IJCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-7, Issue-8, Aug 2019 E-ISSN: 2347-2693

Business Analytics Architecture Stack to Modern Business Organizations

Palanivel K[1*], Manikandan J[2]

[1]Computer Centre, Pondicherry University, Puducherry-605014
[2]Department of Computer Science, Pondicherry University, Puducherry-605014

Corresponding Author: kpalanivel.cce@pondiuni.edu.in

DOI: https://doi.org/10.26438/ijcse/v7i8.275287 | Available online at: www.ijcseonline.org

Accepted: 13/Aug/2019, Published: 31/Aug/2019

Abstract— Business Analytics is a set of techniques and processes that can be used to analyze data to improve business performance through fact-based decision-making. Business analytical applications are designed to retrieve, analyze, transform and report data for business intelligence. These business analytics applications give the organization a complete overview of the company to provide key insights and understanding of the business. So smarter decisions may be made regarding business operations, customer conversions and more. A business analytics architecture is used to build business analytical applications for reporting and data analytics. The existing business analytics architecture is designed with Data Warehouse and the data flow among various business components is unidirectional. Today, Data Lake offers an optimal foundation for modern business analytics. In the literature survey, no business analytics architecture is available with Data Lake solutions. Hence, it is proposed to design a business analytics architecture with Data Lake to meet the needs of modern business organizations. The proposed business analytics architecture supports all standardized business analytic reports with Big Data analysis. The proposed business analytics architecture provides many advantages when it comes to scalability, speed, data quality, and flexibility.

Keywords: - *Business analytics, business intelligence, business environment, business architecture, Data Lake.*

I. INTRODUCTION

In the digital era, data is the lifeblood of businesses. Data piling up from various customer have to be efficiently collected and managed for a business to thrive and prosper [15]. In short, enterprise data management (EDM) affects all core business functions like Human Resource (HR), Customer Relationship Management (CRM), Entrepreneurship Resource Planning (ERP), or supply chain management (SCM). The data strategy and its subordinate activities in businesses are mitigating risks, improving data quality, and streamlining business processes while reducing operating cost, developing and executing advanced analytics for business gain, leveraging and monetizing data assets, complying with regulatory policies, preventing data breaches or cyberattacks, and enabling new products or services. Thus, collectively, an organization's data architecture and data strategy play key roles in running the business efficiently.

The world has come a long way since the early days of data analysis where a simple relational database, point-in-time data, and some internal spreadsheet expertise helped to drive business decisions [17]. Today, enterprises focus significant resources to tap into the enormous promise of artificial intelligence (AI), machine learning (ML) and data analytics to drive disruptive innovation and transform their businesses. Data is at the core of how these modern enterprises are unlocking the potential of AI to change their business. The current business environment is constantly evolving. The global economic scenario is providing opportunities as well as challenges. The factors affecting the business environment are consumer needs, globalization, and government policies, etc. In such a business environment, the organization can be reactive, anticipative, adaptive, or/and proactive. For this, a business organization can develop a new strategy; get into a partnership, etc. Today most of the businesses are having computerized business support in the form of decision support system, business analysis, etc.

Business intelligence framework is designed to support the decision-making process. This framework combines architecture, database, analytical tools and applications. Business intelligence helps the organization achieve commercial success along with sound financial management. The main components of business intelligence are data warehouse (DWH), business analytics (BA) and business performance management (BPM) and user interface. Business analytics forms an integral part of business intelligence. Data warehouse holds data obtained from internal sources as well as external sources.

Business analytics has many practical applications. It differs from business intelligence that involves communicating results of data analysis through statistical operations, the formation of predictive models, and application of optimization techniques [2]. This sector helps private banks compete successfully through data mining; it helps credit card companies ascertain the risk of applicants, and it helps insurance companies estimate risk. In modern business organizations, analytics has become an invaluable player in helping companies plan and predict future performance. Business analytics architecture refers to the applications, infrastructures, tools, and leading practices that enable access to and analysis of information to optimize business decisions and performance as shown in Fig. 1. Creating a sustainable business analytics architecture depends on understanding the different components that are involved with developing successful business intelligence tools. Business analytics creates a report as and when required through queries and rules. This business performance is then broadcasted to an executive decision-making body through dashboards and share-point.

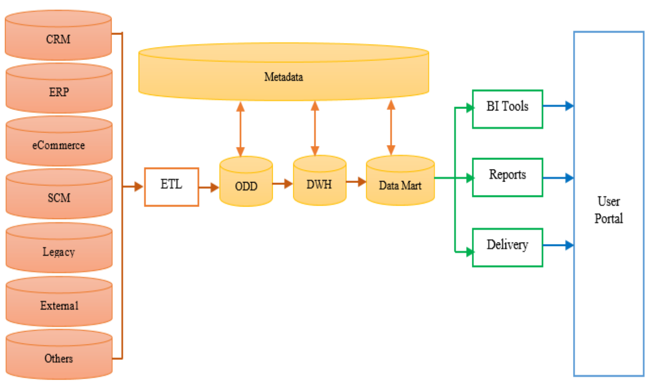

Fig.1. The traditional business analytics architecture

The traditional business analytics architecture defines how business data is acquired, stored, processed, distributed and consumed. Business challenges in supporting large numbers of business users at business speed and its lack of a business user interface require a new kind of data architecture. The traditional business analytics architecture has the challenges of inflexibility, complex architecture, slow performance, old technology and lack of governance.

A. Motivation

Some of the motivations behind to propose a business analytics architecture are discussed below.

Today business requires great data storage and runs complex transformations on massive amounts of data make it less suitable for interactive business analytics. The ability to manage large and complex sets of data has not diminished the appetite for more size and greater speed. The rapid growth in data volumes from a wide range of new sources indeed offers a disruptive opportunity to those who can put it to use. There is a change in mindset among IT organizations and data architects, who now look to capture all data, keep it longer, and prepare to use the data in new ways as business conditions evolve. These changes create dramatic pressure on traditional data architectures, which were built to support structured data with modest growth.

Real-time data analytics is not just a process for storing a large volume of data in a data warehouse, it is about the ability to make better decisions and take meaningful actions at the right time. It is about detecting fraud while someone is swiping a credit card, or triggering an offer while a shopper is standing on a checkout line, or placing an ad on a website while someone is reading a specific article. It is about combining and analyzing data so you can take the right action, at the right time, and the right place.

