ISSN: 2321-7782 (Online)
Volume 3, Issue 7, July 2015
International Journal of Advance Research in Computer Science and Management Studies
Research Article / Survey Paper / Case Study
Available online at: www.ijarcsms.com

An Incremental Framework for Internet of Things (IOT) and Big Data - A business Intelligence (BI) Prospective

Gadasu Preeti[1]
Assistant Professor,
Dept of Computer Science Engineering,
Methodist College of Engineering &Technology,
Hyderabad, Telangana. India

Sandeep Ravikanti[2]
Assistant Professor,
Dept of Computer Science Engineering,
Methodist College of Engineering &Technology,
Hyderabad, Telangana. India

***Abstract:** Internet of Things (IoT) covers every aspect of our lives and to generate a paradigm shift towards a hyper-connected society. IoT has become a mainstream topic. This growing credit is due to the impact the IoT had on business analytics and the prospective that still remains untouched. Each day, new machines, sensors, and devices come online and feed information into data systems. As organizations get on new IoT initiatives and work to abstract more insight from huge data volumes, an architectural approach is shown. Traditional databases and analytics architectures will always be fundamental, but the IoT calls for specific capabilities to handle varied data constantly streaming from various numbers of sources. In this paper we address technological rootmap from IoT – Bigdata to Business Intelligence that leads to bigdata analytics.*

I. INTRODUCTION

"The Internet Of Things Takes Over Big Data As The Most Hyped Technology". Big data has moved down the "trough of disillusionment," replaced by the Internet of Things at the top of the hype cycle.The term Internet of Things was coined by the British technologist Kevin Ashton in 1999, to describe a system where the Internet is connected to the *physical* world through ubiquitous *sensors*.Today, the huge amounts of data we are producing and the advances in mobile technologies are bringing the idea of Internet connected devices into our homes and daily lives.RFID and sensor technology enabling computers to observe, identify and understand the world without the limitations of human-entered data.However, people took it beyond the capture of "physical" events/data. An envision of network of things that was wholly dependent on human beings for information and expanded to involve anything that touched a person.Capturing the behavior of people will need broader collection of data beyond sensor technology ,beyond the "physical" that is web server clickstream data, e-commerce transaction data, customer service call logs, search logs, video surveillance, documents, etc.To understand the behavior of people, you need to capture data from any touch point, gaining a holistic view of that person. Gaining a 360 degree of your customers or a 360 degree view of your business by leveraging an environment of structured data that can be analyzed. IoT of physical devices becomes a subset of the data sources available to such a project.

Big Data: [1] Is a term used to describe the massive growth and availability of structured and unstructured data. While the term refer to the volume of data, it also refers to the technology (tools and processes) that an organization requires to handle these data volumes and storage facilities. Big Data spans three dimensions: Volume, Velocity and Variety. **Importance of Big Data: [1]** Everyone is talking about Big Data trends, from challenges to the tools required for Big Data projects. Businesses understand that Big Data infrastructure will help them make better decisions. When Big Data platforms is effectively and efficiently captured, processed and analyzed, organizations gain a clear and complete understanding of their business which would lead to efficiency improvements, lower costs, increased sales and better customer service. **Business Intelligence:** Business intelligence (BI) [2] is an umbrella term used to encompass the processes, methods, measurements and systems businesses use to more easily view, analyze and understand information relevant to the history, current performance or future projections for a business.The goal of BI is to help decision-makers make more informed and better decisions to guide the business. Business intelligence software and software-as-a-service (SaaS) solutions achieve this by making it simpler to aggregate, see, and slice-and-dice the data. Thus makes it easier to identify trends and issues, discover new insights, and fine-tune operations to meet business goals.

II. RELATED WORK

1) IoT a Subset of Big Data

A simple incremental view of how IoT feeds Big Data which then feeds a broader analytic platform. Imagine IOT as a bunch of customized data sources (sensors and machines) leveraging customized collectors that feed a comprehensive platform (e.g. Hadoop vendors like Cloudera and Hortonworks) which, in turn, allow us to feed downstream analytic, BI, and visualization platforms.

