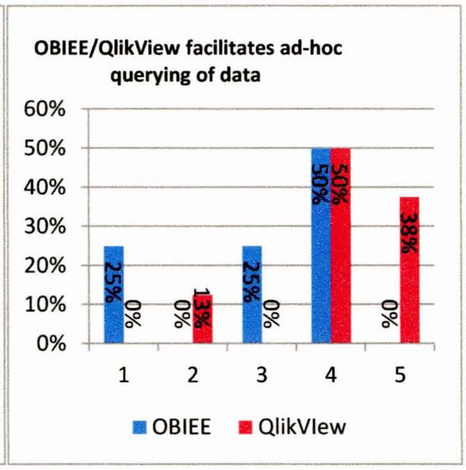
The results indicate that in terms of data visualisation the users strongly preferred the QlikView tool for its ability to drill to details from a summarised view of the data and the data discovery aspect of the tool.

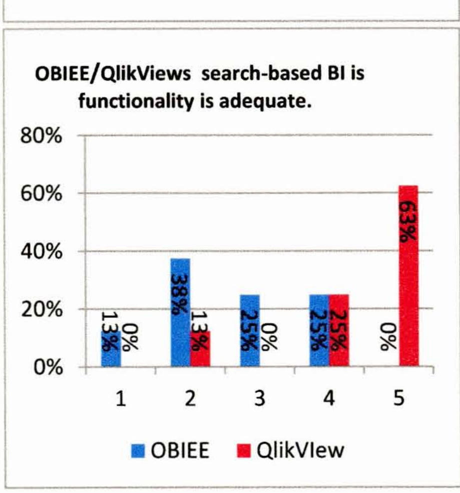
REPORTING ANALYSIS











The percentage of users that 'agreed' or 'strongly agreed' on the reporting criteria are summarised below:

	OBIEE	**QlikView**
Reporting/ Downloading	**76%**	**63%**
Formatting	**75%**	**63%**
Manipulating downloaded data	**50%**	**78%**
Ad-hoc querying/reporting	**50%**	**55%**
Search based feature	**25%**	**88%**

The results indicate that in terms of reporting criteria the users preferred the standard formatted reports from the OBIEE tool over the QlikView tool. However, the search feature of the QlikView tool was strongly preferred by the users

APPENDIX H

USER LIST

Qlikview-OBIEE User Survey - WITS Procurement User List

First Name	Surname	E-mail
`	Mcneill	Lorraine.Mcneill@wits.ac.za
Cornelia	Laubscher	Cornelia.Laubscher@wits.ac.za
Bonolo	Mpshe	Bonolo.Mpshe@wits.ac.za
Sharon	Pillay	Sharon.Pillay@wits.ac.za
Lettah	Maphoto	Lettah.Maphoto@wits.ac.za
Lillian	Robinson	Lillian.Robinson@wits.ac.za
Belinda	Harry	Belinda.Harry@wits.ac.za
Lelau	Ramphisa	Lelau.Ramphisa@wits.ac.za
Innocent	Mamvura	Innocent.Mamvura@wits.ac.za
Emmanuel	Mashandudze	Emmanuel.Mashandudze@wits.ac.za
Amer	Nazir	Amer.Nazir@wits.ac.za
Zarina	Hassim	Zarina.Hassim@wits.ac.za