The motivation behind to propose a new business analytics architecture is shown in Fig. 2.

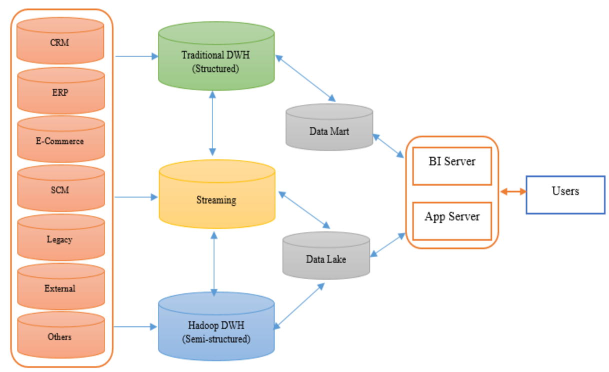

Fig.2. The motivated business analytics model

Today, a rapidly emerging universe of newer technologies has dramatically reduced data processing cycle time, making it possible to explore and experiment with data in ways that would not have been practical or even possible a few years ago. Despite the availability of new tools and systems for handling massive amounts of data at incredible speeds, however, the real promise of advanced data analytics lies beyond the realm of pure technology. Every day it seems that a new technique or application is introduced that pushes the edges of the speed-size envelope even further. Today, the modern data platform delivers on all the requirements for the next-generation data warehouse. Enabling business organizations to radically simplify their existing legacy or overly complex solutions to lower running cost improve agility and deliver real business value.

B. Proposed Solution

The traditional business analytics architecture has analytical processing first pass through a data warehouse. Hence, the objective of this article is to design a business analytics architecture stack to modern business organizations. The goal is to design and build a modern business analytics architecture that provides a flexible, multi-faceted analytical ecosystem for each unique business organization. In the new, modern business analytics architecture, data reaches users through a multiplicity of organization data structures, each tailored to the type of content it contains and the type of user who wants to consume it. The modern business analytics architecture can analyze large volumes and new sources of data and is a significantly better platform for data alignment, consistency and flexible predictive analytics. Thus, the new BI architecture provides a modern analytical ecosystem featuring both top-down and bottom-up data flows that meet all requirements for reporting and analysis.

This research article is organized as follows. The *Background section* describes the details that required writing this research article. The *Literature Review* section reviews various research papers on business analytics. The *Proposed Architecture* section describes the proposal that consists of requirements, business analytics model and the proposed business analytics architecture. Finally, the *conclusion* section concludes this paper and outline the future works.

II. BACKGROUND

This part gives the background details that required writing this research article. This includes an introduction to business intelligence, business analytics, business analytics platforms, business frameworks and architecture.

Every business has to collect, store, organize, and process vast amounts of inflowing raw data before that data can transform into usable information. Raw data has limited value to businesses, while "information" or insights" flashed through marketing dashboards have tremendous benefits for managers and staff. Thus, data performs some defensive actions when it shields itself from breaches and corruption, and some offensive actions when it delivers actionable insights or increased revenue [15]. It needs an enterprise business intelligence or analytics platform that delivers the speed, self-service, and agility front-line business workers demand, and the scale, security, and control to meet rigorous corporate data standards.

A. Business Intelligence

Business Intelligence (BI) is a set of processes, architectures, and technologies that convert raw data into meaningful information that drives profitable business actions [8]. It is a suite of software and services to transform data into actionable intelligence and knowledge. BI has a direct impact on an organization's strategic, tactical and operational business decisions. BI supports fact-based decision making using historical data rather than assumptions and gut feeling. BI tools perform data analysis and create reports, summaries, dashboards, maps, graphs, and charts to provide users with detailed intelligence about the nature of the business. BI catalogues the past and makes it available for future insight. There are three types of BI reports - static reports, dynamic reports, and data mining.

- *Static reports* are by far the most common form of BI. Most businesses have summarized standard reports already laid out and printed to assist in managerial decision-making. For example, educational institutions use enrollment reports to gauge which departments might need to hire more faculty.
- *Dynamic reports* look similar to static reports but online and interactive. A business manager curious as to where a certain summary number on his dashboard comes from can drill down to expose the detail that contributed to that number. For example, if sales in one region are down, and then drill down to discover a problem in the other region.
- *Data mining* uses computer programs and statistical analyses to search for unexpected patterns, correlations, trends, and clustering in the data. One often-cited example of data mining was the discovery that beer and diapers are frequently purchased on the same trip to the grocery store. Marketers can use this information to place the two items nearby in the store.

BI technologies provide current, historical and predictive views of internally structured data for products and departments by establishing more effective decision-making and strategic operational insights through functions like online analytical processing (OLAP), reporting, predictive analytics, data/text mining, benchmarking and business performance management (BPM). These technologies and functions are often referred to as *information management*.

B. Business Analytics

Business Analytics (BA) refers to the skills, technologies, practices to gain insights using the collection, organization, analysis, exploration, investigation, interpretation and presentation of business data tools and technologies [1]. Business analytics drives business planning and the objective is to discover new insights and relate business performance to business data. Business analytics provides fact-based, analytics strategies help in the discovery of a meaningful pattern in data and then to communicate insight using the data visualization of the recorded information. Business Analytics is "*the study of data through statistical and operations analysis, the formation of predictive models, application of optimization techniques, and the communication of these results to customers, business partners, and college executives*."

Business Analytics requires quantitative methods and evidence-based data for business modelling and decision-making; as such, business analytics requires the use of Big Data. It is the iterative, methodical exploration of an organization's data, with an emphasis on statistical analysis. Companies that are committed to making data-driven decisions use business analytics. Data-driven companies treat their data as a corporate asset and actively look for ways to turn it into a competitive advantage. Successful business analytics depends on data quality, skilled analysts who understand the technologies and the business, and an organizational commitment to using data to gain insights that inform business decisions.

Business analytics is the crystal ball. Business analytics makes use of existing infrastructure such as data gathered in BI and aims for deeper insight into the data, frequently dealing with predictive analytics about the future. Business analytics makes use of data to provide predictive capabilities. The differences between the two tend to stem from the types of questions each attempt to answer and the tools and processes used to answer those questions.