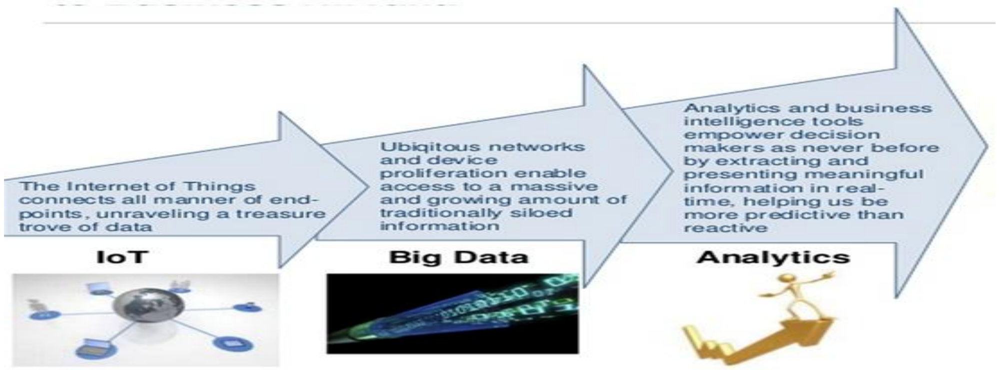

Figure1: Reference image is taken form ref [12] : Incremental View

2) Sensors the core of Iot

A sensor[13] is technically any device which converts one form of energy to another form, the end being usually being an electrical form mainly for measurement, control or monitoring purposes. Consider a typical temperature sensor like a gas pressure based tube sensor which expands or contracts to convert the temperature into a mechanical motion which can be displayed, recorded or used for control as required. Translation a thermostat as used in a refrigerator. The raw electrical signal from a physical sensor is usually in analog form, and can be conveniently processed further and displayed on a meter or other suitable indication device or recorded on paper or other media such as magnetic tape or more advanced digital systems as required. The sensor[13] is classified as per its application and there could be many different types of sensors, with their own inherent advantages or disadvantages for a particular application. Putting it simply, the sensor generates an output which can be conveniently displayed, recorded or used to control or monitor the application at the point where the sensor is installed. What's so special about sensors? You can translate the analog physical world into a digital computer world ,where we convert the sensor's analog signals in to digital signals so, the computer can be able to read it and then we feed that with other digital signals into a Big Data platform."Technologies that operate upon the physical world are a key component of the digital business opportunity." I think IoT requires a lot of talent on the many types of physical sensors and how they are ultimately converted into a form that the emerging Big Data platforms can consume and analyze them.

3) IoT Needs a Big Data Platform

Getting your fridge to talk to you through sensors is one thing, getting your plants to talk to your heating system and to you is quite another. As we map the spread of the IoT, it starts to get more difficult and barriers appear with the centralized big data platform or halt progress JeffHagins Founder and [5] CTO of SmartThings, described the data platform he has been working on that should help expand the IoT and help product designers work out new ways of connecting machines and people. He believes that the Internet of Things has got to be built on a platform that is easy, intelligent and open. The evolving Big Data platforms will become a standard for IoT-based applications and IoT is just that, a set of specialized sensor connectors and Big Data applications. The blurring the physical and virtual worlds are strong concepts in this point. Physical assets become digitalized and become equal factors in understanding and managing the business value chain alongside already-digital entities, such as big data systems and next-generation applications.

III. BUSINESS INTELLIGENCE AND INTERNET OF THINGS

For business intelligence (BI) [8] we inten

d the set of methods and models that explore the data in order to obtain information and then knowledge .Internet of Things: a global network of interconnected objects. In the year 2008, for the first time more object than people were connected to Internet. Merging the concepts from those two fields will provide new ideas and methods to solve problems.

