POLITECNICO
MILANO 1863

Industrial and Information Engineering

Laurea Magistrale in *Electrical Engineering*

Conception, Development and Evaluation of a Data Management System for Digital Twinning of Power Systems

Supervisor:	Prof. Dirk Westermann
Co-Supervisor:	Christoph Brosinsky
POLITECNICO DI MILANO Coordinator:	Prof. Alberto Berizzi

Name: Aitizaz Umer

Personal Code: 10578013

DEDICATION

I dedicate this work to my family; specially my parents, wife and children who have sacrificed a lot and encouraged me to pursue higher education.

ACKNOWLEDGEMENTS

I would like to express my deep gratitude to my Supervisor **Dr. Dirk Westermann**, my Co-Supervisor **Mr. Christoph Brosinsky** and my POLITECNICO DI MILANO Supervisor **Dr. Alberto Berizzi** for their patient guidance, enthusiastic encouragement and useful critiques of this research work.

I would like to specially thank my Co-Supervisor **Mr. Christoph Brosinsky** for his advice, continuous guidance and assistance in completing this work. It was his sincere guidance and critique that persuaded me to work harder.

I would also like to thank TU ILMENAU for letting me in as an **Erasmus Exchange** student. Finally, I wish to thank my parents for their support and encouragement throughout my study.

Table of Contents

List of Tables

Table of Figures

ABBREVIATIONS

EMS	Energy Management System
SCADA	Supervisory Control and Data Acquisition system
WAMS	Wide Area Management/Monitoring System
DT	Digital Twin
CIM	Common Information Model
DBMS	Database Management System
RDBMS	Relational Database Management System
RTU	Remote Terminal Unit
PMU	Phasor Measurement Unit
IED	Intelligent Electronic Device
AI	Artificial Intelligence
HMI	Human Machine Interface
IOT	Internet of Things
IMM	Information Model Manager
UML	Unified Modelling Language
CC	Control Center
PDP	Phasor Data Processor
PDC	Phasor Data Concentrator
LFC	Load Frequency Control
AGC	Automatic Generation Control
HVDC	High Voltage Direct Current
FACTS	Flexible AC Transmission Systems
AVR	Automatic Voltage Regulator
PSS	Power System Stabilizer
VSC	Voltage Source Converter
SVC	Static Var Compensator

TSCS	Thyristor Controlled Series Compensator
SSSC	Static Synchronous Series Compensator
ORM	Object Relational Mapper
OPC	Open Client Communication
EPRI	Electric Power Research Institute
IEC	International Electrotechnical Commission
XML	Extensible Markup Language
XDF	Extensible Data Format
SQL	Structured Query Language
NoSQL	Not Only SQL

ABSTRACT

Reliable access to electrical power plays an important role in the development of any country because it helps in setting stronger industries that fasten the rise of the middle class. In the recent past, electric power industry has experienced an exciting digital transformation. This digital transformation is a result of technological and macroeconomic forces. It has brought about new opportunities and challenges for the power industry.

A digital twin is one of such opportunities offered by digital transformation. The idea of the digital twin was presented by Dr. Grieves in 2002 and he defined digital twin (DT) [1] as "A set of virtual information constructs that fully describes a potential or actual physical manufactured product from the micro atomic level to the macro geometrical level". Digital Twin is a close to the real-time digital image of a physical asset/object or a process which helps in improving its performance. Digital twin may also be defined as an evolving digital copy of an asset based on current and previous behaviors of the asset. The creation of the digital twin of a power system will help in improved power system operations, reduced interruptions and add new business models to generate revenue.

By creating a digital representation of a physical object, how the object is performing and how it will behave under different circumstances, can be visualized. So, by creating a digital representation of the power system, there will be a need to process a lot of data. Therefore, data management is one of the key challenges of the digital transformation of the power system because of the increasing data. Data management is difficult because the data is usually increasing with time, incomplete in nature, difficult to merge, heterogeneous and arrives at different rates. The data accumulated by digital twin from different sources produces a picture of past, present and future of power system components and power system as a whole. Also, data represent diverse views of power system such as over a whole substation or load at feeder level. A system needs to merge all data sets to provide a common view [2]. A Database management system (DBMS), if based on the Common Information Model (CIM), is one such type of system that presents a common view of power system for different utilities. Because CIM makes it easier to make communications between different utilities. Database management systems can store this diverse data and help in the processing of data.

A database management system is one of the key components of a power system control center. With the evolution of power system control centers from SCADA based IP/TCP based communication to PMU based control centers the database management systems have also evolved [3]. Control centers have moved from the strictly centralized to distributed hierarchy. The future control centers will be based on a fast data acquisition system employing the synchronized phasor measurement units, distributed data acquisition and processing [4]. Increasing applications of IoT, machine learning and Industry 4.0 has urged that innovation may also be introduced in the power system control centers. Increasing use of sensors and intelligent electronic devices in the power system may pave the way for the introduction of digital twin concept in control centers. As the core of power system control center is considered to be the, DBMS, therefore a suitable DBMS should be designed and implemented. Considering the benefits, such as increased reliability and efficiency of power system, of employing digital twin in power system control center, the DT may revolutionize the future control center technology, therefore, a novel data

management system is the need of the hour which fulfills the needs of next-generation power system control centers.

After careful considerations and comparisons between graph and relational databases, a relational database schema is chosen for the implementation of this thesis task. The huge amount of power system data and the complexity of equipment models is a challenge in the employment of digital twin in power system. A balance has been maintained while choosing the models of power system equipment such as transformers, transmission lines and regulators so as not to over complicate things while maintaining necessary information to make a digital image of a real power system. A Python-based database configuration script has been utilized for setting up a configurable database schema in a digital twin centered database management system. A PostgreSQL based database has been designed and adapted to the needs of the Digital Twin database. Checking data integrity is vital for the correct functioning of proposed DBMS, therefore, a cross-check validation algorithm for checking data integrity is also employed. To make communications easy between the utilities, PostgreSQL database schema is based on the Common Information Model and the database schema has been translated into shareable XML files.

1. INTRODUCTION & MOTIVATION

Recent technological advancements in the field of communication technology and the availability of intelligent electronic devices dictate that the innovation in the database management system for power system control centers based on the digital twin is the need of the hour. A digital twin may be defined as the digital image/copy of a real object/asset. It receives inputs from the real object and tries to store all information about it in a database. In this way, it not only models the real object accurately in a digital form but also possesses the ability to predict future states of the object thus it can also give a glimpse of future thereby predicting the performance of object in future and also any possible failures. For example to make the digital twin of an aircraft a lot of smart sensors need to be employed on the aircraft reporting all its design, geometry, operational data to database that process that information in a way that the complete model/structure of a real aircraft is replicated in digital form. Digital twin of an aircraft would be able to predict future component or engine failures based on the amount of data previously accumulated. More about digital twin is presented in the next sections of this thesis. As digital twin has to process and store large quantities of data, therefore, a suitable database management system needs to be realized. Therefore, a database management system (if designed properly) will be capable of meeting the needs of digital twinning of power system thus it will greatly improve the power system reliability, provide more monitoring and control actions to the operators as will be discussed in this work.

According to [4], the central nervous system of a power system is believed to be the control center because it is able to sense the pulse of the power system, adjusts its condition, coordinates its movement, and provides defense against exogenous events. The first generation of control centers employed digital computers for rapid checks on the stability and safety of system elements. This computer-based control center, called the Energy Management System (EMS), achieved a quantum jump in terms of intelligence and application software capabilities. An energy management system (EMS) is a system of computer-aided tools used by operators of electric utility grids to monitor, control, and optimize the performance of the generation and/or transmission system [4].

Early control centers used specialized computers offered by vendors whose business was mainly in the utility industry. The traditional control centers were mainly centralized in nature however at present and in the future most control centers are thought to be decentralized. Distributed control centers consist of [4]:

- Separated supervisory control and data acquisition (SCADA), energy management system (EMS), and business management system (BMS)
- IP-based distributed SCADA
- Common information model (CIM)

Information and communication technologies have revolutionized various fields along with the field of power systems [4]. The centralized SCADA systems of the past are being transformed into decentralized systems. Past control centers based on SCADA used centralized personal computers but with the increase in data due to more and more sensor-based monitoring (Phasor Measurement Units and Intelligent Electronic Devices) these centralized control centers cannot meet the demands of current and future data. Future control centers can be built utilizing Web services and grid computing architectures. Grid computing is a processor architecture that combines computer resources from various domains to reach a main objective whereas Web services defined as a special type of Service-oriented architecture that can efficiently operate on Web utilizing XML protocols. An application that can be accessed with the help of a programmable interface is known as a Service. In grid computing, the computers on the network can work on a task together, thus functioning as a supercomputer.

Future control centers are characterized by [4]:

- a fast data acquisition system;
- greatly expanded applications i.e. dynamic security assessment and better visualization tools;
- distributed data acquisition and data processing services;
- distributed control center applications expressed in terms of layers of services;
- standard Grid services architecture and tools to manage ICT resources.

With the introduction of a lot of sensors (PMU\Remote Terminal Unit) for monitoring and control, usage of distributed and renewable energy sources and deployment of advanced communication networks, electric power grids have become quite complex and stressed. Keeping this in mind the current power system controls slightly lack in addressing the grid requirement for renewable energy sources and customer demand. To address these issues a large-scale power system control center needs to be implemented that can solve the problems inherent in the evolution of modern power grids. This large-scale power system control must implement distributed intelligence (It refers to separating the processing in a large system into multiple subsystems) on vertical and horizontal levels [5]. Figure 1 below shows the distributed intelligence architecture.

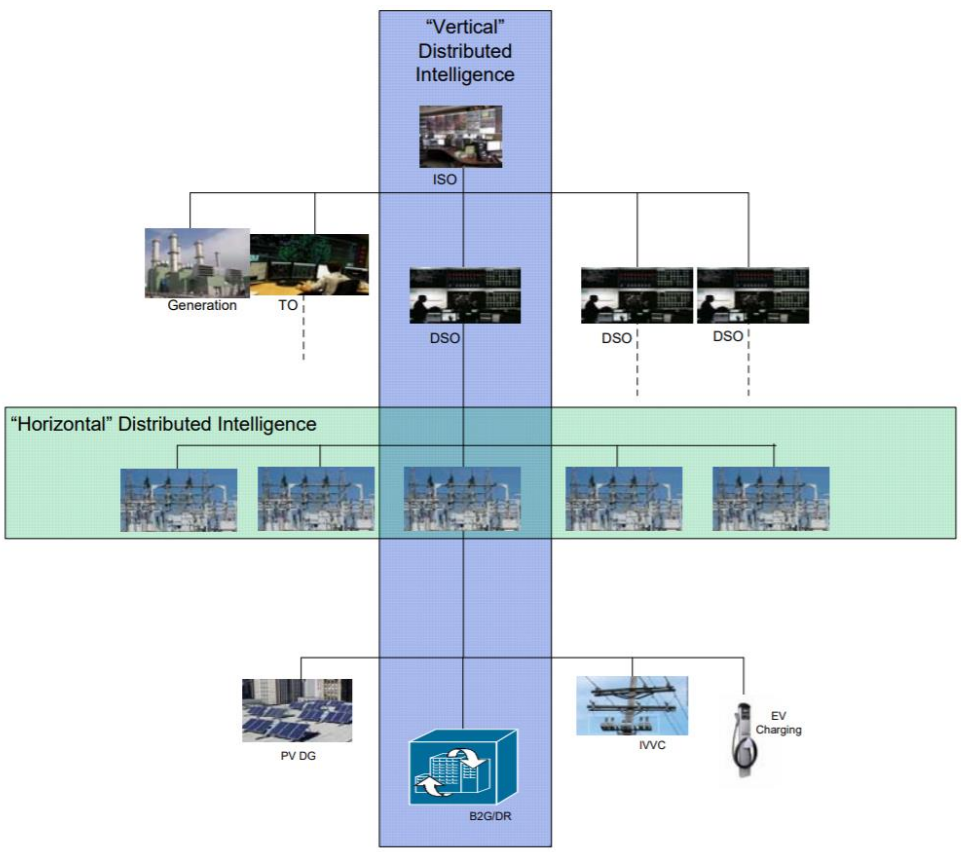

Figure 1 Distributed Intelligence [5]

According to [5] communication network techniques such as IP protocol suite may help in implementing distributed intelligent control systems. This distributed control approach provides for decomposition of large complex control problems down into a series of smaller sub-problems. Because of the huge amount of data generated at each substation, not all the data can be sent to one central location. Phasor data has also imposed further constraints, given a control center with hundreds of PMUs installed, archived sensor data can accumulate to the order of terabytes, within only a 30-day period. Therefore, it becomes extremely inefficient to utilize a single centralized coordinator to collect and archive all the available data from corresponding balancing areas. One approach to alleviating the burden on computing resources has been to distribute the computation across sub-areas [6]. As can be seen in Figure 1, distributed intelligent systems use software and physical agents operating autonomously so that these systems can achieve more flexibility. As a consequence of this decentralized or distributed approach, a need arises for storing the data at various levels [4]. The DBMS for digital twin can be based on this distributed control approach where we decompose large and complex power system networks into smaller problems. This can be achieved by following the Common Information Model and IEC guidelines where a lot of abstract classes are used just to breakdown the complex power system into a simple and easy to understand/manage system.

Today Artificial Intelligence (AI) is finding applications in all walks of life and it can also help in monitoring and control of power systems. Also, it may help in providing clean, affordable and reliable energy. AI can support the use of drones for inspecting generation, transmission and distribution networks. Thus, AI may help in recognizing malfunctioning equipment and pinpoint places requiring repair. Drones coupled with AI will provide 1000's of images of any equipment so more the data the better it is for training algorithms for identifying faulty equipment. Artificial intelligence is going to play a key role in the development and operation of integrated grid control architecture combining centralized and distributed resources. Real-time integration of diverse power system technologies heavily rely on data and communication and AI tool may help here as they process data and react quickly. EPRI is currently exploring a lot of potential applications of artificial intelligence in power systems [7].