Specific types of business analytics include descriptive analytics, predictive analytics, and prescriptive analytics. The descriptive analytics, which tracks key performance indicators (KPIs) to understand the present state of a business. The predictive analytics, which analyzes trend data to assess the likelihood of future outcomes. Also, prescriptive analytics, which uses past performance to generate recommendations about how to handle similar situations in the future.

The technology drivers of business analytics are presented in statistical analysis, data mining, quantitative analytics, qualitative data analysis, business intelligence, predictive analytics, text analytics, etc. Business analytics can be implemented in the areas of business including measurement, analytics, reporting, collaboration platforms and knowledge management. The application of business analytics to business data helps to describe, predict, and improve business performance. The use of analytics methods can make organizations better, faster, and more consistent decisions. It also provides the business with the data for the primary source of competitive differentiation.

Business intelligence and business analytics are both disciplines inside of a wider field of being a data professional. They deal with deriving meaning from data.
A useful distinction can be made [22] among descriptive, predictive, and optimization analytics. As the names suggest, the different kinds of analytics describe a situation, make inferences about it, or suggest an optimal solution. The list categorizes the variety of analytics are optimization analytics (mathematical programming e.g., linear, integer, simulation), predictive analytics (decision trees, cart, genetic algorithms, neural networks) and descriptive analytics (data visualization, dashboards/scorecards, drillable/OLAP reports, published reports and SQL queries.

C. Traditional Business Analytics Platforms

The traditional business analytics platforms are providing users with comprehensive historical reporting and user-friendly ad-hoc analysis tools. The availability of this functionality is largely due to the underlying data architecture, which consists of a centralized data storage solution such as an enterprise data warehouse (EDW). EDWs form the backbone of traditional data platforms and often connect an immense web of source systems into a central data repository. Data is then standardized, cleansed, and transformed in the EDW before being pulled into various reports and dashboards to display historical business information. While traditional business analytics offers a basis for these types of dashboards and ad-hoc reporting, this IT-developed solution has presented its unique challenges.
Some of the challenges [16] associated with traditional business analytics solutions are lack of on-demand analysis, need for predictive analyses, analysis of mixed data types, etc. These challenges require an analytics platform and strategy that goes beyond the breadth of traditional BI platforms. In the existing platform, customer tired of logging into multiple tools to see organization performance. It can view all of organization performance data in one place so the customer can spend less time checking data and creating reports and more time acting on insights. Mix and match metrics from different sources in one Databoard and get a more complete view of company performance at a glance. The traditional business intelligence and analytic approaches are constrained by capability, capacity and availability of the organization's resources such as skilled staff, infrastructure and software. Many organizations that lack the people, processes, and technology necessary to expand their data analysis capabilities to the next level become discouraged.
Enterprise data warehouses continue to play a key role in existing data platforms, providing the thoroughly cleansed, organized, and governed data needed for most businesses. However, the challenges associated with traditional business analytics are creating demand for augmenting an EDW with another form of architecture optimized for quick access to ever-changing data.

D. Modern Business Analytics Platforms

Business organizations looking to modernize their analytics platforms have started to adopt the concept of *data lakes*. Data lakes store information in its raw and unfiltered form, be it structured, semi-structured, or unstructured. Data lakes themselves perform very little automated cleansing and transformation of data, allowing data to, therefore, be ingested with greater efficiency, but transferring the larger responsibility of data preparation and analysis to business users. Data lakes offer a low-cost solution for efficiently storing and analyzing many types of data in its native form. A data lake solution coupled with an enterprise data warehouse defines the next generation of business analytics and offers an optimal foundation for data analysis, as shown in Fig. 3.

Fig.3. a data lake solution coupled with an EDW

In the modern business analytics platforms, the EDW receives system data from various sources through an extract-transform-load (ETL)/ extract-load-transform (ELT) process. After being cleansed, standardized, and transformed, the data is ready for analysis by a wide variety of users via reports and dashboards. Meanwhile, the data lake collects raw data from one, many, or all of the source systems, and data is ingested and immediately ready for discovery or analysis. The result is a wider user base exploring and creating relationships between enormous amounts of diverse data for individual analyses, on-demand.

E. Business Analytical Applications

Business analytical applications can be used to support both internal and external business processes. Business applications that are equipped with analytical capabilities allow business users to gain insights into improving the performance of business operations. Business analytical applications provide functionalities such as modelling, forecasting, sales analysis, and what-if scenarios. By employing analytical applications, decision-makers can also identify and understand what factors drive their business value, and thus able to leverage opportunities faster than their competitors [13].

The working principle of business analytical applications is stated below:

i. Data originating from internal and external sources have to be extracted, transformed, and loaded into the data warehouse. The data from the warehouse is then sent to data marts to fulfil specific operational needs.
ii. When the data passes through the ETL/ELT, it can flow to both directions, either to the operational data store (ODS) or to data warehouse directly.
iii. At end user, data in the ODS, data warehouse, data lake and data marts can be accessed by using a variety of tools such as query and reporting tools, data visualization tools, and analytical applications.
iv. Finally, a centralized metadata repository is connected with various components such as the ETL, the data warehouse, and the end-user.

The data flows among the components are multi-directional. This overcomes limitations of unidirectional data flow in many existing business analytics architectures. A multi-directional flow can enhance query performance and improve accuracy because data error at one layer can be returned to the previous layer for clarification if an error occurs.

F. Business Analytics Framework

Business Analytics framework provides a blueprint for success, particularly for small and midsize organizations getting started with BA. A key benefit of the framework is “to surface key decisions, integration points, gaps, overlaps and biases that business leaders may not have otherwise prepared for” when getting started with BA. The framework defines the people, processes and technologies that need to be integrated and aligned to take a strategic approach to business intelligence, analytics and performance management initiatives – particularly when considering an enterprise approach.

Gartner [11] recommends that organizations should use the framework to develop a strategy and an implementation plan and to surface key decisions, integration points, gaps, overlaps and biases that business leaders may not have otherwise prepared for. Gartner also points out, that a portfolio of information management, analytics and decision-making capabilities will be needed to meet the diverse requirements of a large organization. Again, the intent of the framework is simply to illustrate graphically all of the components that must be considered including people, processes, and technology when considering Business Analytics. In addition, there is no single or right instantiation of the framework; the framework based on business objectives and constraints can support different configurations.