Technology roadmap of IoT

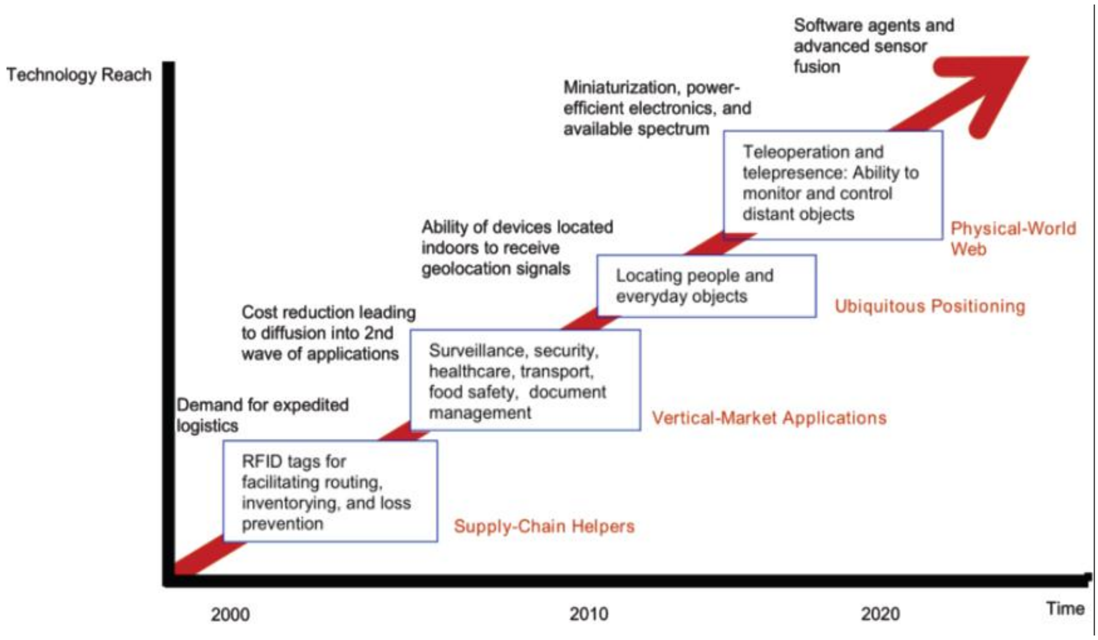

Figure2: Technology moves of Iot

Enabling Factors of BI and IoT

Tagging, sensing, shrinking, connecting have made easier to access and share large amounts of data.

- » Data available from many sources, but heterogeneous in origin, content and representation.
- » Commercial transactions, financial, administrative.
- » Transport & energy.
- » Clinical data.
- » Their presence opens scenarios and opportunities that were unthinkable before.

IV. BIG DATA TO BUSINESS INTELLIGENCE

In older tools and in most current solutions BI tells you what happened in a specific segment of your business. With how quickly business is moving today, that kind of BI is as problematic as driving down the freeway by looking only in your rear-view mirror. With new technology and new expectations, BI is moving toward a more predictive model that shows you what will happen. New BI systems are now beginning to show how all the various parts of your organization work together to produce an outcome, and business leaders can finally see the big picture and make faster, better-informed decisions.

V. Architecture And Related Discussions

The data components of a BI architecture [3] include the data sources that corporate executives and other end users need to access and analyze to meet their business requirements. Important segment in the source selection process include data currency, data quality and the level of depth in the data. Both structured and unstructured data may be required as part of a BI architecture, as well as information from both internal and external sources. Information management architectural components are used to transform raw transaction data into a consistent and coherent set of information that is suitable for BI uses. For example, this part of a BI architecture [3] typically includes data integration, data cleansing and the creation of data dimensions and business rules that conform to the architectural guidelines. It may also define structures for data warehousing that aggregates information in virtual databases instead of physical data warehouses or data marts.

The technology components are used to show information to business users and enable them to analyze the data. This includes the BI software suite or BI tools [4] to be used within an organization as well as the supporting IT infrastructure – i.e., hardware, database software and networking devices. There are various types of BI applications that can be built into an architecture: reporting, ad hoc query, data mining and data visualization tools, and online analytical processing (OLAP) software, business intelligence dashboards and performance scorecards.