Internet of Things (IoT) provides many advantages to the power utilities in the highly competitive modern market. IoT is a huge interconnection of devices (can be cell phones, washing machines, lamps, asset components i.e. power generation unit, measurement devices or engine of an aircraft) to the internet by means of an on/off switch. The idea is that in future a very huge number of devices/gadgets will have many smart sensors associated with them through which the will be able to communicate with each other using the internet as a link [8]. As is expected a large number of interconnected devices are going to share huge volumes of data therefore among many others two major advantages of IoT concept in the field of power system are the ability to collect near real-time data and the availability of new sensing devices.

IoT is further enhanced by combining it with AI technology. IoT is one of the drivers behind the concept of digital twin as digital twin is basically based on the advanced IoT sensors. Without these sensors realization of digital twin is not possible because digital twin needs inputs from the real asset in order to mirror it and only the smart sensors can provide those large measurements to DT. The IoT and wireless sensors make it possible to "sensorize" selected pieces of power system equipment. Doing so enables power system utilities to collect and integrate real-time operational data from the equipment and integrate this with enterprise digital information. Such integrations have paved the way for the digital twin. IoT devices/sensors are one of the core enablers of the digital twin concept. A digital twin is a virtual representation of a physical asset that is virtually indistinguishable from its physical counterpart. It includes design and engineering details that describe its geometry, materials, components, and behavior or performance [9]. The digital twin concept is built on large amounts of cumulative and real-time operational data measurements across an array of dimensions. These measurements can help create an ever-evolving digital profile of the asset that may provide vital inputs on system performance leading to actions in the physical world. Non-availability of information from any of the data sources affects digital twin accuracy greatly. So, to accurately model a power system, we need more and more data about it, the realization of a digital twin based on the AI-enhanced IoT technology greatly helps in this regard [10].

In the past and existing control centers consisting of SCADA, Geographic Information System (GIS), Outage Management System (OMS) and EMS, a lack of common model or common schema is present because each utility maintains network model and database schema according to its needs thereby utilities maintain their own multiple master copies of data that meets their needs. From a storage point of view data obtained from the new sensors (smart meters/PMU) may be in Terabytes after a single year of operation. From a velocity point of view, a typical PMU can produce 50/60 samples per second [11]. This high amount and velocity of data are overwhelming for any database. However, storage and processing of such large amounts of data is not an issue with modern drives having capacities in Zettabytes and processors having amazing computing capabilities. The need for large scale real-time computing, communication, transfer and storage of data generated by intelligent grid technologies is expected to be addressed by cloud computing services. Utility companies can use cloud model to store and process large quantities of data collected from smart meters and appliances, as well as sensors deployed across the power grid. The primary reason is that cloud data centers can accommodate the large-scale data interactions that take place on modern grids and are better architected than centralized systems to process the huge, persistent flows of data generated across the utility value chain [12]. An important issue as mentioned previously also is the lack of a common data model for storing, retrieving and exchanging of data among applications within the utility and among other utilities too for better grid planning, operation and control due to interconnected grid systems. To mitigate this problem a database management system must be realized keeping in mind the Common Information Model guidelines [13].

Data is the focus of modern power grids. Utilities exchange a lot of data between different software systems which enable them to properly operate, plan and maintain the grid. A typical utility spends a lot of time and resources to manually maintain, update, and exchange information amongst different systems. Inconsistencies during data exchange and even the lack of data exchange can lead to dramatic consequences like suboptimal system performance and even system-wide blackouts. In this era of digital world, data accuracy, model complexity and automation are the foundation to maintain optimal performance and maximize future investments. This is where digital twin of a power system comes in handy as it provides utilities with a single source of truth for data exchanges [14]. Digital twins will greatly enhance manufacturing processes, maintenance and creation of connected products in the future. Digital twins will enable asset monitoring, optimized maintenance and reduction of costs in the future power system control centers. Digital twin is a virtual replica of physical object/asset and it will help to analyze all scenarios of the physical system thus helping us in detecting any issues. The digital twin concept will enable organizations to better understand, predict and optimize the performance of their installed assets [15].

A database management system capable of handling the needs of future power system control centers has been designed using the PostgreSQL which possesses the advantages of improved security, enhanced scalability, improved performance, robustness and ease of use.

1.1. The Database Management System in a Control Center

A database is, from a logical point of view, a model of the knowledge structure of the power utility, as it exists to support the operating and planning activities. The database is a common body of data that is used by many different SCADA/EMS functions. The database is one of the important components of a control center. All the data required by the control center software and by the operator are located in the database. In a power system environment, a significant proportion of the data requires frequent updating. It is often necessary to expand an existing database. Snapshots of the data at a certain time are needed to ensure that the data used by the application programs are consistent [16].

SIEMENS has introduced the Spectrum Power Information Model Manager (IMM). "The Spectrum Power IMM is a powerful desktop data modeling, data maintenance and data exchange tool specifically designed to manage all power system model data for a variety of supervisory control and data acquisition (SCADA), communications, power applications and other enterprise information". Using IMM information related to power system models can be generated and maintained in a single central location. IMM serves as the centralized data storage and maintenance platform for operational and all other types of data required by the utility [17].

In a power grid, the main function of control center is to monitor and control the transmission network in real-time to produce warnings in advance for securing the power grid. For this purpose, control centers facilitate cooperation among several specialized groups to optimize the transmission operation by collecting and analyzing the operational information. This helps in decision making and ensure reliability, security, flexibility and efficiency of electric power network. All of the above-mentioned tasks heavily rely on accurate and adequate data; therefore, a database management system plays a key role in the efficient and intelligent power system dispatch [18].

Database management includes collecting, analyzing, storing and providing data to users and applications. Database management system has to deal with data identification, validation, accuracy, consistency and updating. Database management system not only deals with model data (including static parameters) and measurement data but also operational, security and power system configuration data. Different specialized people and group of people (i.e. dispatcher, security check engineer, planner and protection engineer) in a power grid operation have different responsibilities and are interested in the different aspects of power system network model. For example, for real-time monitoring and control current network model is preferred but for the security checks, stability calculation model including detailed generator, excitation system and governor is used. If each group makes the model according to their own demands, this will result in a lack of common network model, unified parameters and operational data. To curb this situation, the emphasis is made on utilizing the common information model architecture [18].

The fundamental requirements of a database management system [18] are listed below:

- To satisfy the requirements of different specialized groups to provide past, present and future network models for specialized analysis.
- To save fundamental network model in different forms i.e. XML files and relational database.
- To accommodate all data sources which may include legacy system, substation configurations and other application systems.
- To present the complete view of whole power system model.
- To be stored in distributed databases but managed in a unified way.

Power system models can be stored in different ways in a database management system. Some of these ways include storing models in a relational database, time-series database, file and memory database depending upon the models and use case. Relational databases are used to store network parameters of real-world power network models. Time-series databases are used to store measurements and memory databases are used to store future virtual models based on the current models and parameters [18]. Generally, a database management system in a control center can be implemented through either relational or graph databases. A more detailed analysis on the relational and graph databases is discussed in the third chapter while a discussion on database management systems in a power system is made in the second chapter.

Some important attributes of a modern DBMS are listed below [16]:

- **Integrity:** data must be accurate and consistent.
- **Resiliency:** data must not be lost or destroyed. Errors must be detected and corrected.
- **Upgradability:** the relationships between and among data must be well understood, documented and defined. A DBMS must support the creation of new relationships.
- **Shareability:** support software must provide simultaneous sharing of data resources while optimizing their use.
- **Security:** the architecture must provide appropriate security controls to defend the system against malicious actions.
- **Independence**: the DBMS should permit the separation of programs from the 1ogica1 and physica1 data structure.
- **Accessibility:** this attribute means the facility to handle the data base by non-computer orientated people.
- **Flexibility:** a database should be tuned as the requirements change.
- **Administration**: the DBMS includes tools that allow the database resources to be managed.

1.2. Objective

The objective of this thesis is the conception, development and evaluation of a data management system, which fulfills the needs of the next generation of power system control centers. To make a database management system able to handle future requirements of power system control centers a PostgreSQL based database will be utilized. This database in combination with PowerFactory as the data processor, Matlab as the database front end and OPC server as a tool to mimic SCADA/WAMS will act as a database management system test setup for digital twin-based power system control center. To make the exchange of power system data possible between the different departments of a utility company and between different utilities, XML based files will be generated according to the CIM-UML and IEC guidelines.

1.3. Thesis Overview

Chapter 1 provides the introduction and motivation for the thesis. Chapter 2 describes the state of the art and contains information about the existing power system control centers and future trends. Chapter 3 presents a suitable database structure which contains information about the appropriate database structure required for future power system control centers based on digital twin. It also describes power system data, models and parameters. It gives information about power system component models and related parameters. It also gives an introduction of different types of data in a power system. Chapter 4 deals with the implementation of a database management system and outline how to implement the database suitable for digital twin applications. It also describes the XML format appropriate for communication purposes in the utilities. The conclusion of the thesis work is given in chapter 5 and chapter 6 deals with future recommendations.

2. STATE OF THE ART

International Organization for Standardization (ISO) defined the data as “a representation of facts, concepts or instructions in a formalized manner suitable for communication, interpretation or processing by human beings or by automatic means”. Whereas, the definition of the Oxford English dictionary provides a more detailed definition of database as “a structured collection of data held in computer storage; especially one that incorporates software to make it accessible in a variety of ways” [19]. Considering the definition of European Union Database Directive “A database shall mean a collection for information purposes, in a fixed form, consisting of independent works, data or other materials, arranged in a systematic or methodical way so that they are individually accessible by electronic or other means” [20].

Data is stored in a database management system in such a way that it makes it simple to restore, manipulate and create information. Database Management System or DBMS relates to technology of storage and retrieving user's data with great efficiency along with suitable safety measures. DBMS enables its customers to build their own databases according to their needs. These databases are extremely configurable and offer a range of choices. Database management systems are intended to facilitate the management of databases. The database management systems are used to record, store and manage the data in the database [21]. A DBMS based on the digital twin should be able to remove any flaws in the existing DBMS’s such that it should be able to react quickly to the high rates of data provided by PMUs and other smart sensors.

The Power Grid utilizes digital components and real-time communications technology installed on a grid to monitor the grid electrical characteristics which are self-tuning to ensure that it operates always at an optimized level. Owing to the introduction of Intelligent Electronic Devices (IED's) and smart sensors for monitoring and control of different power system components, the power grids are now capable of automatically looking into potential issues created by storms, catastrophes, human error or even sabotage [22]. With the growth of digital devices and real-time communication, the quantity of data is quite large, and it needs a modern database management system to manage data efficiently. Reliable data management is the main factor for the successful operation of the power system [2].

Due to the incomplete, heterogeneous and different rates of arrival of data, management of data is, therefore, one of the major issues in power systems. Data transformation is needed for effective use in data analytics and other related activities. Only when the data is stored continuously in a database, data transformation is possible. In the past, various database systems, including real-time SCADA-based database management systems and relational database management systems were used [23]. The ongoing transitions and improvements in sensing, measuring, communication and computing have now entered an era in which the digital data acquisition architecture which served for over half-century to monitor and control the power grid needs important changes to fully utilize the new technologies and realize the promise of a new generation of applications [24].

2.1. Common Information Model

These days, there is a lot of data exchange in Power system utility and between different companies due to the bidirectional nature of data. Operators give instructions and receive measurement data from different sources in a power system. Requirements of data exchange are increasing with higher renewable energy sources integration and introduction of smart grids; therefore, Common Information Model (CIM) standards are also evolving with time [25]. The Common Information Model (CIM) was developed by Electric Power Research Institute (EPRI). The CIM defines most of the main objects in the electric utility enterprise based on object-oriented data modeling technology. It is aimed at covering a standard data model for the energy management system (EMS) and distribution management system environments to facilitate data exchange. The CIM defines a comprehensive logical view of the power system elements and represents their attributes as well as the relationships using a set of Unified Modeling Language (UML) Class Diagrams. The Common Grid Model Exchange Specification (CGMES) is an IEC technical specification (TS) based on the IEC Common Information Model (CIM) family of standards [26]. The CIM for grid model exchange enables exchanges for the data necessary for regional or pan-European grid development studies, and for future processes related to network codes. Grid model exchange is a complex process covering a variety of use cases, which include the exchange of:

- Equipment information, which contains power system equipment data; Topology information, which contains topology related information for the grid elements;
- Information on power system state variables, which contains the results from initial load flow simulation of the system; and/or
- Steady-state hypothesis information, which is valid for newer standards and provides information on load and generation values as well as other input parameters necessary to perform load flow simulations.

The deregulation of the power industry has resulted in multiple utilities, running software from a number of different vendors, having to exchange large data sets on a regular basis. The use of proprietary, custom formats complicates this exchange, requiring complex translation between each of the custom formats. Problems occur when companies need to exchange data between software applications from different vendors, and/or have multiple versions of the same software running within their company. Such a scenario requires a company to either [26]:

1. Maintain multiple copies of the same data in multiple formats
2. Store the data in a format compatible with every piece of software, requiring the removal of application-specific data and a subsequent loss in precision
3. Store the data in a single, highly detailed format and create software to translate from this highly-detail format to the desired application file formats
4. Use a highly detailed format that is compatible with every application and whose standard format contains the basic data required to represent the power system while simultaneously allowing additional, detailed, application-specific data to be contained without invalidating the format.

The CIM is an extensive model that includes such diverse classes as Assets, Work, Customers, Metering, Wires, Generation, Load Model, Outage, and Protection. Each class contains properties and attributes; for example, the Metering class defines meter registers, reading values, service point locations, and so on as well as relationships to other classes such as Customers and Load Control [13]. The objects represented in the CIM are abstract in nature and may be used in a wide variety of applications. The use of the CIM goes far beyond its application in an EMS. This standard should be understood as a tool to enable integration in any domain where a common power system model is needed to facilitate interoperability and plug compatibility between applications and systems independent of any particular implementation. The IEC 61970-301 standard defines the CIM base set of packages which provide a logical view of the functional aspects of an energy management system including SCADA [27].

