Super SCADA Systems: A Prototype for Next Gen SCADA System

Mr. VijayaRamaraju Poosapati MCA, (Ph.D), Prof. Vedavathi Katneni MCA, M.Phil, Ph.D,
Mr. Vijay Killu Manda, MCA, MBA.
Gitam University, Visakhapatnam, Andhra Pradesh
Vijaramaraju.Poosapati@gmail.com
Vedavathi@Gitam.edu
MVKillu@gmail.com

Abstract: Supervisory Control and Data Acquisition (SCADA) systems are basically industrial control systems, with real time processing and distributed intelligence. They process the data and help to control and monitor the operations in an industry. The industries needs are now changing with the advances in information communication and computational technologies and with the availability of low cost sensors, wi-fi enabled premises, open source software, and hardware to handle huge volume of data. The new digitalization movements like Industry 4.0 demands enhanced capabilities of Supervisory Control and Data Acquisition (SCADA) systems by leveraging the recent advances to improve the productivity of an industry at a low cost

The key features of a SCADA systems like schematic visualizations, alarming, data logging, real time control and data archiving would continue to adopt and evolve entirely in a new way. A 'Super SCADA System' as we call it, is designed by integrating traditional SCADA system with industry ERP and other data collecting software applications to use machine learning libraries for predictive analysis and open source web based applications for real time monitoring and controlling. In this paper, we presented the SCADA systems architecture and discussed the technologies used to build a prototype of a Super SCADA Systems with predictive analytics and real-time monitoring capabilities

Key words: Open Source SCADA, Predictive Analytics, Machine Learning, Analytics, Industrial Automation

Highlights:

1. Cost effective industrial automation solution
2. Integration of ERP, SCADA and Transactional databases of an industry
3. Time to Failure prediction for Zero-Unplanned outages using complete industrial data.
4. Open source data visualization tools and web based controls
5. Platform independent by using modern systems designs

1. INTRODUCTION

Supervisory Control and Data Acquisition systems (SCADA) are basically a computer based software or applications that are used in industries to collect the operational data, process it and control network operations as per pre-programmed standards. The advance in technology and ease of integration of multiple data sources created an opportunity to enhance the capabilities of SCADA to improve the overall efficiency of machines. The generated data from multiple sources of an industry is useful when processed appropriately and used to solve critical problems like manpower utilization, unexpected break down of machines etc.

Small and medium size companies often do not use data from all possible sources. This is due to the capital expenditure involved in integrating and managing huge volume of data and being reluctant to share the data outside their network to use cloud based solutions. Therefore, there is a need for a cost effective, on premises and a platform independent solution for predictive analytics.

Hervé Panetto and Arturo Molina1 have discussed the best practices to integrate manufacturing systems with enterprise systems and challenges faced in the process. Akhil Dixit2 have used the combination of PLC and SCADA system to continuously gather information from the system on real time basis and displays it on the computer screen through various trends, graphs and alarms. R. Subramani and C. Vijayalakhsmi3 proposed a novel SCADA-based decentralized approach for power management in today's electricity market for efficient energy power management system. Aamir Shahzad and Young-Gab Kim4 studied and analysed the potential security mechanisms for securing SCADA-IoT and found cryptography to be a noteworthy security solution, based on the proposed system requirements and its communication demands. Mian Du5 introduced an auxiliary decision-making method by considering the parameter selection for modelling a wind turbine's behaviour through machine learning techniques. Abdalhossein Rezai6 have investigated important security threats in SCADA networks, and reviewed the ongoing works on SCADA key management scheme in the SCADA networks. Pen Sun7, have used three types of assessment indices namely Neural networks, Abnormal Level Index and Back propagation neural networks in monitoring the parameters obtained from the SCADA system. Philip Church8, proposed the adoption of event-driven communication protocols to reduce network transfer because the protocol conversion is close to the field devices, and replication of remote servers to ensure polling intervals are met. Traian Turc9 expressed the WEB services from different SCADA systems, The main contribution of the author consists in the SCADA software application, database architecture, and WEB services design. Ravina Sharma10, proposed a prototype model of automated storage and retrieval system with the help of PLC and SCADA. S. S. Saha11 developed a low cost, high efficient embedded system that collect the data from the sensor and made it available to the user from remote location anytime and discussed online temperature monitoring and control through a webpage. Mahmoud Shaker Nasr12 presented the design of a new generation of SCADA, named moveable SCADA, based on IoT cloud.