The business analytics framework updates Gartner's [11] previous business intelligence, analytics and program management framework. The business framework is to be used by IT architects, system developers and program managers that lays out the components in terms of the people, processes, platforms and performance that should be aligned as part of a strategic solution.

G. Traditional Business Analytics Architecture

Business analytics technologies are mainly used in online analytical processing, data mining, process mining, benchmarking, text mining, ad-hoc queries and predictive analytics. Hence, business analytics architecture plays a vital role in business projects by influencing overall development and implementation decisions. Business analytics architecture seeks to help organizations and businesses make better decisions.

Business analytics architecture refers to the applications, infrastructures, tools, and leading practices that enable access to and analysis of information to optimize business decisions and performance. The main components of business analytics are a data warehouse, business analytics and business performance management and user interface.

i. Data warehouse holds data obtained from internal sources as well as external sources. The internal sources include various operational systems.
ii. Business analytics creates a report as and when required through queries and rules. Data mining is also another important aspect of business analytics.
iii. Business performance management is a linkage of data with business objectives for efficient tracking. This business performance is then broadcasted to an executive

decision-making body through dashboards and share-point.

It is pointed out that several common characteristics [9] and values of enterprise BA architecture. These requirements are fundamental to business intelligence systems that will be deployed broadly across the organization. Some of the qualities are agility, business view, deployability, interoperability, manageability, openness, reliability, scalability, usability, leverages existing infrastructure and security. All of these qualities are delivered largely through the underlying architecture. Therefore, creating a sustainable architecture depends on understanding the different components that are involved with developing successful business intelligence tools.

Business analytics architecture may include both structured and unstructured data, as well as information from both internal and external sources. The data that can be used by business intelligence systems are usually transformed from raw transaction data into logical and lucid information by the information management architecture.

The traditional business analytics architecture has the challenges of inflexibility, complex architecture, slow performance, old technology and lack of governance. This traditional business analytics architecture is depicted in Fig. 4. It has been used for decades now and is still the most dominant architecture for business analytics.

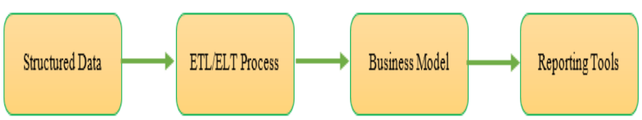

Fig. 4: Traditional business analytics architecture.

The traditional BI architecture comprises of the following:

- an extract-transform-load (ETL) or an extract-load-transform (ELT) tool that periodically integrates data from structured data sources and transforms and reorganizes the data into a Business analytics data model made of dimension and fact tables, suited to efficient reporting
- a data warehouse (DW) that is, essentially, a relational database that stores the data transformed by the ETL/ELT tool into the business analytics data model
- a reporting tool that leverages the data in the DW by using the underlying SQL engine to create dynamic visualizations that are organized into reports and dashboards for business users.

In traditional business analytics architectures, the EDW is a key component in the analytics chain, ingesting data from systems of record, processing it and then feeding key analytic capabilities across data marts, apps and dashboards. As the landscape of data sources and systems has grown, however, several limitations in this approach have become apparent. They are data is stored inexpensive, high-performance data systems across the architecture, relatively low-value ETL/ELT workloads, new types of data, such as clickstream, sensor, and server log data that do not fit predefined schema in the data architecture, EDW are poorly managed, etc.

H. Modern Business Analytics Architecture

The modern data platform delivers on all the requirements for the next-generation data warehouse. Today's modern analytics software requires the ability to power both organization decision making and comprehensive growth. To support modern analytics capabilities, today's analytic software must power the following components of data analysis:

- *Data ingestion***.** It describes the tools and software that collect and store the various types of data and making them available for analytics. Logs and streaming data require different ingestion mechanisms than data residing in a database.
- *Data preparation* or *data wrangling*. It is the cleansing, consolidation, and standardization of data before data analysis that is typically performed in analytics platforms.
- *Data discovery***.** It is used for analyzing patterns and relationships through summary statistics, what-if analysis, and visualizations and is performed in analytics platforms.
- *Advanced analytics***.** It consists of a collection of data analysis techniques that expand beyond historical reporting and trend analysis to gain deeper insights, actionable intelligence, and next steps from diverse sets of data. Methods such as machine learning, artificial intelligence, data and text mining, network/cluster analysis, sentiment analysis, and random forest regressions are changing and shaping the future of modern analytics.

Business analytics service (BAS) connects businesses, platforms, customers, employees, and partners, by integrating data and delivering insights across the global digital ecosystem. Expert consulting, implementation and operational services cover the complete information management lifecycle for customers. These business analytics services leverage deep data science capabilities to deliver actionable insights, resulting in positive business outcomes.

The modern business analytics technology [18] is a true cloud architecture that provides many benefits to businesses, speeding time to value, reducing total cost of ownership and increasing agility. The modern business analytics architecture supports centralized IT teams and decentralized lines of business within large enterprises supporting the flexible demands of thousands of users and petabytes of data. The modern business data architecture provides linear scale storage and compute so that it can scale to meet the needs of not just new data sources but more advanced workloads and data science methods that have emerged.

The EDW became a standard component incorporate data architectures because it provides valuable business insights and powerful decision analytics for front-line workers, executives, business analysts, data scientists, and software developers. For years, the EDW has been the core foundation on which analytics are built within the enterprise. EDW solutions are mature and extremely effective for reporting and data analytics, especially for known transactional data. As new data types and new analytic tools have emerged, however, a new approach for broader data and analytics architecture in the enterprise has become necessary. With the burgeoning volume and variety of these new data sources and types, the centre of gravity for modern data architectures is shifting. This is causing the need not only for new agile tools but also for an integrated ecosystem solution. Table I lists the business analytics platform that can provide services related to data, discovery and deployment.

Table 1: The services of the business analytics platform.