Today there is a need for common data exchange format for exchanging data between utilities and related companies. Common information model is a common data exchange format for energy markets of present and future. A central CIM based database is used for large amounts of common data sharing between applications and companies. The central CIM database schema is translated from the CIM specification. The database should reflect the relationship, such as aggregation and association in CIM/Unified Modelling Language (UML) diagram. Data can be mapped from relational database using SQL approach to CIM/XML and vice versa. For convenience, conventions and procedures have been developed to translate the CIM specification into a workable, “real-world” relational database and vice versa [23].

The Common Information Model (CIM) for Power Systems has the potential to maintain multiple copies of the same data in multiple formats while the eXtensible Markup Language (XML), combined with the Resource Description Framework (RDF) offers a means of storing the data in a format compatible with every piece of software, requiring the removal of application-specific data and a subsequent loss in precision [27].

2.2. Power System Control Center

The power system control center has been traditionally the focus for the interconnected electric transmission and generation system utilizing different sensing components. The control center offers the required functions to monitor and coordinate the physical and economic operating of the power system minute by minute. Significant coordinated decision-making is required to maintain integrity and economy of the interconnected power system. The main task of the Control Center, therefore, is to regulate and control the interconnected grid's physical operation [28].

Following are the typical components of a Control Centre (CC) according to [28]:

- Energy Management System
- SCADA System
- Communication Network

Communication networks are usually considered in the SCADA system and connect SCADA to EMS. The monitoring components of SCADA systems are usually positioned in substations and consolidated into what is called remote terminal unit (RTU). RTUs are generally fitted with microprocessors having memory and logically capable to provide communication links with the Control Centre, along with some form of telemetry. Relays within the RTU, on the command from the CC, open or close selected control circuits to perform a supervisory action. Such actions may include, for example, opening or closing of a circuit breaker or switch, modifying a transformer tap setting, raising or lowering generator MW output or terminal voltage, switching in or out a shunt capacitor or inductor, and the starting or stopping of a synchronous condenser [28]. This power system control center architecture is explained in Figure 2. Control centers have traditionally been based on SCADA, but currently, the use of PMUs has increased a lot due to their high sampling rates which helps in the power system protection and dynamic behavior analysis. Currently, both RTU and PMU systems are being used, therefore both WAMS and SCADA systems are available at the control center and act in parallel [3]. In a control center, phasor data processor (PDP) processes PMU measurements. Phasor data concentrator receives phasor measurements from PMU's. IEEE C37.118 standard data exchange communication protocol shall be supported by a PDC. A PDC should also be capable of exchanging data via standard protocols such as IEC 60870-5-101/104, DNP3, etc. This connectivity is intended to send data to a control center (SCADA / EMS systems) and to receive command from a control center [29]. WAM, WAC and WAP stand for wide area monitoring, wide area protection and wide area control respectively. They receive phasor measurements from phasor data processors and send them to the human machine interface (HMI). The following sections of this chapter provide a more thorough description of these parts.

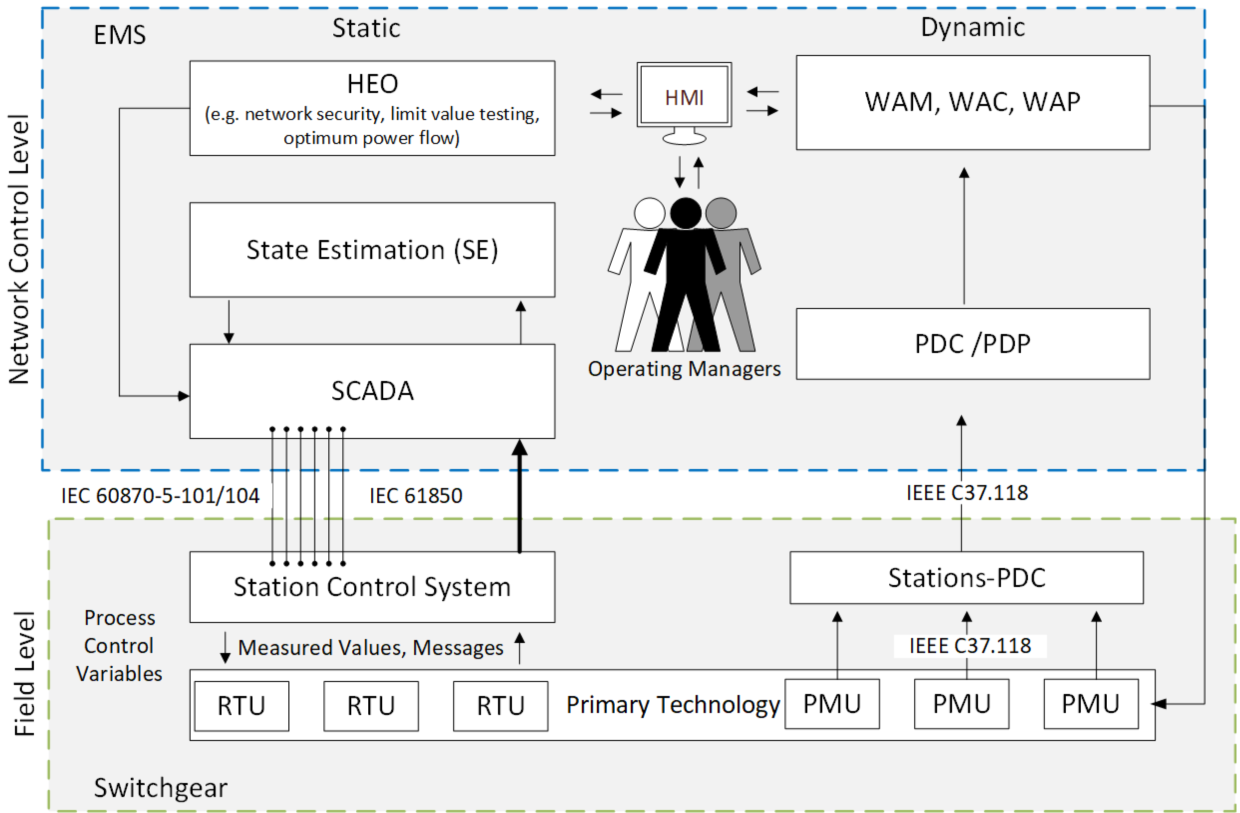

Figure 2 Power System Control Center Architecture [30]

SCADA requires telemetry as a communication form. Telemetry is the measurement of a quantity in such a way so as to allow interpretation of that measurement at a distance from the primary detector Telemetry may be analog or digital. In analog telemetry, a voltage, current or frequency is created and sent to the receiving location in a communication channel in accordance with the amount measured. The earliest type of SCADA telemetry signal circuit consisted of twisted-pair cables; although easy and cost-effective, they have reliability issues due to breakages, water ingress, and potential grounding risks in case of defects. The leased wire systems then arrived and recent systems have been using microwave radio, fiber optic technology and satellites-based VHF/UHF systems [31].

For the transmission of data, SCADA systems use protocols. Modbus RTU, RP-570, Continel and Profibus can be included in these SCADA protocols. SCADA protocols are usually vendor-specific. The standard protocols IEC6870-5-101 (also known as T101) and 104, IEC61850 and DNP3 are also commonly adopted and used. Almost all main suppliers recognize these standard protocols. Improved version of these controls includes TCP/IP communication. In Europe and America, the T101 and DNP3 are used respectively [32].

2.2.1. Energy Management System

The EMS is a computer-assisted instrument used by electric utility grids operators to track and optimize the performance of the generation and/or transmission system. It adopts advanced technology of network management, databases and visualization. The Energy management system is one of the main programs in a control center [33]. The Energy Management System (EMS) has its importance in the need for electric utility companies to run their systems economically without compromising on the system security. In the past, most utilities bought their EMS from one or more EMS suppliers. To operate the system as cost-efficiently as possible, all producing units usually had their features in one place, with the aim of providing the most efficient units dispatched properly together with the less efficient. The scheduling of generators with limited fuel or water supplies were also incorporated in energy management systems. This allowed utilities to further reduce the cost of operation by taking advantage of cheaper fuels or hydropower [28]. Today this is mostly left to market to decide what EMS meets their demands and how they want to operate but the security and reliability of system is desired more than all the other measures.

EMS consists of these major functions, some of which are at the market discretion; however, there can be many more functions available in individual EMS vendor platforms [28], [31], [33]:

- Network Model Building
- Security Assessment
- Automatic Generation Control
- Automatic Voltage Control
- Economic Dispatch
- Reserve Monitoring

2.2.1.1. Network Model Building

To verify that working circumstances are secure under the current topology as well as in case one or more parts fail and are out of service, a network model is essential. The network model must represent the right topology and working conditions in comparison with the current network circumstances. The data available for building the network model contains the status indicators and the analog measurements available from the SCADA. The outcome of the network model builder is a model of power flow. The fundamental tasks of network modeling are to identify a network model, to set up a grid model database and to define the physical network framework for software for network analysis. Two steps are used to build network model [31]:

- Topology Processing
- State Estimation

2.2.1.2. Security Assessment

The security assessment determines first whether the system is in an appropriate state currently and then whether the system would react in an acceptable manner and reach an acceptable state following any one of a pre-defined contingency set. The rest of the power system will have one element less if a circuit breaker is damaged i.e. so, the general power system will weaken, and the remaining system may go through undesirable effects. This is a static security assessment; another sort of security assessment is known as dynamic security assessment [28]. Dynamic security assessment provides fast assessment of single and even multiple contingencies and delivers results corresponding to operational irregularities [34].

2.2.1.3. Automatic Generation Control/Load Frequency Control

One of the key tasks of a Transmission System Operator (TSO) is real-time control of mismatches between power consumption and production of electric power, i.e., frequency control. Load frequency control (LFC) or automatic generation control (AGC) is one of the main problems in the design and operation of electric power systems for excellent quality and reliable electric power supply [35]. System frequency is an excellent indicator of the active power balance of the grid power system. The frequency deviation is proportional to the power imbalance.

Automatic Generation Control (AGC) can determine the true power output for interchange and frequency maintenance of all generating units in some systems. The fundamental energy system generation control is achieved by controlling the electrical frequency in one of the system's high-voltage buses and interconnection lines. If the frequency falls below the nominal, the generation needs to be increased and if the frequency increases above the nominal frequency, the generation needs to be decreased [28].

The European Union (EU) has taken several measures through a number of different directives to liberalize the electricity market. The liberalization of power markets has caused major changes in the production and use of electricity. In a traditionally monopolist and conservative sector, liberalization developed competition to allow fresh company models to arise and fresh players to challenge incumbent utilities. Competition has brought great efficiency and also have brought great challenges for market players. In November 2016, an updated Electricity Market Regulation, revised Electricity Directives, new Risk Preparedness Regulations, and new EU Market Design Regulations were issued by the Commission. These regulations ensure energy markets include more renewables, empower consumers and efficiently manage energy flows across the EU. In March 2019 EU commission passed a new regulation which insures clean energy for all Europeans. These regulations guarantee that power markets include more renewables, empower customers and handle power flows throughout the EU effectively. A new Regulation ensuring clean energy for all European citizens was proposed in March 2019 by the EU. This new rule represents a significant step towards the clean energy transition for the EU and its Member States. The EU intends to minimize carbon output while preserving its worldwide competitiveness and growth rate through this new regulation. The new regulations aim to enable energy customers to play an active part in driving the shift to energy and take full advantage of a less centralized, more digital and durable power system [36].

2.2.1.4. Automatic Voltage Control

Automatic Voltage Control is an automatic optimal control of network voltage/reactive power (Volt/var) that network losses are reduced as much as possible to improve economic operation while ensuring voltage quality, and in the meantime the amount of operating equipment is concentrated to the minimum to reduce control costs and fault probability [28].

2.2.1.5. Economic Dispatch

Almost all of the control centers in the past used economic dispatch calculation before the energy industry was restructured to determine setpoint power levels of all generators to meet demand. In some areas of the globe, such a system still exists. In this calculation, the cost of generation in the computer system is shown as curves, usually in part as quadratic, and the overall calculation reduces operating costs by finding a point in which the generator's total output is equivalent to the total power required, and in which all generator incremental power production costs are equal. In some areas, a market dispatch is made by using an auction system, whereby the optimization algorithm is like financial dispatch except for the generator's cost curves are replaced by generator owner bids [28].

2.2.1.6. Reserve Monitoring

In the past, reserve monitoring was done to calculate the actual available energy capacity in the scheme to fulfill changes in demand. Reserve monitoring (RM) was used to control danger of accidental power grid defects (for example, loss of the biggest unit or the biggest power exchange) [31]. Today however this function is at the market discretion.

As technology progresses rapidly, the wide-spread substation deployment of intelligent electronic devices (IEDs) comes into the scene. These computer-based systems can record and store a large volume of data, both operational and non-operational, depending on the desired purpose. They are much more sampled than the RTUs, and data is synchronously sampled vs. non-synchronous RTUs data. The fast development of computer, communication network, database technologies, and substation IEDs, as well as the new demands of electricity markets, makes developing a new generation EMS highly desirable [24]. Also, due to different design decisions for RTU and SCADA, several data monitoring constraints are present in the current EMS systems.

2.3. Supervisory Control and Data Acquisition System

The term Supervisory Control and Data Acquisition (SCADA) is used for any centralized system that retrieves data from remote locations and, if permitted, may issue control commands. Typically, these SCADA systems are in the utility control center, but they can include a SCADA engineering system that retrieves protection data or disturbance data, or a maintenance SCADA system which monitors both power system and the communication device [37]. Control part of SCADA system offers electrical equipment control and monitoring in an EMS.