Considering all the above an attempt is made to design a Super SCADA System which integrates multiple data sources with ERP and SCADA and helps in predicting the time to failure of the machines using machine learning algorithms customized as per the business requirements. These systems enable real-time processing of data and are reliable for generating business intelligence reports. The paper outlines a brief description about the components of SCADA system, the evolutionary generations of SCADA Systems and an introduction of proposed Super SCADA System, its architecture and system design concepts.

2. COMPONENTS OF A SCADA SYSTEM

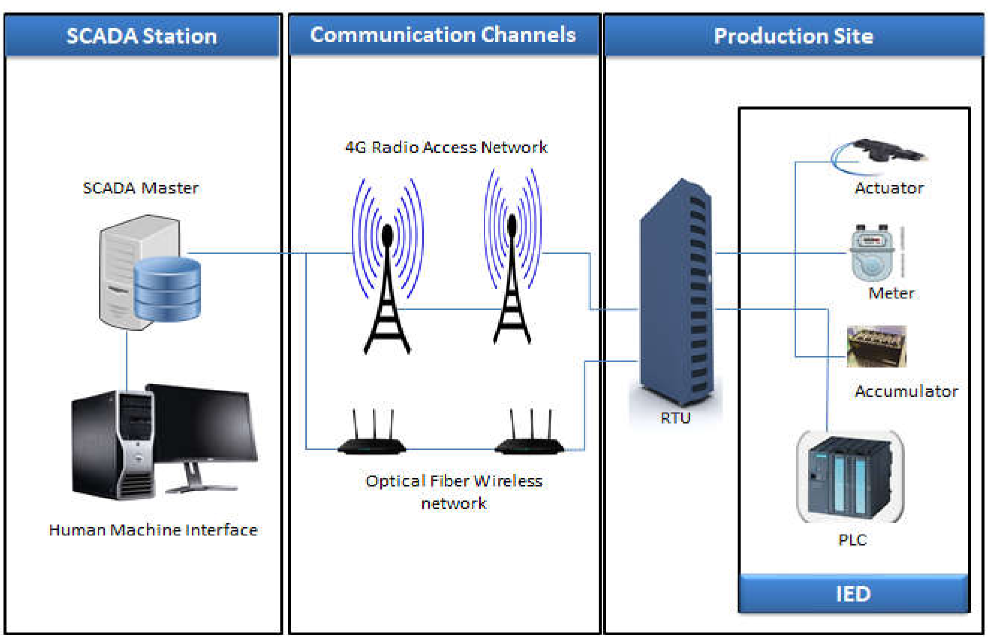

Figure-2.1: SCADA Components

2.1 HMI: It is a device which shows the processed data collected from sensors to operators. It helps to control the devices remotely by sending the instruction from operators to machine.

2.2 SCADA Master: A computer server collects, stores and processes all the data from multiple devices and sends back the control commands.

2.3 RTU: Remote Terminal Unit (RTU) or Intelligent Electric Devices (IED) acts as an interface between SCADA and physical devices. They help to convert sensor signals to digital data and transmit the digital data to SCADA master or supervisory control.

2.4 PLC: Programmable logic controller (PLCs) are low cost alternative to RTU. PLC are used as field devices. They are versatile, flexible, and can be configured when compared to RTU.

2.5 Communication Channels: These are Communication infrastructure that is connecting the supervisory system to the remote terminal units.

2.6 Physical Devices: Sensors, Meters, Actuators.

3. EVOLUTION OF SCADA

SCADA Systems were first employed in early 1980s where computation was mostly done using a mainframe computer. Over last few decades the core architecture of SCADA did not change much when compared to the changes that are noticed in communicational and computational technologies.

3.1 First Generation SCADA Systems:
These were evolved during the Mainframe era, in which all the data was stored in a single mainframe computer and entire computation was done on single mainframe system. In this generation, the systems were not interlinked. The communication protocols to transfer and receive data from RTU were usually developed by RTU vendors. They perform functionality of scanning and controlling data within the remote devices. These systems were well known as Monolithic as all data was around one single mainframe device.