S/N	Business Scenario	Various Stages		
		Data Discovery	Data Processing	Data Deployment
1	Data Analytical	-	-	√
2	Dashboards	√	-	-
3	Data Capture	-	√	-
4	Data Cleansing	-	√	-
5	Data Lake	√	-	-
6	Data Preparation	-	√	-
7	Data Quality	-	√	-
8	Data Transformation	-	√	-
9	Database Analytics	-	-	√
10	Decision Making	-	-	√
11	Forecasting	√	-	-
12	Machine Learning	√	-	-
13	Optimization	√	-	-
14	Reporting	√	-	-
15	Stream Analytics	-	-	√
16	Text, Audio & Video Analytics	√	-	-
17	Visualization	√	√	√
18	Workflow	-	-	√

The benefits of modern business analytics architecture include:

i. move rarely used data to the modern data platforms and access it on-demand, saving on overall storage costs
ii. store more data longer to enhance analytics with deeper information providing better results
iii. store and process new data sources and feed transformed data into your EDW to augment or create wholly new analytic value
iv. onboard ETL/ETL processes in the modern data platforms to take advantage of computing and operational efficiencies

Today, businesses require analytics-driven insights and real-time dashboards to run efficiently. Predictive analytics enable successful companies to compete and win. All the while, driven by new competitive pressures, new data sources, and budget constraints, leading organizations are augmenting traditional data architectures with Hadoop [20] as a way to extend capabilities with new sets of data and analytic applications while containing or reducing costs.

Hence, it is proposed to devise an architecture that answers the challenges of cost, complexity, and expansion, organizations are turning to modernize their data architectures and at the same time enhance the value of their existing EDW implementations.

III. LITERATURE REVIEW

This section reviewed some business analytics architectures to propose new business analytics for today business organizations.

Business Analytic Applications are complete, prebuilt BI solutions that deliver intuitive, role-based intelligence for everyone in an organization from front line employees to senior management that enable better decisions, actions, and business processes. They are designed for heterogeneous environments; these solutions enable organizations to gain insight from a range of data sources and applications including Siebel, Oracle E-Business Suite, PeopleSoft, and third-party systems such as SAP.

Some of the business architectures reviewed in this article are discussed below.

The Oracle Business Analytics Warehouse (OBAW) [14] architecture comprises the client, server and repository components. The server executes the instructions from the client. The server manages data warehouse processes, including loading of the ETL and scheduling execution plans. It dynamically adjusts its actions based on information in the repository. The repository stores the metadata that represents the data warehouse processes.

The Pentaho Stack [6] consisted of a presentation layer, data layer and server layer elements. The data available in the presentation layer is viewed as reporting, analysis, dashboards and process management. The data layer can be used to connect to any data sources. The server layer is the middle layer in BA Architecture. The application runs on a server and serves as a middle layer. It can deploy the report and dashboards and make it available to the end-user.

The emerging architecture [10] consists of a data source layer, analytics layer, integration layer and data visualize layer. It is important to note that each layer is associated with different sets of users and that different sets of users will define "real-time" differently. Each layer enables a critical phase of real-time analytics deployment. Table II compares various existing business analytics architecture.

Table II: The services of the business analytics platform.

S/N	Business Scenario	OBAW[14]	Pentaho Stack[6]	Emerging Architecture [10]
1	Client Service	√	√	√
2	Server Service	√	√	√
3	Storage Service	√	√	√
4	Integration Service	×	×	√
5	Any kind of Data Source	√	√	√
6	Reporting	√	√	√
7	Dashboard	√	√	√
8	Process Management	×	√	×
9	Mobile Apps & Alerts	×	×	×
10	Visual Analytics	×	×	×
11	Real-time	×	×	√
12	Data Mart	×	×	×
13	Data Lake	×	×	×
14	Machine Learning	×	×	×

Traditional BA architectures usually rely on single-server row-based RDBMSs as the technologies used by their DWH [19]. These work well enough with small amounts of data i.e., less than few Terra Bytes, but with larger data sets these technologies struggle. Some new developments have thrived in recent years and new database technologies have appeared while traditional RDBMSs have evolved to include: columnar storage formats, compressed formats, in-memory processing and massively parallel processing. Therefore, in BA architectures leveraging Big Data, also, it is common to replace the underlying RDBMS technology of the DWH with an enhanced solution that leverages the above-mentioned optimizations. In some cases, an improved database with the above-mentioned enhancements is used alongside the DWH and acts as an accelerator for BA queries.

Today, it is essential to consider a lot of business analytics applications that can be built into an architecture: reporting, ad-hoc query, data mining and data visualization tools, online analytical processing (OLAP) software, dashboards and performance scorecards.

- *Ad hoc report/query:* It enables users to ask their own questions of the data, without relying on IT to create a report. In particular, the tools must have a reusable semantic layer to enable users to navigate available data sources, predefined metrics, and hierarchies and so on.
- *Dashboards:* A style of reporting that graphically depicts performances measures. It includes the ability to publish multi-object, linked reports and parameters with intuitive and interactive displays. Dashboards often employ visualization components such as gauges, sliders, checkboxes and maps, and are often used to show the actual value of the measure compared to a goal or target value. Dashboards can represent operational or strategic information.
- *Excel integration:* Often Excel acts as the reporting or analytics tool. Various BI tools provide integration with Microsoft Excel, including support for native document and presentation formats, formulas, charts, data “refreshes” and pivot tables. Advanced integration includes cell locking and write-back.
- *Mobile Apps and Alerts:* It enables organizations to develop and deliver content to mobile devices in a publishing and/or interactive mode, and takes advantage of mobile devices' native capabilities, such as a touchscreen, camera, location awareness and natural-language query.
- Reporting: It provides the ability to create highly formatted, print-ready and interactive reports, with or without parameters.
- *Visual Analytics:* Before visual analytics, organization users relied on data analysts to build reports. Visual Analytics can also take on big data. Staff analyze data on their own to build accurate reports for management, saving significant time and money. Typical visual analytics platforms include SAS, Tableau Software, Qlikview and Tibco Spotfire.

Modern network analytics architecture stack to enterprise networks [21] offered a breadth of analytics solutions architecture that enables network providers to gradually adopt analytics. This had been designed with a set of architecture principles that are commonly accepted as best practices in the network industry. This architecture believed that it can improve agility, while at the same time having control over data integration and distribution.

Form the literature, the technology components are used to present information to business users and enable them to analyze the data. This includes the business analytics suite or tools to be used within an organization as well as the supporting IT infrastructure.