The SCADA System's communication media can contain nearly any sort of device, as long as 1-second response time is accommodated. SCADA Systems generally provide system broad data every 2 to 10 seconds on power system hardware. In Italy, the SCADA scan rate is 4 seconds. Typical scan rates are every 1–2 seconds for generation and interchange data and circuit breaker status indications; every 2–15 seconds for line flow and voltage measurements, and every 15 minutes to one hour for energy values. While typically used in real-time generation, transmission and distribution activities, many distinct systems, applications and personnel at the control center can use the data obtained from the SCADA system. In power systems, there are different kinds of SCADA systems. All these SCADA systems do not have to be present on each Power System. This list contains various kinds of SCADA systems based on their use in a power system [37].

- Operations SCADA System
- SCADA System for Data Management
- Engineering SCADA System
- Maintenance SCADA System
- Planning SCADA System

2.3.1. Operations SCADA System

Power system operations SCADA system receives real-time data from power system equipment and issues control commands to power system equipment via [37]:

- **Remote Terminal Units**
  An RTU gathers data from power system by converting analog sensor data into digital data. This digital data is then transferred through a communication channel into a system in the form of status and analog "points." An RTU issues control actions by turning a system's digital commands into electro-mechanical or solid-state action on power system devices.
- **Electronic Devices (present inside substations and along the feeders)**
  The power system device is controlled by an IED or a controller (not quite as smart as an IED). IEDs are basically microprocessor-based controllers of power system equipment. A few examples of IEDs are digital protective relays, circuit breaker controllers and capacitor bank switches. It generally also monitors data from the power system that is important to its possible control activities. It can be situated in substations, along feeders, at client premises, or wherever there are power systems located. The IEDs are connected in a substation via a communication network, such as the LAN.

- **Generation management systems (may also include distributed sources)**
  Generation management systems are used to control, monitor and to increase the efficiency of generation and transmission systems. In the generation management systems AGC/LFC are employed to optimize performance of generating units. Generation management systems manage critical functions such as multi-area/market generation control and dispatch, load forecasting and real-time cost monitoring.
- **Substation masters**
  A substation master collects data from IEDs, controllers, and power system equipment in substations. It could also gather data from distribution feeder device, but this is rarely achieved, as most substation masters are used as part of substation automation systems for large transmission substations. The control commands obtained from other systems can also be passed.
- **Other control centers and manual entry**

2.3.2. Data Management SCADA System

Data Management SCADA System receives power equipment configuration data from devices. They may have their own communication channels to remote locations, or they can obtain this information via the SCADA distribution system [37].

2.3.3. Engineering SCADA System

Engineering SCADA System gets event of sequence data, oscillographic data (Special Handling Required), historical data and statistical data. It has its own communication channels on distant locations, or it may acquire this data through the distribution operations SCADA system [37].

2.3.4. Maintenance SCADA System

Data related to health of the power system equipment and communication equipment are received by the SCADA system maintenance system. They can have their own communication channels to remote locations, or they can obtain this information via the SCADA distribution system [37].

2.3.5. Operational Planning SCADA System

Operational Planning SCADA System receives data for statistical analysis of measurements of the power system: limits, minimum levels, averages, trends, profiles and metrics of power quality, necessary for short and long term planning [37].

2.4. Wide Area Monitoring Systems

Wide area monitoring systems assist operators of power systems to continually analyse all characteristics of a large power network in real-time. Using phasor measurement units (PMUs), information can be recorded and monitored for the purpose of detecting disturbances and improving knowledge of network behaviour under dynamic conditions, allowing system operators to maximize and network stability and security [38]. In power system control centers, WAMS are still emerging tools. WAMS includes measurements from the whole grid, advanced visualizing instruments and modern data analytics capabilities. WAMS can supply complete power system data in real-time in a control center. WAMS have the capability to provide dynamic states of the system in real-time, thus helping in making its digital image.

2.4.1. Phasor Measurement Unit

Phasor Measurement Units also called Synchrophasors are devices that measure phasors synchronized in time. PMUs are the backbone of WAMS, some other components of WAMS are Phasor Data Concentrators and communication links between PDCs and PMUs. With the use of PMUs at various places in the power grid, it is possible to measure voltages and current streams simultaneously. When put on a bus, PMUs include the voltage phasor of all the branches coming from the bus with time-synchronized readings. Therefore, the "phasors" that a traditional PMU usually measure are:

- Voltage magnitude and angles/real and imaginary components of voltages
- Current magnitude and angles/real and imaginary components of currents

In the phasor measurement technique, samples of sinusoidal waveform data are collected each period of the fundamental frequency of power system. In present commercial PMUs, quite high sampling rates are used. PMUs generally provide system-wide data in milliseconds to power system equipment. This provision of timestamping has enabled system operators and planners to measure the state of the electrical system and manage power quality. The use of global positioning system (GPS) satellites, makes time synchronization easier. Since these measurements are synchronized, they can be used in real-time to assess system security. The capacity of PMUs to measure complex voltages and/or currents directly provides PMUs a huge advantage over other telemetry systems, because complex voltages and/or currents are the principal entities on which decisions are made in a power system [39].

2.4.2. Phasor Data Concentrator

One of the main components of WAMS is known as a phasor data concentrator or PDC. PDCs collect and distribute data to other applications. PDC's can be easily integrated into power system monitoring and control platforms. PDC's support different communication protocols such as IEEE C37.118, IEEE 1344 (for communicating with PMU devices and other PDC's) and IEC 60870-5-101/104 or DNP3 (for communicating with SCADA/EMS and devices in substations).

2.4.3. WAMS Applications

Wide area monitoring systems (WAMS) have been widely applied in power system networks throughout the world using synchronized phasor measurement units (PMUs). PMU-based measurements offer new ways to achieve real-time stability and control [39]. For a number of applications in power systems, PMUs have been used collectively called wide area management system (WAMS)-based applications; these can be broadly classified into 4 categories:

- **Model Validation and Parameters Identification**
  The simulation findings from power system models are based on many planning and operational activities such as the identification of working capacities for the transmission line, network extension, relay configurations and integration of renewable generators. It is of major significance to validate and calibrate power system models and parameters as they can differ over time due to either deliberate or gradual modifications. For model validation, PMUs deliver high-quality measurements [40].

- **Wide Area Monitoring**
  PMU data are mainly used in post-event analysis and state estimation for power system monitoring reasons. State Estimation of the power system from real-time measurements is a key element in energy management systems. The state of the power system is defined by collecting voltages of all network buses acquired at the same time. In the past and some current situations, the state estimation technique presently used is based on non-synchronous SCADA system measurements. These non-synchronous measurements coupled with low rates of scanning and slow SCADA calculations are a significant obstacle to providing the precise dynamic states of the power system. PMUs are able to measure the power frequency once per cycle so that the dynamic state of the power system can be correctly predicted [40].
- **Wide Area Protection**
  PMUs are able to assist and protect power system equipment like circuit breakers, transmission lines and generators by assisting the operators to implement corrective measures rapidly due to the rapid scanning rates. The easy availability of synchronized measurements using GPS technology and improved communication technologies makes it possible to consider true, differential protection of transmission lines and cables which was very difficult to achieve before because input signals cannot be synchronized at both ends of the transmission line [40].
- **Wide Area Control**
  The power system control was mainly by local signals before the implementation of PMU's. Feedback controls were used extensively for control devices with these locally accessible measurements. In other cases, control actions were made without actual system measurement on the basis of a mathematical system model. The emergence of phasor measurement allows control based on the measured value of remote quantities [40]. HVDC and FACTS were locally controlled before the arrival of PMU power system controllers such as power system stabilizers.

2.5. Digital Twin Introduction

The Digital Twin has been defined in an array of styles in Literature. One of them is: "The digital twin is a comprehensive digital representation of an individual product. It includes the properties, condition and behavior of the real-life object through models and data. The digital twin is a set of realistic models that can simulate its actual behavior in the deployed environment. The digital twin is developed alongside its physical twin and remains its virtual counterpart across the entire product lifecycle" [41]. An example of an aircraft digital twin has been provided already to understand the concept. So, a digital twin may be defined as the digital image/copy of a real object/asset. It receives inputs from the real object and tries to store all information about it in a database. In this way, it not only models the real object accurately in a digital form but also possesses the ability to predict future states of the object thus it can also give a glimpse of future thereby predicting the performance of object in future and also any possible failures. Concepts similar to digital twin are already present what makes it unique is the integration of IoT (smart advanced sensors), Industry 4.0 and latest technological advancements. Industry 4.0 is known as the 4th industrial revolution. "Industry 4.0 is the next phase in the digitization of the manufacturing sector, driven by four disruptions: the astonishing rise in data volumes, computational power, and connectivity, especially new low-power wide-area networks; the emergence of analytics and business-intelligence capabilities; new forms of human-machine interaction such as touch interfaces and augmented-reality systems; and improvements in transferring digital instructions to the physical world, such as advanced robotics and 3-D printing" [42].

Figure 3 illustrates concept of digital twinning of a real-life asset.

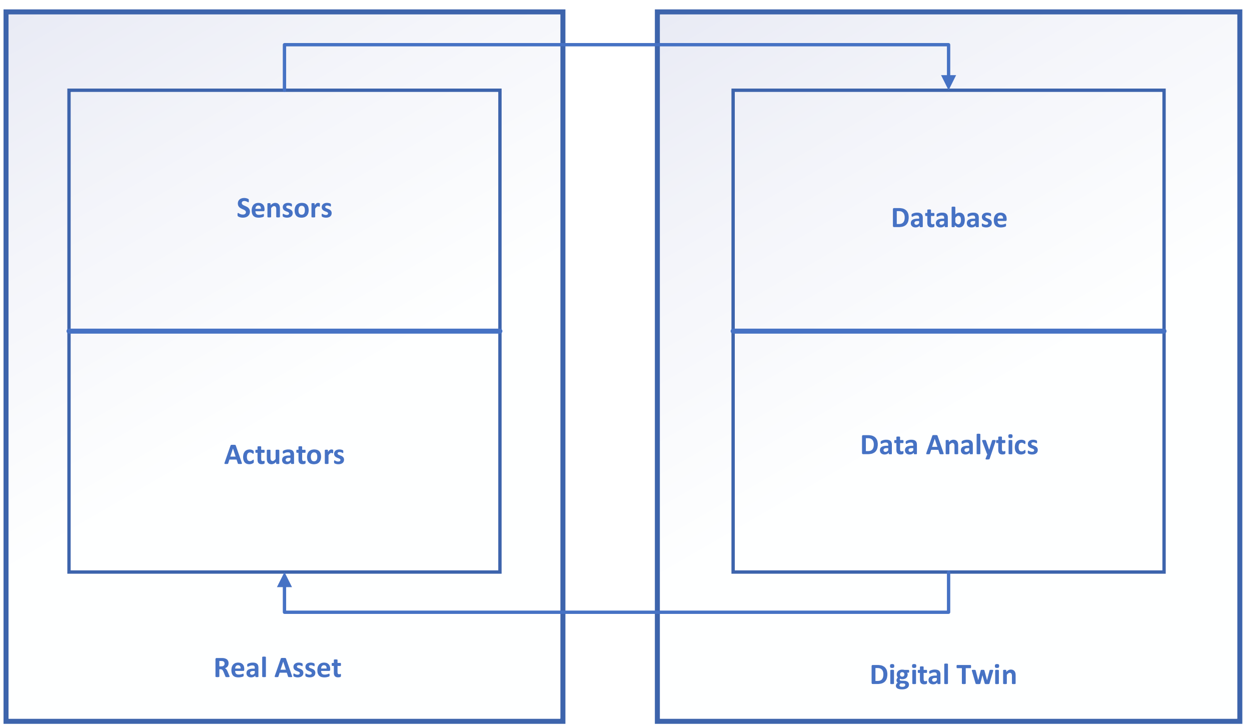

Figure 3 Digital Twin of an Asset [43]

The World-leading tech research group Gartner named digital twins among the "Top 10 Strategic Technology Trends for 2018," predicting the existence of a large number of digital twin applications in near future and estimating that the use of linked sensors and endpoints is estimated to reach 21 billion by 2020 [44].

A Concept similar to digital twin known as device twin is being used by Microsoft in the Microsoft Azure Internet of Things (IoT) as part of their device management solution [45]. Amazon also uses a similar type of idea called the device shadow. "The device shadow is a JavaScript Object Notation (JSON) file containing the state, meta-data, timestamp, unique client token and the version of a device connected to device shadow". In addition, IBM is also working to introduce advancements in the digital twin concept. Industrial suppliers such as GE Predix and Bosch are also working to develop and have developed digital twins for different applications .

In many sectors like construction management, healthcare, intelligent towns, petroleum and gas, intelligent buildings and many more, the Digital Twins idea has great potential. Digital twin predicts the problem before it happens in physical systems through extensive cooperation of artificial intelligence, machine learning, and information analytics. It's like understanding the future and being able to shape it. It is an integrated way of optimizing and monitoring efficiency virtually. Digital twins can assist in addressing cyber security threats in terms of increasing cyber-physical systems and interconnectivity. Corporate IT and operational systems (OT) are exposed through the Industrial Internet of Things to an ever-growing number of external networks and devices and as more assets are remotely monitored, controlled and maintained the need for digital twin systems has increased [46].

2.6. Digital Twin Applications in Power Systems

The worldwide demand for electric power has increased, which brings further pressure on reliable production, which has led to an increased shift to renewable sources of energy that are processed differently than traditional sources [44]. While the trend toward renewables offers energy resource diversity, long-term "cycling" of traditional plants, operating fossil fuel power sources up and down and switching them off and on repeatedly increases the stress on equipment, boosting maintenance costs and increased risk of unplanned power outages. Digital twin technologies allow businesses to test techniques to combine these energy sources more efficiently without risky activities.