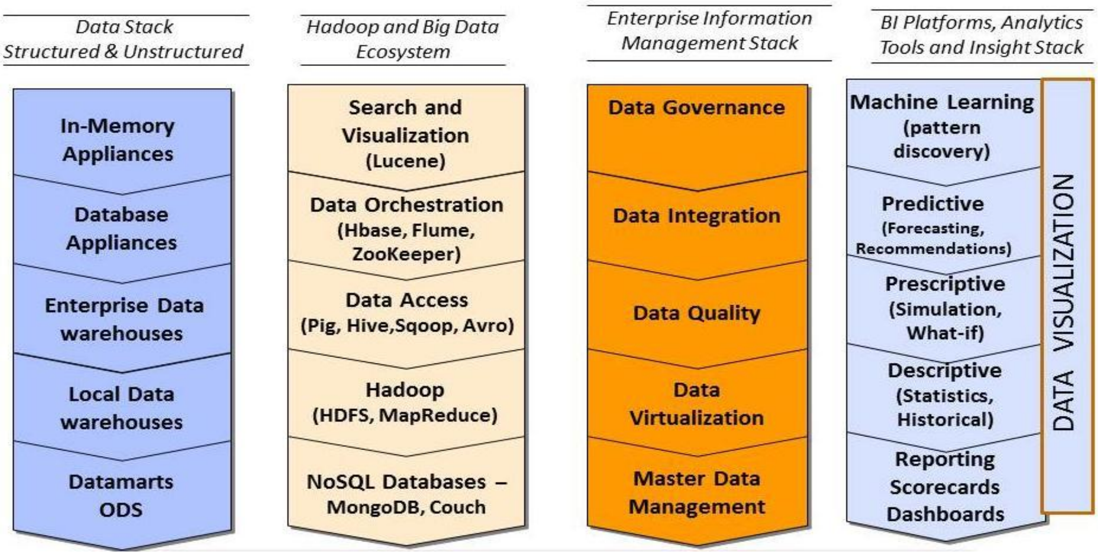

Figure3: Structured and Unstructured Data [3], BI architecture

Whether the IoT and sensors are Big Data: The discussion [4] is on architecture challenges–from device to data to data consumer–created by the start of millions (or billions) of connected sensors and smart things. Connected devices bring back some standard engineering challenges back. How do you transmit data securely and with low power consumption.How do you handle lossy networks and cut-off transmissions. Everything is not smartphone application transmitting JSON over HTTP. How do you handle communication myriad protocols, each of which could be using a near-infinite variety of data encoding formats. IoT data is messy. Devices get cut off in mid-transition or repeat. How do you detect this–and clean it up–as data arrives. IoT data is of high volume. By 2020, we will have 4x more sensor and IoT data than enterprise data. We already get more data today from sensors than we do from PCs. How do we scale to consume and use this. In addition, connected devices are not always smart or fault-tolerant. How do you ensure you are always ready to catch all that data (i.e., you need a zero-downtime IoT utility).IoT and sensor and of itself is not terribly useful. It is rarely in a format that a business analyst would even be able to read. It would be incredibly wasteful to store all this as-is in a business warehouse, DropBox repo, etc.. IoT and sensor data needs context. Knowing device Knowing that FE80:0000:0000:0000:0202:B3FF:FE1E:8329 is at GPS location X,Y is of no use. You need to marry it to data about the "things" to get useful insights. IoT data is in two view points: what does this mean right now and what does this imply for the big picture. The Lambda Architecture [6] is an ideal tool to handle this.

VI. TECHNOLOGIES

The Lambda Architecture, first proposed by Nathan Marz, attempts to provide a combination of technologies that together can provide the characteristics of a web-scale system that can satisfy requirements for availability, maintainability, and fault-tolerance. Big data analytical ecosystem architecture [7] is in early stages of development. Unlike traditional data warehouse / business intelligence (DW/BI) architecture which is designed for structured, internal data, big data systems work with raw unstructured and semi-structured data as well as internal and external data sources. Moreover, organizations may need both batch and (near) real-time data processing capabilities from big data systems. Lambda architecture provides a clear set of architecture principles that allows both batch and real-time or stream data processing to work together while building immutability and recomputation into the system.