IV. PROPOSED ARCHITECTURE

The primary focus of this article is to design a business analytics architecture for modern business organizations. This includes studying the requirements, quality attributes, proposed business analytics model and architecture and some use cases. This section presents the design principles considered in architectural design. It also provides a detailed description of its main components.

A. Requirements and Goals

The business analytics architecture is an umbrella term for an enterprise-wide set of systems, applications, and governance processes that enable sophisticated analytics, by allowing data, content and analyses to flow to those who need them when they need them. The BA architecture must be able to quickly provide users with reliable, accurate information and help them make decisions of widely varying complexity. It also must make information available through a variety of distribution channels, including traditional reports, ad hoc analysis tools, corporate dashboards, spreadsheets, e-mail, and pager alerts.

The things considered when identifying the analytics architecture are sources of data, diverse sets of data and structure formats, latency of data, volume of data, data delivery modes, multi-platform data architecture, agile delivery method, use of traditional platforms, data virtualization techniques, well-defined ETL/ ELT, analytics use, support for all types of users, flexibly and quickly, adapt to change, etc.

The goal of business analytics and intelligence software is to help businesses access, analyze and visualize data, and then communicate those insights in meaningful dashboards and metrics. Unfortunately, the reality is that the majority of software options on the market today provide only a subset of that functionality. In addition, those that provide a more comprehensive solution, tend to lack the features that make it user-friendly. The modern analytical requirement needs an architecture mix of traditional approach for sustainability and ability to cater for a fast-paced business depends on data analytics, which is, must for business to survive and become a Data-driven decision-making organization.

B. Qualities of the Architecture

Regardless of the technique used to elicit the requirements, the desired qualities of the business system to be constructed determine the shape of its structure. The architectural qualities considered the proposed architecture are listed below:

- *Agility***.** The proposed modern business analytics architecture adapt accordingly.
- *Business view***.** The proposed modern business analytics architecture maintains a single view to ensure user confidence.
- *Deployability*. By making a few changes, the proposed modern business analytics architecture is easily deployed.
- *Interoperability***.** The proposed modern business analytics architecture has a single interface for all business analytics capabilities with the ability to navigate through scorecards, dashboards or reports.
- *Leverage existing infrastructure.* The proposed modern business analytics architecture is designed to support existing environments and leverage everything those environments have to offer Web infrastructure, databases and OLAP data sources, security providers, application servers and more.
- *Manageability.* The proposed modern business analytics architecture can administer efficiently and proactively ensuring that potential problems are identified early and avoided, thus keeping the system operating effectively.
- *Openness.* The proposed modern business analytics architecture support open in terms of the data that can access and for integration with existing and new applications, portals, security systems and more.
- *Reliability.* The proposed modern business analytics architecture supports operating on a 24x7 basis with redundancy for all capabilities and services.
- *Scalability***.** The proposed modern business analytics architecture linearly supports scalability to thousands and millions of users across a global organization.
- *Searchability:* The proposed modern business analytics architecture is highly searchable so that business users can leverage business analytics information that the business organization has already created.
- *Security.* The proposed modern business analytics architecture supports with existing security providers to ensure that access to both the business analytics system and the information in that system is always secured as required.
- *Usability.* The proposed modern business analytics architecture recognizes and accommodates different types of business users through common user experience, across all business analytics capabilities and on the full range of technology, including mobile devices.

The above a set of attributes or qualities are applicable across the components of the proposed architecture. Some qualities are easier than others to describe in terms of standards.

C. Proposed Business Analytics Model

With the above requirements and qualities, the proposed business analytics model is presented in Fig. 5.

In the proposed business analytics model, most source data now flows through data ingestion tools, which primarily acts as a staging area and online archive. This is especially true for semi-structured data, such as log files and machine-generated data, but also for some structured data that cannot be cost-effectively stored and processed in SQL engines. From data ingestion tools, data is fed into a data warehousing, which often distributes data to downstream systems, such as data marts, data lakes, operational data stores, and analytical sandboxes of various types, where users can query the data using familiar SQL-based reporting and analysis tools.

The data analysts and scientists can analyze raw data inside data ingestion tools by writing programs using programming languages. The business users can able to query and process data using familiar SQL-based data integration and query tools.

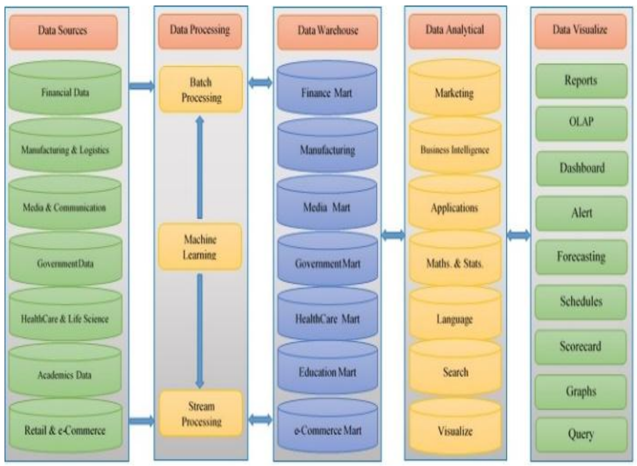

Fig.5. Proposed business analytics model.

With this proposed model, the business analytics architecture is designed to accommodate virtually all data sources. The proposed business analytics architecture and data governance process combine to enable truly integrated reporting, while also remaining flexible and scalable.

D. Proposed Business Analytics Architecture

In the modern business analytics architecture, information is sourced from various systems and tools before undergoing a detailed process to cleanse, conform, consolidate, and audit the data. A single source of common data is then used across multiple systems to establish integrated reports and visual dashboards. These dashboards become the new starting point to investigate, identify, analyze, and answer business-critical decisions.

The modern business analytics architecture creates analytical sandboxes that let power users explore corporate and local data on their own terms. These sandboxes and specialized analytical databases that offload data or analytical processing from the data warehouse or handle new untapped sources of data, such as Weblogs or machine data. The new environment also gives department heads the ability to create and consume dashboards built with in-memory visualization tools that point to both a corporate data warehouse and other independent sources.

The proposed modern business analytics architecture shown in Fig.6 has data source layer, ETL Layer, data warehouse layer, data analytics layer and data visualization or end-user layer.