The applications of digital twin in the field of power systems are limited until now. However, there are numerous applications of digital twin in the other fields such as in transportation, manufacturing, construction and energy. The most useful applications of a Digital Twin for a high value, high complexity asset are [47]:

- Remaining life assessment of component
- Inspection/maintenance planning based on true load history
- Relationship between loads and power production for control system policies
- Early damage detection for pre-emptive maintenance and shutdown prevention
- Visualization and inspection of stresses at inaccessible/hidden locations

SIEMENS and GE are the two leading players in digital twins, and both used digital twin concept in many applications including the application of Smart Grids digital twins and other energy technologies. SIEMENS [14] has expanded the concept of the digital twin for better grid planning and operations. With software, grid operators can analyze, plan and supply-demand requirements in a complex electrical system virtually. The grid operator can use virtual instruments to decrease the effort to gather and verify information, enhance investment planning, accurately predict, integrate renewable and conventional power more accurately, and enhance general grid efficiency. GE has created a power plant digital twin [48]. The GE Digital Twin continuously enhances its capacity to model and monitor the state of the plant as the plant is operated. The Digital Twin allows plant operators to optimize the instantaneous and transient control of the plant for efficiency or performance, make informed decisions regarding performance versus part life, assign loads and lineups through time, and perform the right maintenance tasks at the ideal time.

There is usually only one power grid, but distinct grid models of that power grid can be used by a utility. Each grid model coincides with the distinct areas like asset administration, planning, operation and security. Maintaining data synchronized in such an environment is quite a challenge, as each system contains its own specific information, data formats and a distinctive team to assess data that generates digital data silos. During data exchange between those systems, inconsistencies may sometimes occur, and data interchange may not occur in the worst-case scenario. These situations can lead to poor results of the system, imprecisions in models and entire system blackouts. Renewables, distributed energy and digitalization are flooding the utilities with more and more data and also increasing grid connectivity. The issue is therefore only getting worse for utility companies, despite spending cash and time to prevent information losses [14].

Standards and regulations call for greater reliance on data precision and complicated case studies. The increase in information collection and the need for increased information exchange between systems present certain difficulties and also offer utility companies valuable opportunities. To succeed in this evolving atmosphere, companies must alter their strategy towards the management of databases in order to achieve a sustainable digital future. To address these difficulties SIEMENS has designed a DT for electric power utilities that offers a single source of reality in order to model information across the IT landscape. To run grid simulations for all domains this common network model helps tremendously [14].

The SIEMENS Electrical Digital Twin [14] closely aligns real and virtual worlds by providing utilities with a single source of truth to model data across their IT landscape. Data is synchronized from different devices to a multi-user database via standard adapters/interfaces. The common network model enables grid simulation in all fields appropriate in order to plan, operate and maintain reliable, effective and safe electrical systems. SIEMENS ' DT consists of three primary elements: the engine, the user interface and the adapters. The engine consists of:

- Central multi-user database
- Data management functions (e.g. scenarios, variants, projects, etc.)
- Case builder
- Data synchronization
- Data validation
- Data exchange and communication

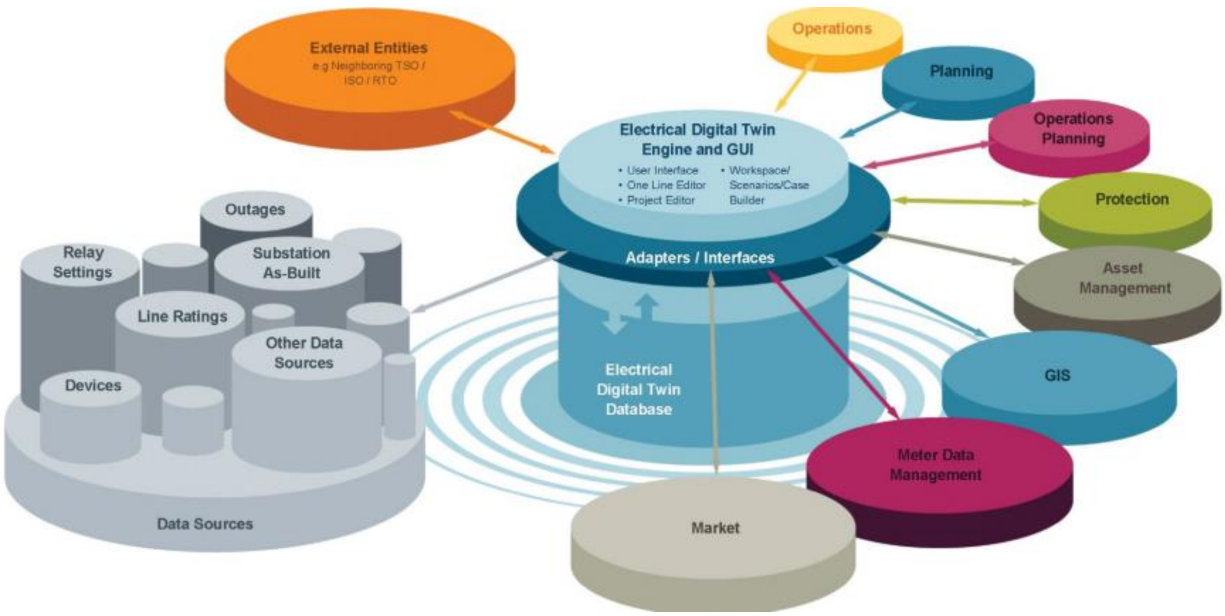

Figure 4 SIEMENS Digital Twin Complete Model Courtesy [14]

2.6.1. Concept of a Digital Twin Based Control Center

With the emergence of Internet of Things and Industry 4.0 sensor data can be coupled with a database. Based on these concepts digital twinning of a power system may revolutionize the future control centers. Dynamic models of the power system are required to make dynamic digital mirror that can reflect states of power system in real-time. Dynamic digital mirror is a simulation engine developed by SIEMENS to mirror power system [49]. Dynamic digital mirror is based on the data available from asset operator. This data contains the current operating states of power system as well as the complete grid model. Dynamic system behavior cannot be studied with the traditional SCADA systems, therefore, WAMS must be employed in Digital Twin centered applications. Dynamic digital mirror becomes a part of power system digital twin. Real-time data utilizing PMU's and RTU's is stored in the digital twin compatible database. This data can be used in a feedback loop to the dynamic digital mirror. The digital image of real power system becomes updated each time real-time data is fed in the database. This real-time data combined with data science algorithms enables dynamic image to reflect the current states and even predict the future states of the power system. Data consistency checks must be implemented to remove any model errors or erroneous measurements to accurately image the real power system [3]. Figure 5 outlines the DT control center Schematic Diagram.

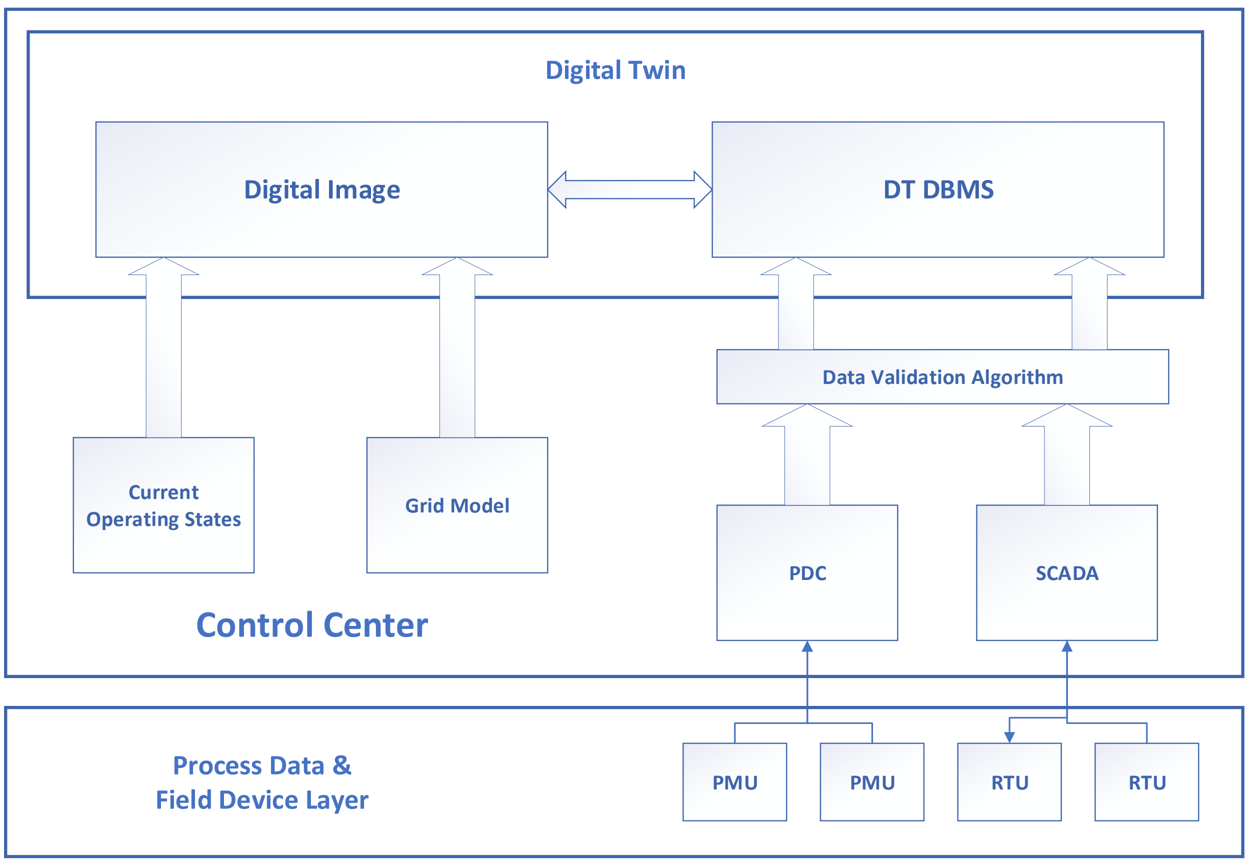

Figure 5 Concept Schematic of a DT Control Center

3. CONCEPTION OF A DATABASE STRUCTURE FOR DT-DBMS

As has been discussed already that a Digital twin is a virtual representation of an object so obviously it represents real object as stored data. Because of the advanced IoT sensors that produce huge amount of data and at quick rates, therefore, this data needs to be stored in a database so that the analysis on data can be performed to predict future states of the system. This huge amount of data helps digital twin to simulate physical object in real-time. The only purpose of data storage is not to perform data analytics but sometimes there is a need to perform tests and simulations as well. The digital twin can be simple or complex depending upon the amount of data it can process and store. For the Digital Twinning of power system, a suitable database structure needs to be selected. There are mainly two types of databases; relational and graph, from which a choice needs to be made.

3.1. Relational and Graph Databases

By implementing a relational database which is a set of formally described tables from which data can be manipulated and accessed without the need of reorganizing the database tables. On the other hand graph databases are NoSQL oriented and can perform different functions such as store, map and query relationships using graph theory [50]. In a graph database, a graph is made of nodes and relations. Each node in a graph denotes an entity, for example, a person, company or object and relations define how these entities are connected to each other.

Relational databases have prearranged columns and rows which allow to store highly structured data in tables having the same type of information. Owing to the complexity of power systems nowadays pure relational databases do not have the ability to store data having large data volume, high velocity and variety as mentioned earlier; relational databases are efficient for storing structured information. Despite this, there are a wide range of databases available in the market which are closely aligned with relational databases concept and introducing some NoSQL features as well. Such databases retain the properties of relational databases and offer flexibility as compared to non-relational databases. Lately, there has been a surge in the use of non-relational databases due to their ability to process large volumes of unstructured and semi-structured data. Due to the flexible, high performance and agile nature of graph databases, there exists a great potential for the use of non-relational databases in the future market [51].

For retrieval and data-intensive storage applications, relational databases have been the first choice for around many decades. A declarative query language, SQL (Structured Query Language) is used to perform retrievals in relational databases. The efficiency of relational database is dependent upon the number of relationships developed by joining larger tables and it decreases considerably when many relationships require joining of large tables. Due to the recent NoSQL movement in database world introduced by renowned projects such as Google’s BigTable and Facebook’s Cassandra the trend toward using NoSQL has increased Other examples of famous NoSQL projects are CouchDB, Dynamo and Project Voldemort all of which reject the relational and object-relational database schemas due to unstructured and very high volumes of data. The NoSQL is preferred for these types of applications because it can handle large volume of data and have the ability to reject the relational and object-relational models which are ACID (atomicity, consistency, isolation and durability) compliant. In order to have a better understanding, a comparison of one such NoSQL graph database called Neo4j with a common relational database system, MySQL, is presented here [50]. MySQL and Neo4j have been chosen for comparison because these are the more commonly used databases from each category i.e. relational and graph.

3.2. Objective measures to compare Graph and Relational Databases

For traversal applications (graph traversal is defined as the process of traversing the graph vertices) Graph databases are much faster than relational databases almost 10 times for certain queries this is because the relational databases are not designed for traversal applications. This is shown in Table 1 below a [50], [52]. Queries of Table 1 are explained first:

S0 is the query to find orphan nodes in a graph. Orphan nodes are defined as nodes that are singletons having no incoming and outgoing edges. S4 query is defined as traversing graph up to depth of 4 and counting the nodes that are reachable. S128 is the query to count the number of reachable nodes by traversing graph to the depth of 128. SO, S4 and S128 queries are best suited for graph databases and they are defined as structural queries. However, S0 for integer-based graph databases does not follow the same trend. The tables below display performance of relational and graph databases for structural and integer-based queries as the database size is increased.

Table 1 Structural Query Results in milliseconds [50]

Database Size	Rel S4 (ms)	Graph S4 (ms)	Rel S128 (ms)	Graph S128 (ms)	Rel S0 (ms)	Graph S0 (ms)
1000int	38.9	2.8	80.4	15.5	1.5	9.6
1000char8K	1.1	0.1	21.4	1.3	1.1	1.1
1000char32k	1.0	0.1	12.5	0.5	1.3	1.0

The query to find orphan nodes results in fairly comparable results between the two systems. The data queries for the integer payload databases demonstrate the efficiency of the relational databases over graph databases.

Now let us discuss queries suited for integer databases. Query I1 is a query to count nodes whose payload data is equal to some value. I2 is counting the number of nodes whose payload data is less than some value [50].