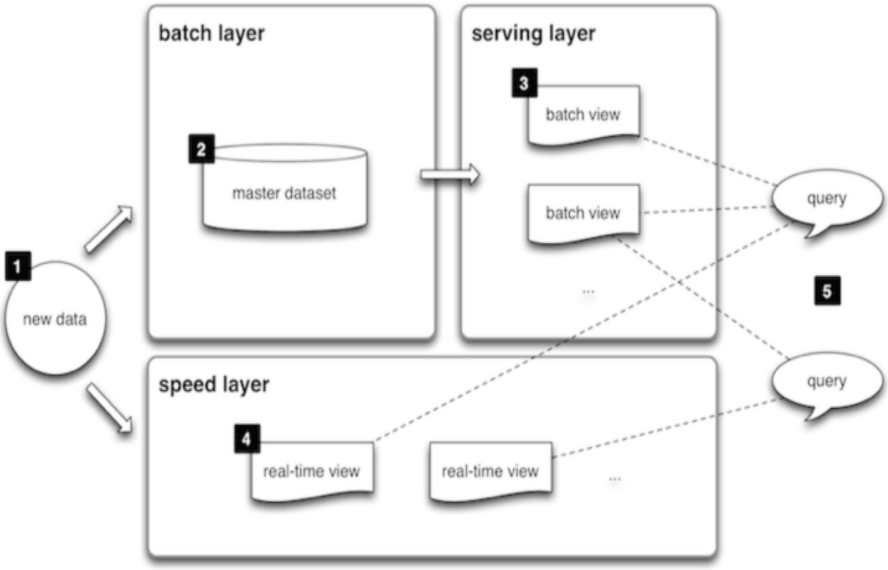

Figure4: Lambda Architecture[7] ; various layers

Batch processes high volumes of data where a group of transactions is collected over a period of time. Data is collected, entered, processed and then batch results produced. Batch processing requires separate programs for input, process and output. An example is payroll and billing systems. In contrast, real-time data processing involves a continual input, process and output of data. Data must be processed in a small time period (or near real-time). Customer services and bank ATMs are examples. Lambda architecture has three (3) layers:

» Batch Layer

» Serving Layer

» Speed Layer

Batch Layer (Apache Hadoop): [7] Hadoop is an open source platform for storing massive amounts of data. Lambda architecture provides "human fault-tolerance" which allows simple data deletion (to remedy human error) where the views are recomputed (immutability and recomputation).The batch layer stores the master data set (HDFS) and computes arbitrary views (MapReduce). Computing views is continuous: new data is aggregated into views when recomputed during MapReduce iterations. Views are computed from the entire data set and the batch layer does not update views frequently , resulting in latency.

Serving Layer (Real-time Queries): The serving layer indexes[7] and exposes precomputed views to be queried in ad hoc with low latency. Open source real-time Hadoop query implementations like Cloudera Impala, Hortonworks Stinger, Dremel

(Apache Drill) and Spark Shark can query the views immediately. Hadoop can store and process large data sets and these tools can query data fast. At this time Spark Shark outperforms considering in-memory capabilities and has greater flexibility for Machine Learning functions. Note that MapReduce is high latency and a speed layer is needed for realtime.