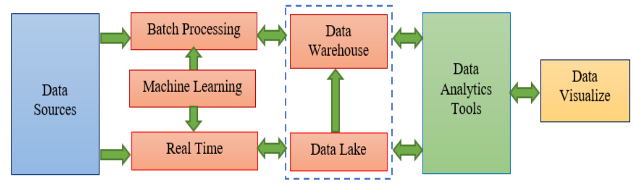

Fig.6. Proposed business analytics architecture.

Data Source Layer. Nowadays, many business applications require the use of structured data, unstructured and semi-structured data to make an effective and timely decision. All these data can be acquired from two types of sources: internal and external. *Internal data sources* include data related to business operations (i.e., customers, products, and sales data). These operating systems are also known as online transaction processing systems because they process a large number of transactions in real-time and update data whenever it is needed. Operational systems contain only current data that is used to support the daily business operations of an organization. *External data source* refers to those that originate outside an organization. This type of data can be collected from external sources such as business partners, syndicate data suppliers, the Internet, governments, and market research organizations. These data are often related to competitors, market, environment (e.g., customer demographics and economic), and technology.
Business organizations need to identify their data sources. Knowing where the required data can be obtained is useful in addressing specific business questions and requirements, thereby resulting in significant time savings and greater speed of information delivery.

ETL/ELT Layer. The ETL/ELT layer focuses on three main processes: extraction, transformation and loading.

- *Extraction* is the process of identifying and collecting relevant data from different sources. Usually, the collected data are not integrated, incomplete, and may be duplicated. Therefore, the extraction process is needed to select data that are significant in supporting organizational decision-making. The extracted data are then sent to the transformation and cleansing process. After that, the data will go through the transformation and the cleansing process.
- Transformation is the process of converting data using a set of business rules into consistent formats for reporting and analysis. Data transformation process also includes defining business logic for data mapping and standardizing data definitions to ensure consistency across an organization. As for data cleansing, it refers to the process of identifying and correcting data errors based on pre-specified rules. If there is an error found on the extracted data, then it is sent back to the data source for correction.

- Once data have been transformed and cleansed, they are loaded into a target repository. Loading is the last phase of the ETL process.

Data Warehouse Layer. The business analytics warehouse is a unified data repository for all customer-centric data. The purpose of the business analytics warehouse is to support the analytical requirements of BI applications. The business analytics warehouse includes data integration engine, an architecture, data extractors, set of ETL and centralized console. A data integration engine that combines data from multiple source systems to build a data marts in a warehouse. An architecture to allow organizations to use third-party analytical tools in conjunction with the business analytics warehouse using the BI server. Prebuilt data extractors to incorporate data from external applications into the business analytics warehouse. A set of ETL (extract-transform-load) processes takes data from multiple source systems and creates the business analytics warehouse tables. It includes a centralized console for administration, loading, and monitoring of the business analytics warehouse. There are three components in the data warehouse layer, namely operational data store, data warehouse, and data marts.

- *Operational Data Store (ODS)*. An ODS is used to integrate all data from the ETL/ELT layer and load them into data warehouses. It provides an integrated view of near real-time data such as transactions and prices. The data stored in ODS are updated frequently.
- *Data Warehouse.* A data warehouse is a central storage that collects and stores data from internal and external sources for strategic decision-making, queries, and analysis. Data are stored in a data warehouse longer than in ODS. Data in a data warehouse is updated regularly. Data warehouses are designed to support OLAP applications by storing and maintaining data in multi-dimensional structures for a query, reporting, and analysis.
- *Data Mart.* A data mart is used to support the analytical needs of a particular business function or department. There can be many data marts inside an organization. It contains historical data that can help users to access and analyze different data trends. The amount of data stored in a data mart is much lesser than the data stored in a data warehouse.

Data warehouses and data marts are built based on the multi-dimensional data model, which consists of fact and dimension tables. Data flows from operational data store to data warehouse and subsequently to data mart.

Data Visualization Layer. The data visualization layer is also called the end-user layer consists of tools that display information in different formats to different users. These tools can be query and reporting tools that are used mostly by the operational management level.

- *Query and reporting tools.* Query and reporting tools are very useful tools, which allow business users to access and query data quickly, and to produce reports for decision-making and management purposes. There are many different types of reports including standard reports, ad-hoc reports, budgeting and planning reports, and metadata reports.
- *OLAP Tools. OLAP* servers can manage data in the data warehouse layer for reporting, analysis, modelling, and planning to optimize business. OLAP server is a data manipulation engine that is designed to support multidimensional data structures. OLAP is a user-friendly graphical tool that allows users to quickly view and analyze business data from different perspectives. Besides that, OLAP also allows users to easily compare different types of data and complex computations.
- *Data Mining.* Data mining process can be achieved with the integration of data warehouses and OLAP servers by performing further data analysis in OLAP cubes. Since the amount of data in an organization is growing rapidly, it is necessary to have data mining to make decisions faster. Data mining techniques have been used to predict future results and summarize details of data.
- *Data Visualization Tools.* Data visualization tools such as dashboard and scorecards can be provided to business managers and executives who need an overall view of their business performance. Business users can also view more detailed information about key performance indicators across their organizations. Business managers can effectively monitor their business performance and progress toward defined goals.

Governance is an internal body that helps organizations oversee changes to analytics solutions and processes, resolve analytics/data issues, and facilitate decision making amongst agency stakeholders. The governance body helps prioritize data sets to be ingested into the data lake, defines best practices for performing analyses and creating efficient self-service data sets, and sets the criteria for publishing data sets for other users. As higher volumes of data are ingested into the data lake, the risk of misinformation and incomplete or undefined data grows, reducing the overall usefulness of the data stored, and ultimately the quality of any downstream analyses produced. This is where metadata management comes in to play.

Effective metadata management not only builds trust through clearly identified data but also enables shared knowledge of how data is defined and related, expediting future analyses. Security also plays a key role in the development and proper use of a data lake solution.

The identity management and authentication systems are key to controlling access to content stored in the data lake. Through these processes, agencies can increase the consistency with which users can locate and trust data, increasing user adoption and trust.