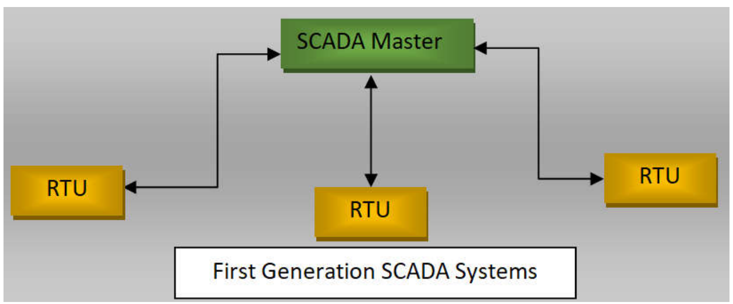

Figure- 3.1: First Generation SCADA Systems

3.2 Second Generation SCADA Systems:
The commercial use of Local Area Networks (LAN's) helped in sharing control functions those were distributed across the multiple systems connected to each other. These were named as distributed SCADA systems. These distributed SCADA Systems were tiny and less expensive when compared with Monolithic systems. The processing performance of SCADA systems improved due to redundancy and use of reliable LAN protocols.

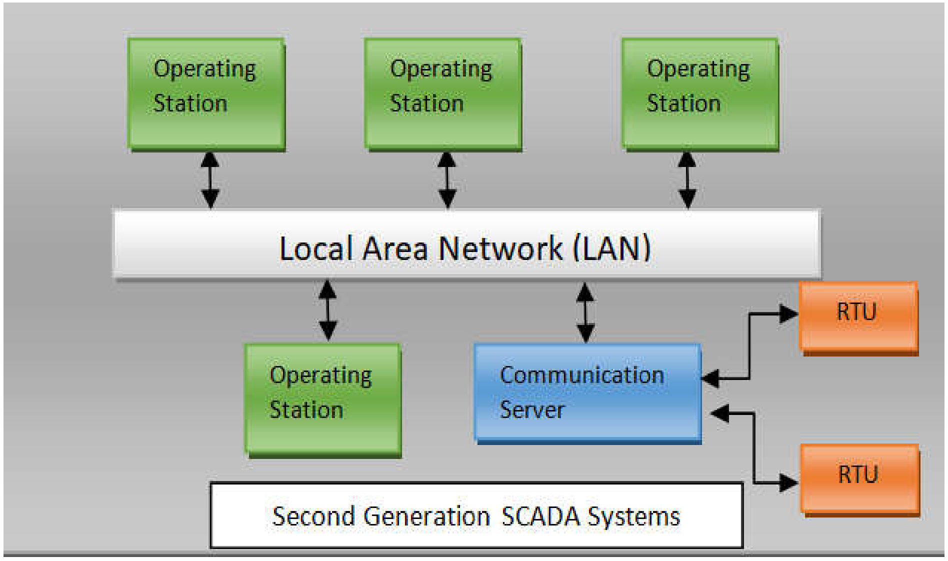

Figure-3.2: Second Generation SCADA Systems

3.3 Third Generation SCADA Systems:

With the evolution of open source Internet protocols for communication, SCADA systems could overcome the limitations of proprietary systems. Third generation SCADA systems were generally networked and communicated using Wide Area Network (WAN) Systems over data lines or phone. They used Ethernet or Fiber Optic Connections for transmitting data and Programmable Logic Controllers (PLC) for monitoring and controlling the devices. Unlike Monolithic and Distributed SCADA Systems, network SCADA systems suffered with risk as they were connected through Internet.