Speed Layer(Distributed Stream Processing):[7] The speed layer compensates for batch layer high latency by computing real-time views in distributed stream processing open source solutions like Storm and S4. They provide:

» Stream processing

» Distributed continuous computation

» Fault tolerance

» Modular design

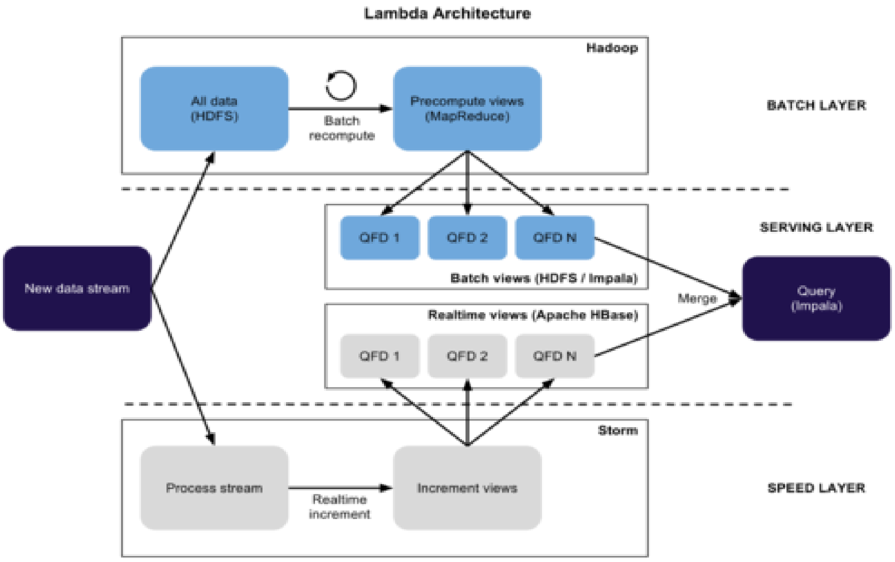

Figure5: Layers in Detail

In the speed layer real-time views are incremented when new data received. Lambda architecture provides "complexity isolation" where real-time views are transient and can be discarded allowing the most complex part to be moved into the layer with temporary results. The decision to implement Lambda architecture depends on need for real-time data processing and human fault-tolerance. There are significant benefits from immutability and human fault-tolerance as well as precomputation and recomputation. Lambda implementation issues include finding the talent to build a scalable batch processing layer. We can work with Hadoop, MapReduce, HDFS, HBase, Pig, Hive, Cascading, Scalding, Storm, Spark Shark and other new technologies.

VII. BIG DATA ANALYTICS INFRASTRUCTURE

Big data analytics [8] refers to the science and analysis of both internal and external data to obtain valuable, actionable insights that allows the organization to make better decisions. Recent surveys suggest the number one investment area for both private and public organizations is the design and building of a modern data warehouse (DW) / business intelligence (BI) / data analytics architecture that provides a flexible, multi-faceted analytical ecosystem.

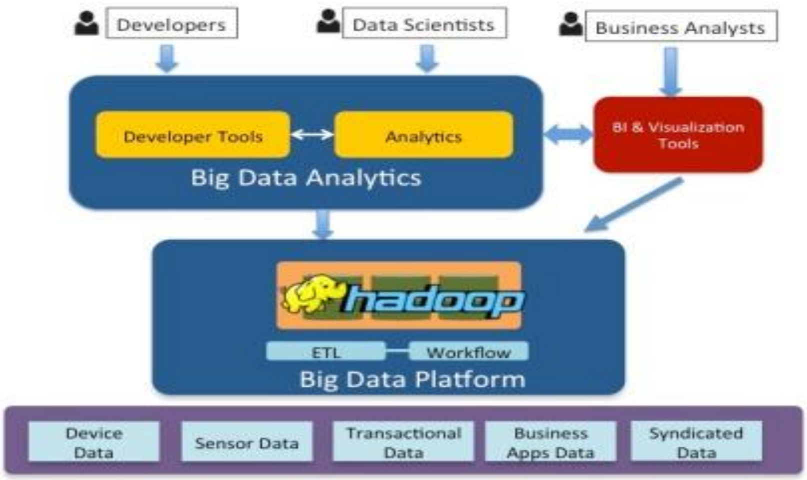

Figure6: Bigdata Analytics[8]

The goal is to leverage both internal and external data , as well as structured and unstructured data - to gain competitive advantage and make better decisions. While traditional business intelligence (BI) technologies provide historical, current and descriptive views of business operations, big data analytics focuses on data science, predictive analytics, data mining, decision optimizing processes and business performance management.