The proposed architecture is a fully multi-tenant solution from both a data processing and data storage perspective. The business users are spread across an infinitely scalable pool of computing resources, leveraging its shared-nothing architecture. The proposed analytics architecture solution provides scalability. The data visualizations in the dashboards are built for end-user performance and remove additional steps in the load process. Multiple queries are sent simultaneously from dashboards, whereas most other products send queries sequentially. This ensures business organizations frequently processing data, can interact with data and make smarter decisions at all times.

This study is qualitative in nature. Qualitative strategies are commonly associated with inductive approaches to do research, whereas deductive approaches usually consist of more quantitative strategies. In inductive approaches, the theory is the outcome of the research. This study has chosen an abductive approach based on systematic combining of theory, framework, and the empirical world. An initial theoretical framework directs the empirical findings, and the empirical results generate a need to modify the theoretical model. Hence, the theoretical architecture is developed and refined concurrently.

E. Use Cases

This research article was conducted as a case study, which is an empirical investigation of a phenomenon within its real context, with the use of multiple sources of evidence. It involves detailed and intensive analysis of a single case, to capture its complexity and particularities. To perform empirical case studies of Business Analytics, this study also deems it as an appropriate method for investigating how business analytics works in, and how it can be developed through, practical applications.

Fuse by Cardinal Health [7] is an innovation lab focused on improving the future of health and wellness by making healthcare safer and more cost-effective. The Fuse team focuses on connected care, building a smarter supply chain, and discovering new insights through analytics. Fuse chose the Hortonworks Data Platform to optimize its data architecture and enrich its existing data with freely available public datasets.

Hortonworks data platform [4] is a leading innovator at creating, distributing and supporting enterprise-ready open data platforms. The mission is to manage the world's data. It has a single-minded focus on driving innovation in open source communities such as Apache Hadoop, NiFi, and Spark. The open connected data platforms power modern data applications that deliver actionable intelligence from all data: data-in-motion and data-at-rest. It provides the expertise, training and services to the customers.

Neustar [4], a telecommunications information and analytics provider used to capture less than 10% of its network data and retain it for 60 days. With Hortonworks Data Platform, Neustar now captures 100% of the network data and retains it for two years. That means 150 times more storage while saving millions.

The Symantec data platform [4] helps consumers and organizations secure and manage their information-driven world. The Symantec Cloud Platform team turned to Hortonworks and HDP to help speed the rate with which it could ingest and process 500,000 security log messages per second (40 billion messages per day). Using HDP in the cloud, the team reduced its average time to analysis from four hours to two seconds.

TRUECar’s mission [4] is to make the car buying process simple, fair and fun. The company looked to Hadoop and HDP to build a modern data architecture that could scale economically and capture more data. TRUECar data in Hadoop makes the market work more efficiently, and all the parties like car buyers, dealers and manufacturers share in the benefits of the efficiency.

V. CONCLUSION

Analytics are providing a limitless business intelligence workspace to support how people think and work, giving them the ability to find the right information, gain insight, share it with others and see the business from any perspective. Today’s business intelligence landscape is rapidly changing and does not show signs of slowing down. This article stated the challenges associated with traditional business intelligence platforms that had driven business leaders to look for modern, forward-looking, flexible solutions.

The proposed architecture is built on a proven technology platform. It is designed to upgrade seamlessly and to cost-effectively scale for the broadest of deployments. It keeps maintenance costs down and improves productivity.

One of the key challenges with most analytics platforms is that business users are unable to interact with data while the underlying data is being processed or loaded from other systems. The future work will be to introduce Always On technique that enables the business users to continuously view dashboards or visualize data while data is being processed or loaded into analytical systems.

REFERENCES

[1]. Business Analytics – Wikipedia, https://en.wikipedia.org/wiki/Business_analytics

[2]. The importance of business analytics, **2017**. https://www.businessblogshub.com/2017/07/the-importance-of-business-analytics/

[3]. Business Intelligence - Architecture, Components and its Benefits.https://www.managementstudyguide.com/business-intelligence.htm

[4]. NEUSTAR, A Case Study, **2017**. https://www.hortonworks.com/

[5]. Hortonworks, Data Architecture Optimization, Data Architecture Optimization Customer Use Cases, **2016**.

[6]. Edureka, Understanding Pentaho Architecture, 2018. https://www.edureka.co/

[7]. Fuse by Cardinal, https://www.cardinalhealth.com/ en/ about-us/ who-we-are/ fuse-by-cardinal-health.html

[8]. Gintarė Vizgaitytė, Rimvydas Skyrius, Business "*Intelligence in the process of decision-making: Changes and Trends*", ekonomika, Vol. **91**, Issue.**3**, **2012**.

[9]. IBM Software Group, Cognos Enterprise Business Analytics, The right architecture for business intelligence. The foundation for effective enterprise BI, IBM Corporation, **2012**.

[10]. Mike Barlow, Real-time big data analytics: Emerging architecture tools and technologies driving real-time big data analytics, 2015.

[11]. N Chandler, et al., Gartner's business analytics framework, Gartner, Inc., **2011**.

[12]. Nenshad Bardoliwalla, Paxata, Paxata Solution and Hortonworks Data Platform, **2016**.

[13]. Nicolaus Henke., et al. The age of analytics: Competing in a Data-driven world, **2016**.

[14]. OBAWA, Oracle business analytics warehouse architecture, **2012**.

[15]. Paramita Ghosh, Data Strategy Needs to Include a Robust Data Architecture, **2018**.

[16]. P Needleman, M K Sternitzke, Modern Business Intelligence: The path to big data analytics, Delloite, **2018**.

[17]. Ulrika Jägare *Unified Analytics Databricks*, Special Edition, John Wiley & Sons, Inc. **2019**.

[18]. Oracle, An Enterprise Architect's Guide to Big Data Reference Architecture Overview, **2016**.

[19]. ClearPeaks. What can Big Data do for BI? https://www.clearpeaks.com/what-can-big-data-do-for-bi/

[20]. Srinivasa Kalyanachakravarthy Vasagiri, Data Architecture Optimization with Hadoop, **2016**.

[21]. K. Palanivel**, "***Modern network analytics architecture stack to enterprise networks*", Intel. journal for research in applied science & engineering technology**,** Volume **7**, Issue **IV**, **2019**.

[22]. Hugh J. Watson, "*Business Analytics Insight: Hype or Here to Stay?",* Business Intelligence Journal, Volume **16**, Issue **1**, **2010**.