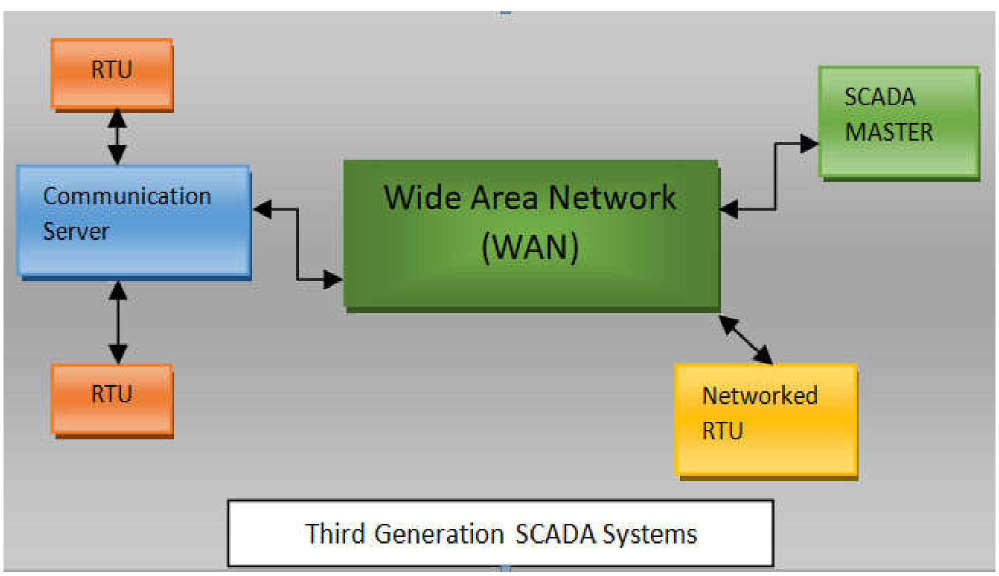

Figure-3.3: Third Generation SCADA Systems

3.4 Fourth Generation of SCADA Systems:

Evolution of cloud based computing methodologies and various open source IoT platforms helped SCADA systems to become more dynamic and cost effective. The current generation of SCADA systems are easy to maintain and can be integrated with other components easily. They became flexible and scalable to handle data of any kind of volume or that require computational challenges.