Table 2 Query results for integer-based database in milliseconds [50]

Database Size	Rel I1 (ms)	Graph I1 (ms)	Rel I2 (ms)	Graph I2 (ms)
1000int	0.3	33.0	0.0	40.6
5000int	0.4	24.8	0.4	27.5
10000int	0.8	33.1	0.6	34.8

Graph databases are much faster than relational databases for character-based queries. Here ‘d’ denotes the length of search string. Generally, for small ‘d’ relational databases perform better than graph databases but for higher ‘d’ graph databases outperform relational databases however some anomalies can be observed in the table below .

Table 3 Query results on character databases, in milliseconds [50]

Database Size	Rel (ms)	Graph (ms)	Rel (ms)	Graph (ms)	Rel (ms)	Graph (ms)	Rel (ms)	Graph (ms)	Rel (ms)	Graph (ms)
	d=4		d=5		d=6		d=7		d=8	
1000char8K	26.6	35.3	15.0	41.6	6.4	41.6	11.1	41.6	15.6	36.3
10000char8K	301.6	38.4	269.0	41.5	257.8	41.5	263.1	42.6	249.9	41.5
1000char32K	59.5	41.5	41.6	42.6	30.9	41.5	31.9	41.4	31.9	35.4

3.2.1. Comparison between Relational and Graph Database

To further help in deciding between choosing a Relational and Graph database a table is presented below which summarizes objective and subjective measures to compare MySQL (Relational database) and Neo4j (Graph database) [50], [53].

Table 4 Relational and Graph Database Comparison [50], [53]

	Relational	Graph
Traversal Applications	Slow	Fast
Integer Payload	Fast	Slow
Character Payload	Slow	Fast
Usability and Scalability	Yes	Yes
Maturity/Level of Support	Yes	No
Ease of Programming	Task Dependent	Task Dependent
Flexibility	Highly Flexible	Not so Flexible
Security	Highly Secure	Can be made Secure
Suitability	Structured Data	Unstructured Data

One of the integral characteristics of the graph databases is the presence of multi-tiered operational processes model that enables them to create a visually effective platform. Although, such visually attractive platform is viable but generally it will have a very high computing requirement due to the visually intensive nature of the database. This means that the database will need high processing power in order to have successful operation at larger scale. As compared to the traditional table-based database this is a prominent drawback, because intensive data requirements have a major cost impact on large scale deployments such as those in a power system grid. The other areas which are a major concern for graph database is the level of node-centric search and its potentially slow application while implementing a large-scale graph database. The overall application from a node search point of view is limited and slow, creating issues when implemented in large scale systems such as power grid platforms but relationship indexing within graph databases is effective. The final issue of concern with graph database is the level of modularity which can cause problems in large scale implementation as in the case of power grid system. Developing a fragmented system based upon graph database has two main issues, the first one being difficult to maintain over time and the second one is difficulty in upgrading. These elements present a clear list of issues that may cause problems when utilizing a graph database in a power grid system [54].

There is no need for predefined database schema in graph database as it stores data as schema-free. So, even when there is a change it can store data instances in a database schema based on CIM profile or any other schema. However, for the database schema based on CIM, as a tradeoff, the CIM semantics have to be linked to the artifacts of Graph database for every data instance. But this process of adding CIM semantics is viably time-consuming, tedious and error-prone [55].

Graph databases are suitable for highly unstructured data where we have to process images and shapes a lot of the time. For highly unstructured data, there might be a temptation to use NoSQL databases but for this thesis work data may be very large, but it will be structured, time-stamped and clearly defined. There are only a few data sources (such as relay fault records) that are unstructured. For this relatively small amount of unstructured data, PostgreSQL has added capability like JSON which can handle unstructured data [13]. Suitable database for this work will be integer based and it will be querying integer-based structured data most of the time. As already discussed in the objective and subjective comparison of graph and relational databases, the relational databases are better for integer-based queries. Finally, it can be concluded that Graph databases can indeed be used for DBMS of a power system application but considering the points discussed above, it can be said that using relational database such as PostgreSQL for this database management system is more suitable. A comparison of different relational databases and suitability of PostgreSQL database for this thesis is discussed in the coming sections.

3.3. Relational Database Management Systems

In order to work with the data, relational models are implemented in Relational Database Systems. The relational model configures structured information to be stored by making appropriate schemas, which are related entities with attributes across tables. In order to contain and work with the data, such types of database management systems require tabular structures. In these tables, each column also known as attribute in database terminology, hold different information having different data types. A row of such a table is known as a record in database terminology and each row is associated with a unique key. Relational model defines these attributes and records thus they are related to each other. In the column of these tables, different type of data is contained. Whereas, a unique key is identified for each record in the database, which is translated to a specific row that links to a table, with each row's series of attributes being characterized as the columns of a table -- all linked together, as defined within the relational model [56]. A Relational database has a unique characteristic, i.e., it contains more than one interconnected table. Thus, a relational database management system has the ability to interconnect data tables on the basis of logic and can also perform search inside these tables for common information.

3.3.1. Comparison of Relational Databases

There are different Relational database management systems based on different relational databases for example SQLite, MySQL and PostgreSQL etc. So, in order to make choice easier a comparison between different relational databases is presented below [56]:

Table 5 Comparison of Relational Databases [56]

Database	**SQLite**	**MySQL**	**PostgreSQL**
Supported Data Types	null, integer, real, text, blob	**Numeric:** int, tinyint, smallint, mediumint, bigint, float, real, decimal, boolean, bit **Date and Time:** date, datetime, timestamp, time, year **String:** char, varchar, binary, blob, longblob, text, longtext, enum, set	**Numeric:** smallint, bigint, serial, smallserial bigserial, doubleprecision, real, decimal, integer **Date and Time:** date, interval, time, time with time zone, timestamp, timestamp with time zone **Character:** char, varchar, text **Geometric:** box, circle, line, lseg, path, point, polygon **Network Address:** cidr, inet, macaddr **Bit String:** bit, bitvarying **Text Search:** Tsquery, tsvector **JSON:** json, jsonb **Other Data:**

			boolean, bytea, money, uuid, xml, pg_lsn, txid_snapshot
Type of Applications	Embedded, Testing	Distributed operations and web applications.	Data Integrity and complex database designs.
Advantage	Portability	Easy to work with and feature rich.	Superior parallel processing, Standards compliant, has NoSQL features and Extensible.
Disadvantage	No User Management and lack of Possibility to tinker performance. Not suitable for large datasets.	Functional limitations and Reliability issues along with problematic concurrent read operations.	Less Popular than MySQL and may underperform for simple applications.

For this thesis work, a Python based database configuration script (i.e. PostgreSQL) for setting up a configurable database scheme will be used, which could be extendable, that possesses advantages of improved security, enhanced scalability, improved performance, robustness and ease of use [57].

3.3.2. PostgreSQL Database Management System

The main goal of PostgreSQL is to be standards-compliant and extensible DBMS. In addition, it is also one of the more advanced and open-source [object]-relational database management systems. It strives to adopt the ANSI/ISO SQL standards by making the revisions. The PostgreSQL can be described as a general-purpose object-relational database management system. So, it is allowed to add customized functions using programming languages such as Java, C/C++, etc. [56]. Compared to other RDBMS, PostgreSQL provides support for object-oriented and ACID (Atomicity, Consistency, Isolation, Durability) compliant relational database functionality.

Owing to the powerful inherent technology, PostgreSQL is proficient in handling huge number of tasks very efficiently. Thanks to the application of MultiVersion Concurrency Control (MVCC), support for Concurrency is achieved without using read locks, which also ensures that ACID compliance is enacted. Even before Oracle was using the MVCC feature, PostgreSQL already became the first database management system that implemented it. The customized procedures known as "stored procedures" make PostgreSQL highly programmable. Such customized features are created for repeated, complex and database operations that require simplification [58]. PostgreSQL is designed to be highly extensible. The PostgreSQL offers user the ability to define his own index types, data types and functional languages. If the user does not like one part of the system, he has the option to develop his own customized plugin in order to fulfill his specifications. The user can even override different parts of the system using the customized plugins to add a new optimizer or to change the execution of commands. PostgreSQL is supported by automated testing for both features and concurrency and it also offers high-quality. When used in default setting, the database offers strong disk-write guarantees. So, it ensures the developers protection of data because for developers' risk of data-loss is serious consideration while selecting database management system. By using CREATE EXTENSION, database extensions are easily loaded, which not only automates versions checks but also dependencies and other aspects of configuration. PostgreSQL supports user-defined data types, operators, indexes, functions and languages. The version 9.4 of PostgreSQL scales well on a single node up to 32 CPUs. PostgreSQL has the capacity to scale well up to hundreds of active sessions, and can also handle up to thousands of connected sessions when using a session pool [57].

PostgreSQL can be chosen over other relational databases [56].

- When reliability and data integrity are an absolute necessity without excuses, PostgreSQL is the better choice.
- If there is a database requirement to perform custom procedures, PostgreSQL, being extensible, is the better choice.
- In the future, if there is a chance of necessity arising for migrating the entire database system to a propriety (e.g. Oracle) solution, PostgreSQL will be the most compliant and easy to handle base for the switch.
- Compared to other open-source and free RDBMS implementations, for complex database designs, PostgreSQL offers the most in terms of functionality and possibilities without giving up on other valuable assets.

3.4. Power System Simulation Data

At present, electrical networks can be simulated in a great amount of depth and detail utilizing, the real-time simulators and advanced programs. The details of the power system simulation are improving constantly with time by the implementation of new technologies, new models and the addition of new application areas. The real-time simulation of power systems is also constantly moving from analog and hybrid technology to pure digital technology [59].

Dynamic modeling and parameter identification of loads, generators, prime mover, excitation systems and governor systems are a foundation for operation, planning, design and analysis in power systems. For the calculation of practical power system computation results, precision in the power system component models and parameters play an important role. Establishment of accurate power system models and identification of precise parameter values is important because parameter values supplied by vendors may not be accurate enough to make power system computations [60].

A classification of power system data can be made, considering type of Data that is stored in a Database.

Real-time Data:

Real-time data depicts the status of power system components and is obtained directly from PMUs or RTUs with rate of arrival different for both sensing devices. Real-time data is periodically stored in the database. To get real-time data monitoring of power system needs to be done in real-time employing the SCADA and WAMS. Along with the measurements of current and voltage the status of circuits breakers is also taken. Samples of these measurements are taken continuously. For example, in the monitoring of circuit breakers, RTUs of a SCADA system are used.

It is not possible to use real-time data in the implementation of DBMS of this work due to non-availability of such data. For the DBMS test setup of this work OPC Server combined with PowerFactory and Matlab will be used to mimic the field measurements.

Topological Data:

This type of data contains the power system topology and the data which is not dynamic such as the data of some transmission line parameters and transformer impedance etc.

Process Data:

This type of data is calculated from real-time data and static data. It is based on different analysis i.e. load flow and short circuit etc. This type of data includes grid model, branch current obtained from state estimation, magnitude and angle of bus voltage.

3.4.1. Dynamic Power System Modeling

Dynamic power system modelling is not an easy task because we need to have detailed models of all the power system equipment. Inclusion of detailed power system models add to the complexity of the power system dynamic simulation. Therefore, some simplifications have to be introduced. If too much simplifications in the models are introduced it reduces the efficiency of digital twin [3]. So, there is a fine line between the models reduction and achieving the full capability of digital twin. Power system components/devices and their models are described in the upcoming sections of the thesis. The models have been organized into different classes which will be later implemented in the database.

Input data for dynamic model consists of three main parts [61]:

- static topological data (available in SCADA and corresponding to data for load flow and short circuits calculations)
- additional data for protections, secondary U/Q control and on load tap changers models (available in Energy management System - EMS)
- data for dynamic models (they include type of model and parameters for it).

Power system dynamic model should contain all components necessary for realistic power system simulation, namely these particular models for:

1. synchronous and asynchronous generator and excitation system for synchronous generators,
2. prime movers (steam, hydro, gas and wind turbines, internal combustion engine and so on),
3. energy sources (steam boilers, hydro reservoirs and so on),
4. energy conversion systems like photovoltaic power plants (called Power Park Modules as well),
5. load (lighting and heating, induction motors, static characteristic and so on),
6. secondary U/Q and f/P (called load frequency control as well) controls
7. on load tap changers for normal transformers and for phase shifting transformers,
8. network protection (distance, overcurrent and so on) and special protection schemes,
9. FACTS and HVDC.

3.5. Power System Classes and Parameters for DT DBMS

Below is the list of different types of data classes needed for power system representation. Data has been divided into different CIM compliant categories/packages/classes so it would be more feasible to make database management system considering these categories/packages/classes of data.

In the following sections, power system models and parameters are mostly taken from [62] which is not a good research practice but this has been done only to remain consistent with the naming conventions and general consistency. Where deemed feasible models have been taken from other sources such as IEC standards and other books etc. too.

3.5.1. Core Power System Class

It contains core data about the power system resources and conducting equipment. Conducting equipment includes the parts of the power system that are designed to carry current or that are conductively connected. Core power system data is further divided into following parts:

- Bay Data
  It contains data of a collection of power system resources (within a given substation) including conducting equipment, protection relays, measurements, and telemetry.
- Conducting Equipment Data
- Power System Resource Data
- Base Power
- Base Voltage
- Terminal Data
  An electrical connection point to a piece of conducting equipment. Terminals are connected at physical connection points called "connectivity nodes".
- Geographical Data

3.5.2. Topology Class

It contains data about how the equipment is connected together. It is further divided into

Bus Name Data: Used to apply user standard names to topology buses.

Connectivity Nodes Data: Connectivity nodes are points where terminals of conducting equipment are connected together with zero impedance.

Topological Node Data: A set of connectivity nodes that, in the current network state, are connected together through any type of closed switches, including jumpers. Topological nodes can change as the current network state changes (i.e., switches, breakers, etc. change state).