Five key benefits to data science and business analytics:

1. Improving the decision-making process.
2. Speeding up the decision-making process.
3. Better alignment of resources with strategies.
4. Realizing cost efficiencies.
5. Responding to user needs for availability of data on a timely basis.

By embracing data science, decision optimizing processes and an analytical approach, companies identify their most profitable customers, accelerate product innovation, optimize supply chains and pricing, and identify the true drivers of financial performance.

VIII. APPLICATIONS

The COMPOSE project: [10] It has built a cloud platform that helps developers find their way smoothly through the booming, Internet of Things (IoT). Their mission is simple: to allow developers to create apps that provide people with Internet services and launch them into the market quickly , whether that be shopping or traffic information systems, or home-based apps .'We give developers a one-stop shop where it is easy to design and deploy IoT applications, providing building blocks so they can create their own dream apps.[8] . We hope that opening the door to this realm for smaller developers will lead to higher innovation. For this, COMPOSE (Collaborative Open Market to Place Objects at your SErvice) has developed and uploaded libraries of software, all of which can be downloaded free from its open source code repository, GitHub, the largest code host in the world. The project is continually adding new items to GitHub so users can acquire COMPOSE smart 'objects' or modules from the repository. They can combine these to create their own apps easily and quickly. It saves building any app from scratch, avoiding wasting time and development money in the process, by using basic blocks that have been developed by programmers in the past and shared through the COMPOSE project.

Shopper behaviour, car sharing and happy skiers: COMPOSE [8] is conducting three pilot projects. One of these involves the start-up U-Hopper, which has won awards for its COMPOSE-based retail analytics platform, RetailerIN, currently on trial at the SAIT-COOP supermarket in Trento, Italy. Shoppers' carts and baskets are tracked to create a heat map of where customers spend their time in the store. From the office, the store manager can monitor the effectiveness of displays and campaigns, and the queues forming at various counters, changing the supermarket's strategies to suit. The second pilot involves car sharing among around 750 staff and students at the University of Tarragona in Spain. Through social media, the app encourages car sharing by linking it to reserving spaces in the university car park, thus reducing the number of vehicles travelling to the university every day. In the third pilot in a resort of Trentino, Italy, skiers get real-time snow and weather conditions fed to their smartphones from a network of meteorological stations. The app, Go2Ski, also helps friends meet up and share photos, and even informs on the length of queues at ski lifts. And it is truly an app for all seasons. When the snow melts it uses the same meteorology network, but for warmer weather sports such as cycling.

IX. CONCLUSION

In conclusion, Internet of Things is the concept in which the virtual world of information is connected to the real world of things. The technologies of Internet of things such as sensors, bigdata make our life more comfortable. This paper discusses the Internet of Things which is one of the upcoming concepts in the field of Internet. We analyzed the architecture, its technologies and real life applications. An incremental view of Iot and Bigdata to Business intelligence leads to Big data analytics [8] that refers to the science and analysis of both internal and external data to obtain valuable, actionable insights that allows the organization to make better decisions.

References

1. http://www.happiestminds.com/Insights/Big-Data/
2. http://www.smallbusinesscomputing.com/News/Software/article.php/3896051/What-is-Business-Intelligence-and-Why-Should-You-Care.htm
3. http://www.savi.com/2015/how-to-architect-for-the-internet-of-things-iot/
4. http://www.b-eye-network.com/
5. www.ctosmartthings.org
6. http://www.dbta.com/Columns/Big-Data-Notes/Real-Time-Big-Data-With-the-Lambda-Architecture-99798.aspx
7. http://www.datasciencecentral.com/profiles/blogs/lambda-architecture-for-big-data-systems
8. http://www.rosebt.com/big-data-analytics-infrastructure.html
9. www.ibmresearch.com
10. http://www.libelium.com/smart-city-urban-resilience-smart-environment/(app final)
11. http://www.meetup.com/Internet-of-Things-DC/messages/boards/thread/49041291/(app)
12. https://www.googleimags.com
13. www.datamation.com