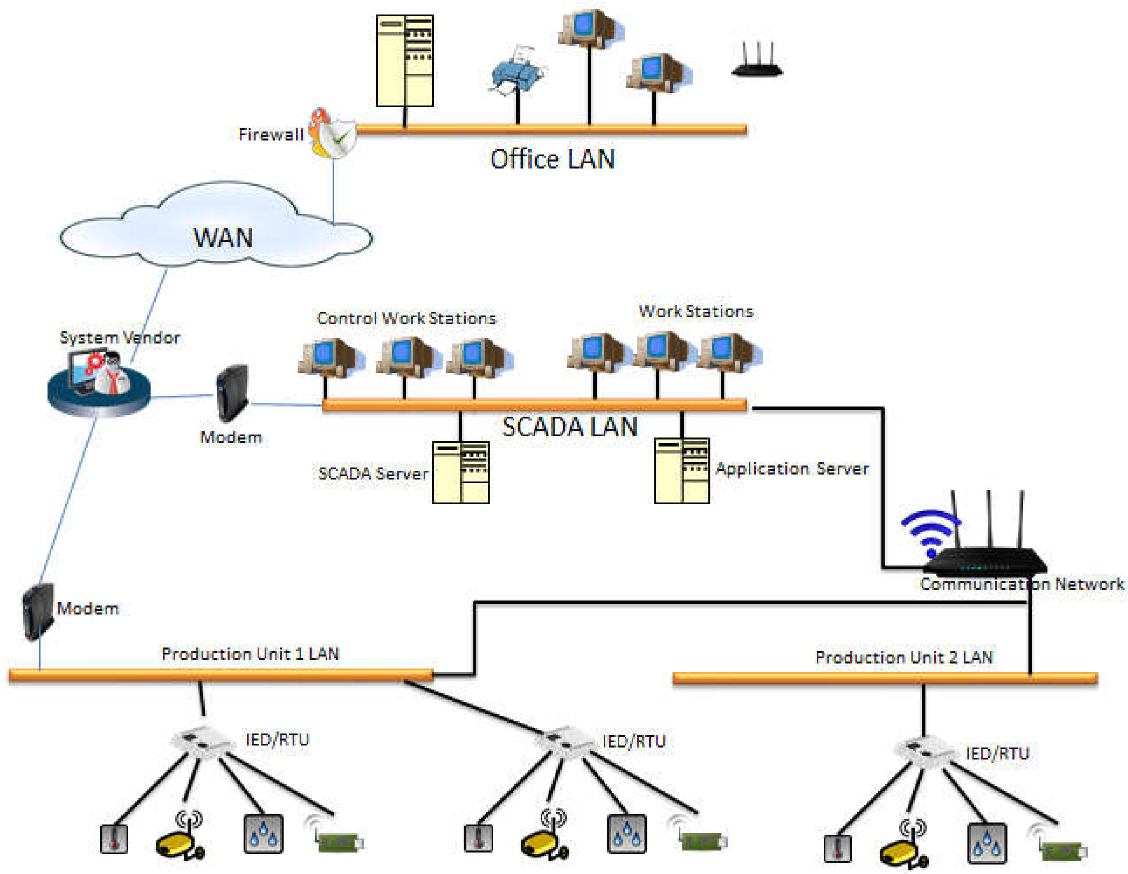

Figure- 3.4: Fourth Generation SCADA Systems (Currently Used)

4. PROPOSED SUPER SCADA SYSTEMS (SSS)

With the adoption of IIoT and by leveraging latest technologies, connected enterprise has now become a reality. The huge volume of data that is being generated in the industry will provide useful information when processed effectively. SCADA System helps to collect data from field devices like sensors and helps in monitoring and controlling, but they operate as a standalone devices and data is not exchanged with ERPs or Decision Support Systems or Data warehouse. Interoperability of data between devices is a major challenge as they use proprietary Software with huge licencing cost. Thus, it has become a bottleneck to small and middle scale companies to leverage IIoT benefits. Industries need the processed information to operate their processes accurately and as and when the information is generated. The processes in an industry are most complex and contain multiple communication interfaces which exchanges information between controllers and a centralized control station.

The proposed Super SCADA System helps to overcome this challenge by connecting people, processes and things in an intelligent way. This system integrates SCADA with Industry ERP and other industrial data sources using open source tools. The data collected from multiple sources is analysed as and when they are generated and displayed on a dashboard which runs as a web application with open source. Predictive analysis and business intelligence reports are built using open source visualization tool for effective decision making. These act as a bridge between operational technologies (OT) and business information technologies by providing flexibility to do real time or historic data analysis.

This integration in the proposed super SCADA helps to enhance the performance of three major functions of an industry namely production planning, quality analysis and maintenance analysis by eliminating the lag in intermediate process which results in effective utilization of machinery, man power and money. This ultimately leads to improvement in productivity and in benchmarking the standards and formulating strategies. Further, all these processes can be configured to connect with many diverse types of mobile/wearable or augmented reality devices.

5. ARCHITECTURE OF SUPER SCADA SYSTEM

The proposed architecture is an extension of fourth generation SCADA systems, built by integrating multiple data generating systems using open source technologies to provide interoperability, real time data access and remote accessibility to view reports and operational data in an industry. This combination of data helps in predictive analysis more effectively.