3.5.3. Power System Resource Class

A power system resource can be circuit, breaker, transmission line, transformer and regulators etc. Power System Resource Class contains data about the electrical characteristics of Generation, Transmission and Distribution Systems. It is further divided into:

3.5.3.1. AC Line Segment/Transmission Line

A wire or combination of wires, with consistent electrical characteristics, building a single electrical system, used to carry alternating current between points in the power system. This data consists of Susceptance, Conductance, Length, Resistance and Reactance data. It contains data about the conductor and type of conductor.

A short transmission line can be modelled by lumped pi model. Short transmission line means line length is much shorter than the wavelength.

$$l \ll \lambda$$

where l is the line length and λ is the wavelength defined as

$$\lambda = \frac{1}{f_n\sqrt{L_L C_L}}$$

where f_n is the rated frequency of the AC system, and L_L and C_L are the per unit length inductance and capacity respectively, of the transmission line.

Figure 6 gives the lumped pi model of short transmission line [62].

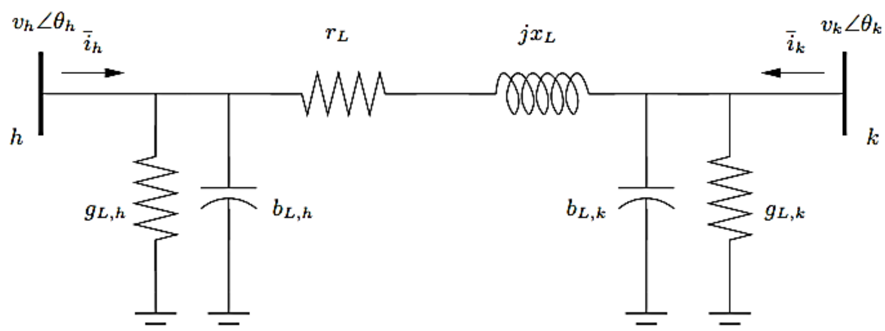

Figure 6 Transmission Line Lumped Pi Model [62]

Typical parameters of a short transmission line are defined in table below.

Table 6 Transmission Line Parameters [62]

Parameter	Description	Unit
$b_{L,h}$, $b_{L,k}$	Shunt susceptances	pu
C_L	Per-unit length line capacity	F/km
$g_{L,h}$, $g_{L,k}$	Shunt conductances	pu
I^{max}	Current limit	kA
L_L	Per-unit length line inductance	H/km
l_t	Total line length	km
P^{max}	Active power limit	MW
R_L	Per unit length line resistance	Ohm/km
r_L	Resistance	pu
S^{max}	Apparent power limit	MVA
x_L	Reactance	pu

Below table represents an estimate of some parameter values and limits which may be helpful in cross check validation algorithm.

Table 7 Transmission Line Parameter Limits [62]

Parameter	Value and Unit
Rated frequency f_n	50 or 60 Hz
Line Length and Capacitance	Depends upon line Geometry. Typically, Line Length (l)=200km-750 km at 50 Hz
Wavelength (λ) (Overhead)	6000 km
Wavelength (λ) (Cables)	2000-2800 km.

When a lossless transmission line is loaded at surge impedance, the reactive power loss of the line is zero and the natural impedance loading is an optimum condition with respect to voltage and reactive power control. The shunt capacitance of the cable depends strongly on whether the three-phase conductors are screened or constitute separate single-phase cables. Typically, the per-unit-length series reactance of a cable is about half that of a similarly rated overhead line.

Line capacitance can be calculated from the below formula

$$\lambda = \frac{1}{f_n\sqrt{L_L C_L}}$$

The lumped series resistance and reactance can be computed as

$$r_L = \frac{R_L l_t}{Z_b}$$

$$x_L = \frac{\omega_s L_L l_t}{Z_b}$$

The per unit length resistance R depends on the temperature, on the section and on resistivity of the conductor. The lumped shunt susceptance can be approximated as:

$$b_{L,h} = b_{L,k} = \frac{\omega_s C_L l_t Z_b}{2}$$

Table below shows an example of the overhead transmission line parameter values.

Table 8 Overhead Transmission Line Parameters [63]

f_n (Hz)	V_n (kV)	R_L (Ω/km)	x_l=ωL (Ω/km)	$b_{L,h}$=ωC (μS/km)	Z_b (Ω)	P_{SIL} MW	Type of Line
50	275	0.067	0.304	4.14	271	279	Short
50	400	0.018	0.265	5.36	222	720	Short
50	230	0.05	0.488	3.371	380	140	Long

PSIL defined here is the surge impedance loading power. Which is given by

$$P_{SIL} = \frac{V_n^2}{Z_b}$$

Table 8 compares overhead transmission line parameters for long and short transmission lines. Transmission Line parameters are dependent on the conductor dimensions and relative placement. The surge impedance of most overhead lines is around 250-350 ohms whereas it is 30-50 ohms for cables. Typical positive sequence inductance and capacitance parameters for a 400-kV overhead line:

L = 1.044mH/km, C = 12 nF/km. The resistance per unit length of this line = 0.0296 ohm / km

Typical values of the transmission line parameters depend upon many factors such as type of conductors, operating, temperature, spiraling and bundling of conductors. Dimensions and characteristics of three-phase high-voltage overhead lines are reported in Appendix II [64].

3.5.3.2. DC Line Segment

DC line segment consists of a wire or combination of wires that are used to carry direct current between points in the DC region of the power system. Data mainly consists of Inductance, Susceptance, Conductance, Length, Resistance and Reactance.

3.5.3.3. Protection Devices

Protection Devices class models information for protection equipment such as relays and circuit breakers. Contains attributes such as normal open and fault interruption current rating, rated current and transient time etc. Protection equipment are associated with conducting equipment and usually operate circuit breakers. Protection devices class has many sub classes i.e. current relay, relay synchronous check and circuit breaker reclosure sequence class.

3.5.3.4. Transformer

The transformer class consists of data about the transformer coolant type, transformer winding type, winding connection and winding test data. It also contains tap changer data for changing transformer winding tap positions.

Two-winding transformers can be modelled as a transmission line with a series impedance $z_T = r_T + jx_T$ and a shunt admittance at the sending-end bus, which models iron losses g_{Fe} and the magnetizing susceptance bμ. From the modelling viewpoint, the main difference between transformers and transmission lines is that transformers can introduce a complex off-nominal tap ratio me^jφ that allows modifying the magnitude and the phase angle of the receiving or sending-end bus voltage. Figure 7 shows the single phase two winding transformer model with off-nominal ratio m.

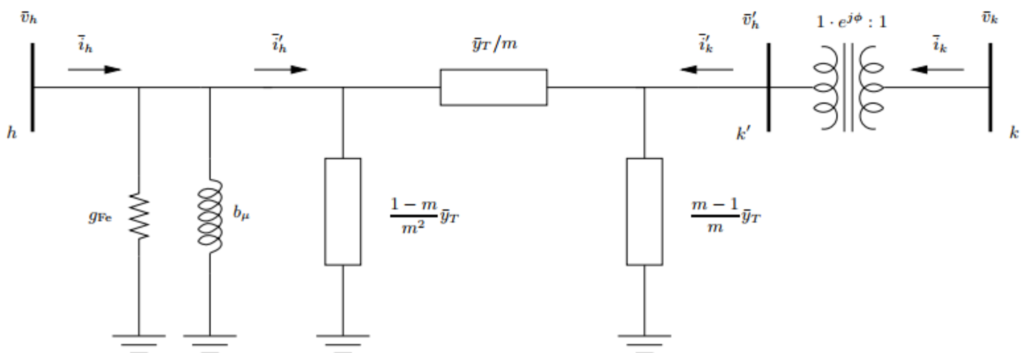

Figure 7 Single Phase Two Winding Transformer Model [62]

The value of parameters in the transformer equivalent circuit can be determined from the no load test and the short-circuit test. In both of these tests the supply voltage, current and real power are measured. The fixed tap ratio m unit is pu/pu since it represents the ratio of the primary voltage in pu by the secondary voltage in pu. For example, if the nominal voltages are $V_{n,h}$ = 220 kV and $V_{n,k}$ = 128 kV, and the actual tap positions of the transformer are V_h= 231 kV and V_k = 128 kV, the tap ratio m is given by

$$m = \frac{V_h \, V_{n,k}}{V_{n,h} V_k}$$

m=1.05

Table 9 describes different parameters required to model 2 winding Transformers.

Table 9 Transformer Parameters [62]

Parameter	Description	Unit
b_μ	Magnetizing susceptance	pu
$k_T = V_{n,h}/V_{n,k}$	Nominal voltage ratio	kV/kV
g_{Fe}	Iron losses	pu
I^{max}	Current limit	kA
M	Fixed tap ratio	pu/pu
r_T	Resistance	pu
P^{max}	Active power limit	MW
$V_{n,h}$	Primary voltage rating	kV
$V_{n,k}$	Secondary voltage rating	kV
S^{max}	Apparent power limit	MVA
x_T	Reactance	pu
Φ	Fixed phase shift	rad

To find parameters of a transformer, standard transformer tests are performed i.e. no-load test, shirt circuit test. Table 10 summarizes the parameters of a two-winding transformer as found in the [65]. According to tests performed in [65], for nominal, minimum and maximum tap position. Values close to actual resistance, reactance and impedance have been obtained by interpolation in excel [65].

Table 10 Transformer Parameters

Tap m	V_n	P_{LL}	I_{max}	Test	I_{Hrated}	Zbase	z (pu)	r_T	Z	R	X
	kV	kW	%	MVA	A	@ MVA			Ω	Ω	Ω
1	151.8	39.952	7.5	15	57.05	1536.216	0.075	0.002663	115.216	4.0917	115.143
9	138	41.66	7.68	15	62.76	1269.6	0.0768	0.002777	97.5052	3.5261	97.4415
17	124.2	37.25	7.41	13.5	62.76	1142.64	0.0741	.002759	84.6696	3.1528	84.6109

3.5.3.5. Busbar Section

Busbar Section is defined as a conductor, or a group of conductors, with negligible impedance, that serve to connect other conducting equipment within a single substation.

3.5.3.6. Generation Data Class

Generation data is further sub divided into two portions.

1. Production
   It describes various kinds of generators that are present within the power system. Contains information about the generating unit, generator operating mode, cost, operating schedule
2. Generation Dynamics
   It contains data about the prime movers, such as turbines and boilers.

- **Synchronous Machine**

An electromechanical device that operates synchronously with the network. It is a single machine operating either as a generator or synchronous condenser or pump. Data consists of Time delay, damping, inertia, coolant type, base reactive power, max reactive power, min and max voltage, operating mode, active power, positive, negative and zero sequence resistance and reactance, direct axis sub transient and synchronous reactance, apparent power, quadrature axis sub transient and synchronous reactance.

The electrical equations for all variants of the synchronous machine models are based on the Synchronous machine equivalent diagram for the direct and quadrature axes. Figure 8 gives one such synchronous machine model for direct and quadrature axes [66]

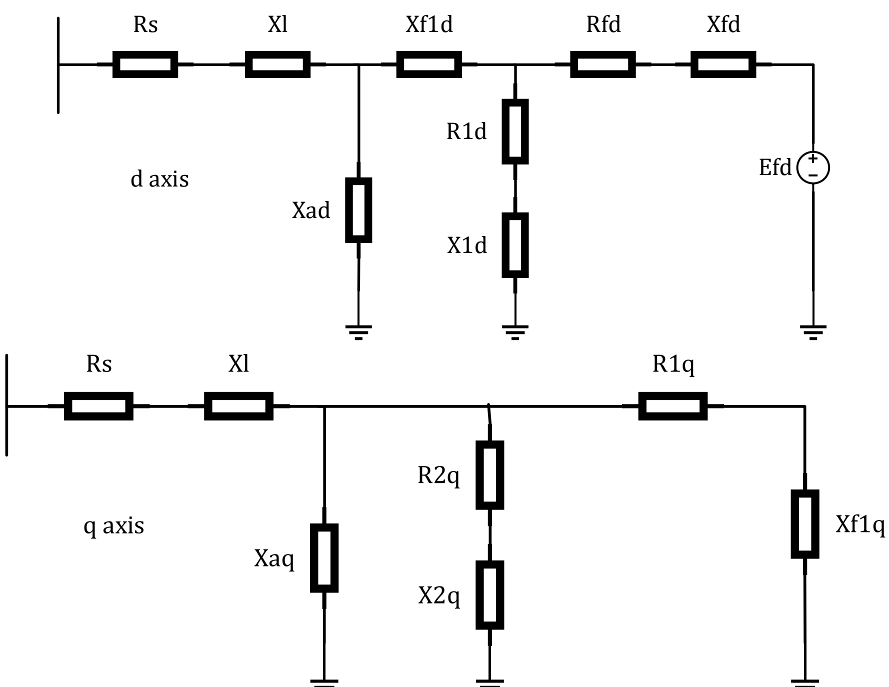

Figure 8 Synchronous Machine Parameters [66]

Relationships are represented with the equations below:

$$X_d = X_{ad} + X_l$$

$$X_q = X_{aq} + X_l$$

Table 11 gives description of synchronous machine parameters and their typical values (IEC 61970-302).