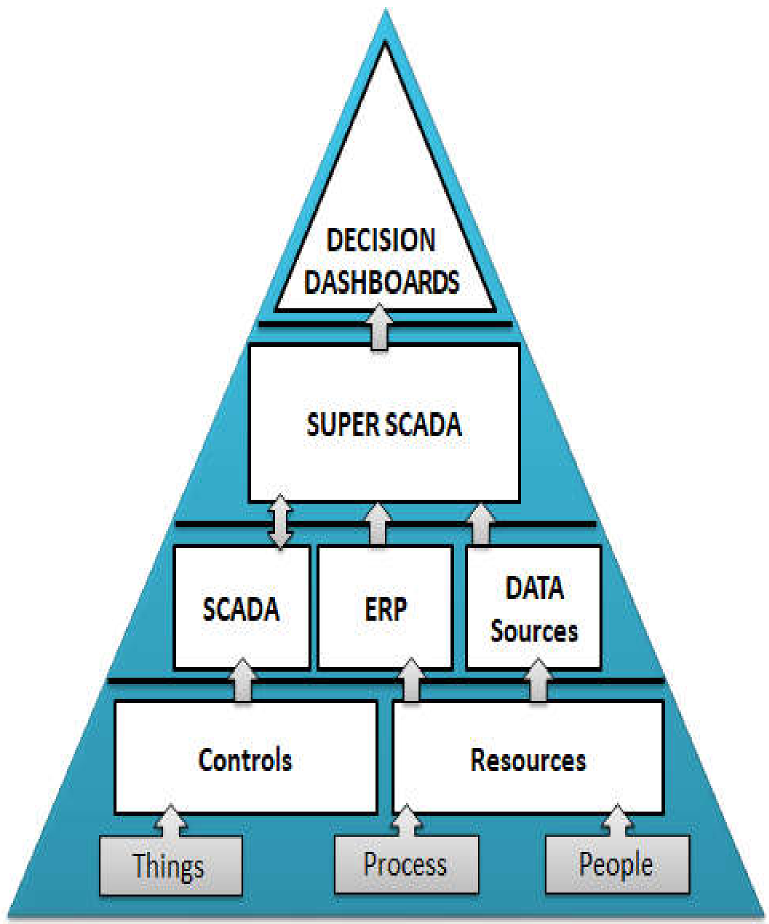

Figure-5.1: Proposed Super SCADA System Architecture

Figure- 5.1 graphically represents the proposed architectural design of the physical components of the super SCADA system and the relationship between the processes. The pyramid highlights various levels of data collections in an industry which helps in information exchange from control systems and information systems. Decision Support dashboards/reports are built using the data from Super Scada System which contains integrated data from all major data sources in an Industry. The present study considered on premises data storage instead of cloud storage to address the needs of few specific industries who do not want to use cloud for data security reasons.

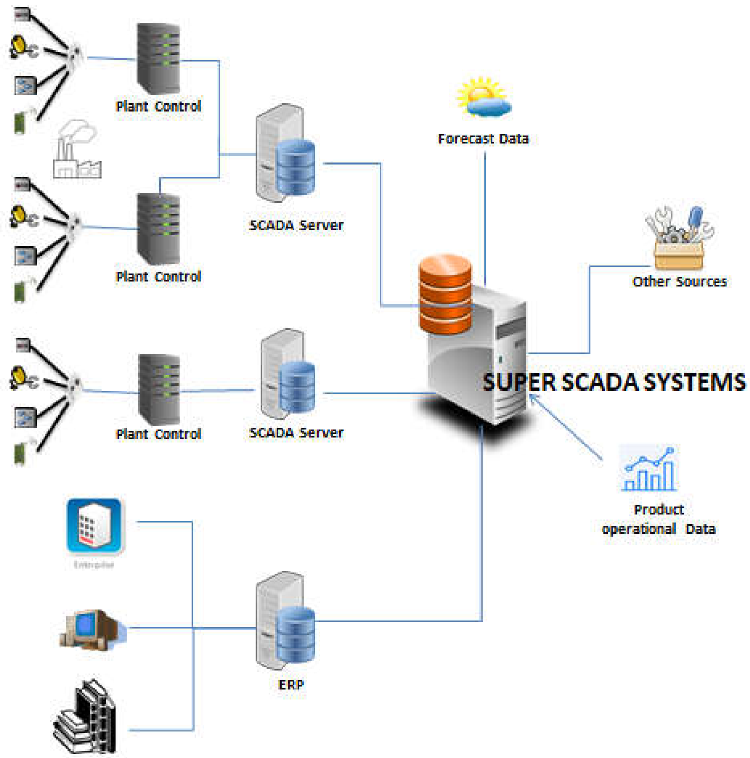

Figure-5.2: Proposed Super SCADA System Description

6. SYSTEM DESIGN AND TECHNOLOGY

The objective of the study is to develop an open source system which can integrate data from SCADA Systems, ERP and other sources into Super SCADA systems and effectively use it to accurately predict the time to failure of a machine in an industry and send proactive alerts.

A Super Scada System consists of a MySQL database, a communication manager, a webserver. APACHE 2.4.29 webserver is installed on an Ubuntu server 16.04.3. MySQL server 5.6 is used for storing the integrated data. Eclipse Neo SCADA, an open source SCADA software is considered for the prototype.

The field devices like sensors or actuators collect information of various components in a plant and sends to RTU. RTU or a PLC collect the information and send it to a SCADA system. Arduino – an open source electronics platform is used as hardware to connect the sensors. Arduino driver connects itself to the ethernet shield of Arduino and polls the data from the board. Socket.IO, a java script based library for real time web applications is used for bi-directional communication to control and gather the information from RTU.

Process level information and business data from ERP are collected and stored in a data archive of a Super Scada System. Data from multiple sources are pushed to Super Scada Systems in a JSON file. JSON is a language independent data format and is compatible with many of the latest programming languages and databases. A cron job (Linux command for scheduling a task) is run periodically to update the data archive.

Graphical user interface (GUI) for Super Scada HMI Module is designed to control the SCADA systems, manage events, send alerts (emails/SMS), alarm triggering and to view production statistics. Angular 5, Node JS, HTML and CSS are used for building the application and controls. Electron is used to build a desktop based S/W.

The data from a Super SCADA system are used for real time predictive analysis. Python scikit-learn, a free machine learning library for python is used to build a model. The predictive data received from scikit-learn is pushed and is shown using D3 JS visualization tool. In a Realtime scenario, to handle huge volume of data, Apache Spark5 is best suited for large-scale data processing. It has the capability to process data 100 times faster than traditional technologies. It is scalable with the growing needs of the data acquisition. Apache Spark comes with an inbuilt machine learning library which can be effectively used for predictive analysis instead of Scikit Learn.

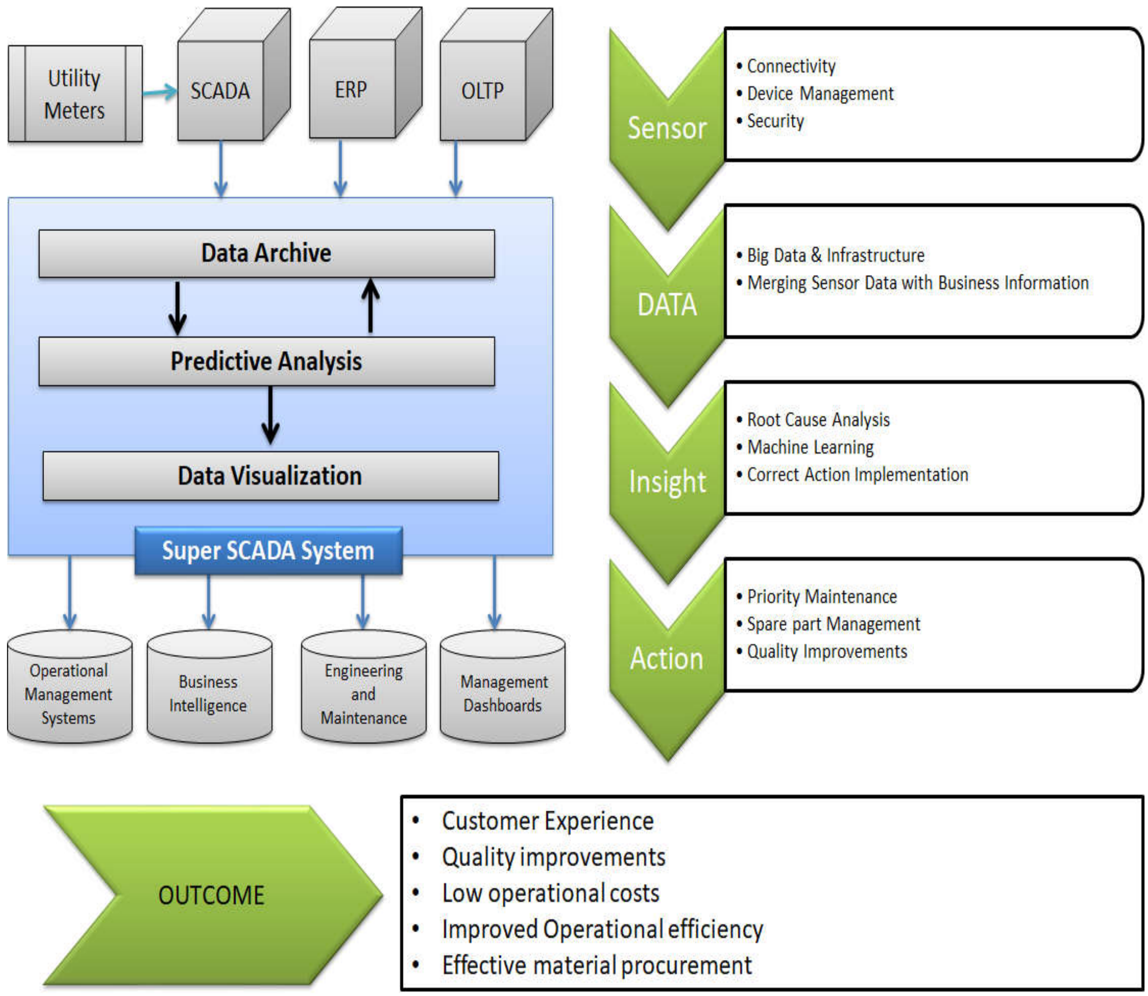

Figure-6.1: Proposed Super SCADA System applications and outcome

7. CONCLUSION AND FUTURE WORK

In this paper, the possibility of building a Super SCADA System that can predict the time of failure of a machine is discussed. The prototype provides a best combination of latest open source technologies. These technologies are to be integrated with multiple data sources to provide an economical solution for middle and small-scale companies to leverage Industrial IOT benefits.

In future a predictive model based on the data from Super SCADA System can be built and the results with predictive data from traditional Scada systems may be compared. Further, the system may be implemented using Apache Spark MLLib and with real industrial data. Finally, the ultimate objective of Super SCADA system is to make a user friendly and affordable technology which will work effectively in predicting the machine condition and helping industries to improve productivity.

References

1. *HervéPanetto & ArturoMolinab. (2008). Enterprise integration and interoperability in manufacturing systems: Trends and issues. Elesvier - Computers in Industry, Volume 59, Issue 7, 641-646.*

2. *Akhil Dixit, Rahul Mendiratta, Tripti Chaudhary & Naresh Kumari. (2015). Review Paper on PLC & Its Applications in Automation Plants. International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463 Vol. 4 Issue 3, 63-66.*

3. *R. Subramani & C.Vijayalakhsmi. (2015). Implementation of Lagrangian Decomposition Model for Energy Efficiency using SCADA System. Journal of applied sciences Research, ISSN:1819-544X, EISSN:1816-157X, 74-79.*

4. *Aamir Shahzad & Young-Gab Kim. (2018). Secure SCADA-IoT Platform for Industrial Automation and Control: A Collaborative-Communication Designed Model. Conference presented at Symmetry 2017—The First International Conference on Symmetry, Barcelona, Spain.*

5. *Mian Du, Jun Yi, Peyman Mazidi, Lin Cheng & Jianbo Guo. (2017). A Parameter Selection Method for Wind Turbine Health Management through SCADA Data. MDPI Energies Journal Article, 10-12.*

6. *Abdalhossein Rezai , Parviz Keshavarzi & Zahra Moravej. (2017). Key management issue in SCADA networks - A review. Elsevier Engineering Science and Technology, an International Journal 20, 354–363.*

7. *Peng Sun, Jian Li, Caisheng Wang & Yonglong Yan. (2017). Condition Assessment for Wind Turbines with Doubly Fed Induction Generators Based on SCADA Data. Journal of Electrical Engineering & Technology Vol.12 No.2, 689-700.*

8. *Philip Church, Harald Mueller, Caspar Ryan, Spyridon V. Gogouvitis, Andrzej Goscinski & Zahir Tari. (2017). Migration of a SCADA system to IaaS clouds - a case study. Journal of Cloud Computing: Advances, Systems and Applications, 6-11*

9. *Traian Turc. (2015). SCADA Systems Management based on WEB Services. Elsevier Procedia Economics and Finance Volume 32, 464-470*

10. *Ravina Sharma, Sapana Navtake and Prof. H. P. Chaudhari. (2017). Automatic Material Storage and Retrieval System using PLC and SCADA. International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 6, 1361-1365*

11. *Prof. S.S.Shah, Sanket Raut, Dipak Jagadale, Amar Khatmode & Hajrat Patil. (2017). IOT Based Industrial SCADA System. International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395 -0056, p-ISSN: 2395-0072 Volume 4 Issue 4, 2432-2435*

12. *Mahmoud Shaker Nasr & Ali Najim Abdullah. (2017). Design and Implementation of IoT Cloud Moveable SCADA Supported by GSM for Industrial Applications. Journal of Babylon University Engineering Sciences No. (2) Vol. (25). 409-424*