Table 11 Synchronous Machine Parameters Typical Values [66]

Parameter	Description	Limit	Typical Value (pu)
D	Damping coefficient	≥0	0
H	Inertia constant	>0	3
S_1	Saturation factor	≥0	0.02
S_{12}	Saturation factor 120	≥Saturation factor	0.12
x_l	Stator leakage reactance	≥0	.15
r_s	Stator resistance	≥0	0.005
K_s	Saturation correction Factor	≥0	0
x_d	d-axis synchronous reactance	$\geq x'd$	1.8
$x'd$	d-axis transient reactance	$\geq x''d$	0.5
$x''d$	d-axis sub transient reactance	$> x_l$	0.2
x_q	q-axis synchronous reactance	$\geq x'q$	1.6
$x'q$	q-axis transient reactance	$\geq x''q$	0.3
$x''q$	q-axis sub transient reactance	$> x_l$	0.2
$T'do$	d-axis open circuit transient time constant	$> T''do$	5s
$T''do$	d-axis open circuit sub transient time constant	> 0	0.03s
$T'qo$	q-axis open circuit transient time constant	$> T''do$	0.5s
$T''qo$	q-axis open circuit sub transient time constant	> 0	0.03s
t_c	Damping time constant for leakage resistance	≥0	0s

- **Synchronous Machine Regulators**

This chapter describes the most relevant synchronous machines primary regulators and limiters. These are the turbine governor, the automatic voltage regulator and the over- and under-excitation limiters.

i. **Turbine Governor**

Turbine Governors (TGs) define the primary frequency control of synchronous machines. When defining the TG data, the droop R and mechanical power limits are often given in pu with respect to the synchronous machine active power rating. If this is the case, during the initialization of turbine governor's data, the droops (The droop R is a measure of the participation of each machine to system losses and load power variations) have to be converted to the system power base, as follows [62] :

$$R_{system} = R_{machine} \frac{S_{ystem}}{S_{machine}}$$

When initializing the turbine governor variable, mechanical power limits have to be checked. If a limit is violated, it means that the turbine governor parameters are not consistent with those of the static generator used in power flow analysis. Turbine governors regulates the production of synchronous machine active powers. There are two common types of Turbine Governors [62]:

TG Type 1:

It includes a governor, a servo and a reheat block. Figure 8 represents the type 1 turbine governor model control diagram.

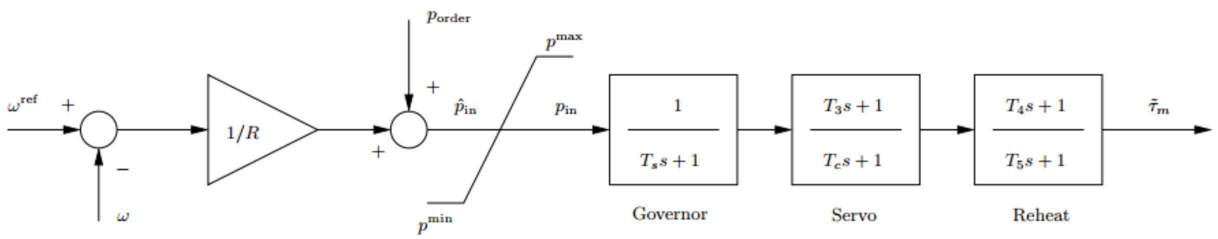

Figure 9 Type 1 Governor Model [62]

The number of blocks of each part of the turbine can be increased to take into account each stage in detail. However, the structure of the control diagram does not change. Table 12 gives the type 1 turbine governor parameters and their brief description.

Table 12 TG 1 Parameters [62]

Parameter	Description	Unit
p^{max}	Maximum turbine input	pu
p^{min}	Minimum turbine input	pu
R	Droop	pu
T_3	Transient gain time constant	s
T_4	Power fraction time constant	s
T_5	Reheat time constant	s
T_c	Servo time constant	s
T_s	Governor time constant	s

TG Type 2:

The turbine governor type II is typically more than adequate for transient stability analysis. TG2 control diagram is drawn below and parameters are listed in a table.

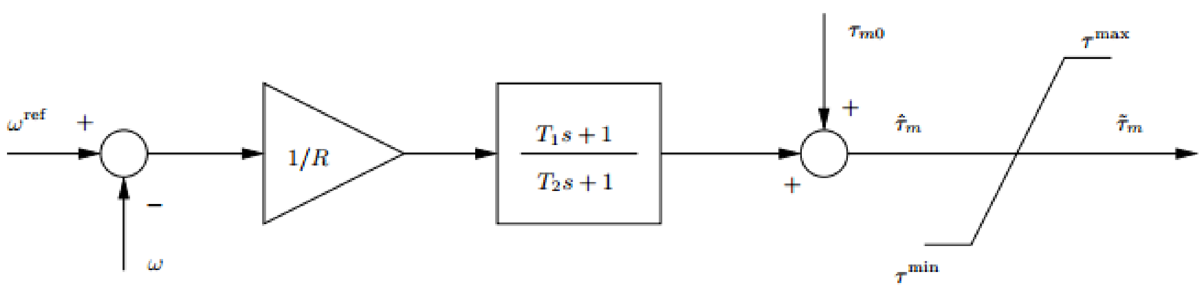

Figure 10 TG 2 Control Diagram [62]

Table 13 gives the type 2 governor parameters and their brief description.

Table 13 TG 2 Parameters [62]

Variable	Description	Unit
τ^{max}	Maximum turbine output	pu
τ^{min}	Minimum turbine input	pu
R	Droop	pu
T_1	Transient gain time constant	s
T_2	Governor time constant	s

ii. **Automatic Voltage Regulator**

Automatic Voltage Regulators (AVRs) define the primary voltage regulation of synchronous machines. Several models have been proposed for AVR's in practice. AVRS are broadly classified into three categories. Detailed description of different models of AVRS is beyond the scope of this thesis. Brief summary of regulator models is described below in a table. Table 14 shows the brief comparison between the three types of automatic voltage regulators.

Table 14 Automatic Voltage Regulators [62]

Model	**Response Time**	**Application**
AVR Type 1	Slow	DC excitation
AVR Type 2	Fast	Static excitation
AVR Type 3	Fast	Simplified stability analysis

A brief overview of type III AVR is given in this section. Type III AVR is a simple model that is useful for simplified stability studies. Figure 11 outlines the control diagram for type III AVR. Here v_h is the generator terminal voltage or a remote bus regulated voltage. The initial field voltage v_{fo} and bus voltage v_o are set during the synchronous machine initialization step. Field voltage is subjected to an anti-windup limiter [52].

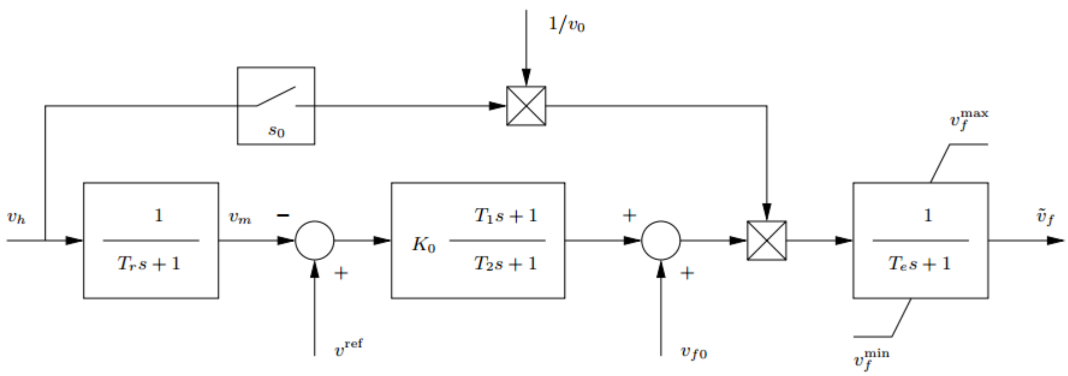

Figure 11 Type 3 AVR Control Diagram [62]

Detailed parameters of a type 3 AVR are listed below in a Table.

Table 15 Type 3 AVR Parameters [62]

Variable	**Description**	**Unit**
v_f^{max}	Maximum field voltage	pu
v_f^{min}	Minimum field voltage	pu
K_0	Regulator gain	pu/pu
T_1	Regulator zero	s
T_2	Regulator pole	s
T_e	Field circuit time constant	s
T_r	Measurement time constant	s
s_0	Bus voltage signal	binary

iii. Power System Stabilizer

Power System Stabilizers (PSSs) are used for damping power system oscillations. Although several PSS models have been proposed in the literature, the rationale behind the PSS functioning is the same for all control schemes. Typical PSS input signals are the rotor speed ω, the active power and also the bus voltage of the generator to which the PSS is connected through the automatic voltage regulator. The PSS output signal is a signal v_s that modifies the reference voltage of the AVR [62].

Power Stabilizers are generally categorized into:

- PSS Type 1
- PSS Type 2
- PSS Type 3

Block control diagram of type 3 PSS is drawn below.

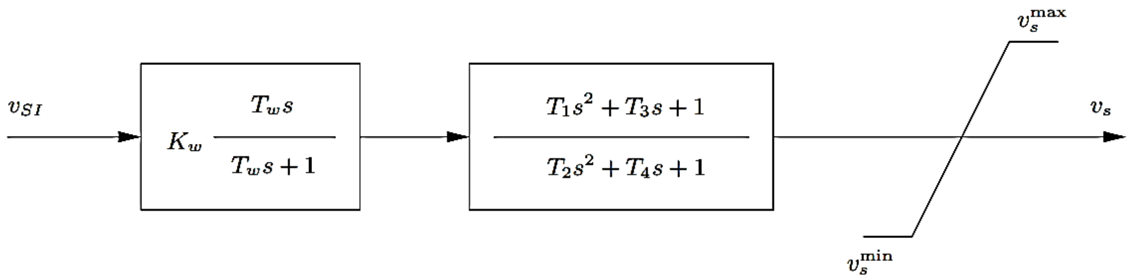

Figure 12 Type III PSS Control Diagram [62]

The typical type III PSS parameters are described in the Table 16 below.

Table 16 Typical PSS Parameters [62]

Variable	Description	Unit
v_s^{max}	Maximum stabilizer output signal	pu
v_s^{min}	Minimum stabilizer output signal	pu
T_w	Wash-out time constant	s
T_1	First stabilizer time constant	s
T_2	Second stabilizer time constant	s
T_3	Third stabilizer time constant	s
T_4	Fourth stabilizer time constant	s
K_p	Gain for active power	pu/pu
K_v	Gain for bus voltage magnitude	pu/pu
K_w	Stabilizer gain	pu/pu

3.5.3.7. Regulation Control Class

Regulation control class specifies a set of equipment that works together to control a power system quantity such as voltage or flow. It contains data about the type of control, input range, output range and target value.

3.5.3.8. HVDC Class

High-Voltage Direct-Current (HVDC) transmission systems consist of a dc line interfaced to the ac network through a rectifier and an inverter. HVDC systems are mainly used for long connections that cannot be obtained with standard ac transmission lines due to distributed parameter issues, e.g., Ferranti's effect and delays. HVDC systems are also the unique choice in case of long submarine connections. HVDC data consists of inverter, rectifier and control parameters.

HVDC controllers have to coordinate the operations of the rectifier and the inverter devices. The controlling variables are the transformer tap ratio and the firing angle on the rectifier side and the transformer tap ratio and the extinction angle on the inverter side. Tap ratio controls are necessarily slower than those of firing/extinction angles. The firing and extinction angles, which are characterized by fast dynamics, are considered algebraic variables, while the tap ratio, whose dynamic is relatively slow, are considered state variables. Figure 13 represents the HVDC control scheme.

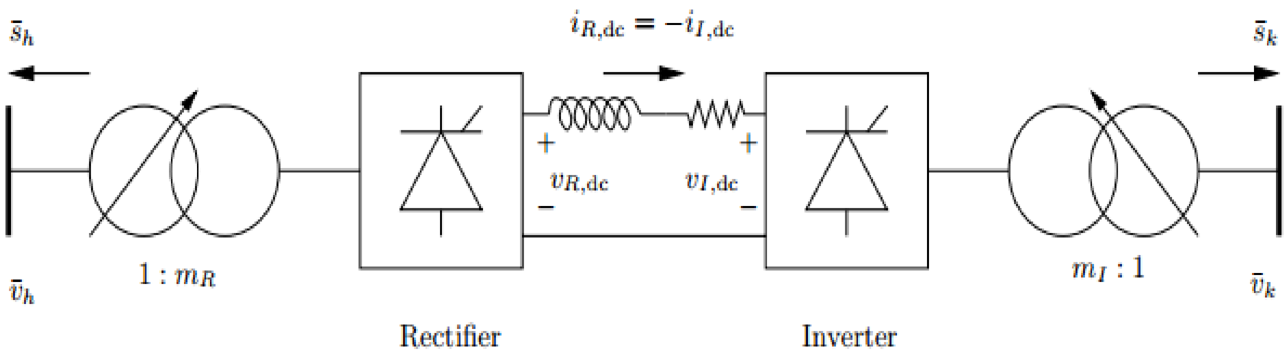

Figure 13 HVDC Scheme [62]

HVDC transmission system can be controlled via 3 schemes:

- Rectifier Current Control Mode
- Inverter Current Control Mode
- Power Control

HVDC parameters are discussed in the table below.

Table 17 HVDC Parameters [62]

Parameter	Description	Units
c	Slope of the voltage/current power control	pu
i_{dc}^{m}	Current margin	pu
i_{dc}^{min}	Minimum dc current	pu
K_i	Integral gain for the current PI control	1/s
K_p	Proportional gain for the current PI control	rad/pu
p^{ref}	Power reference	pu
T_I	Inverter control time constant	s
T_R	Rectifier control time constant	s
v_{dc}^{max}	Maximum dc voltage	pu
v_{dc}^{min}	Minimum dc voltage	pu
v_{dc}^{ref}	Reference dc voltage	pu
v_{ac}^{ref}	Reference ac voltage	pu
α^{ref}	Reference firing angle	rad
γ^{ref}	Reference extinction angle	rad

i. **Voltage Source Converters**

The Voltage Source Converter (VSC) device is similar to the rectifier or inverter devices of the HVDC transmission system. Thus, the VSC provides a link between ac and dc networks. The main difference with HVDC systems is the technology of power electronic switches. HVDC devices are built using conventional thyristors, which provide the turn-on control, but can be turned off only if the current is zero. The electronic switches used for VSC devices are Gate Turn-Off Thyristor (GTO), MOS Turn-off Thyristor (MTO), Integrated Gate Bipolar Transistor (IGBT) or Integrated Gate-Commutated Thyristor (IGCT). These are called turn-off devices since they provide both turn-on and turn-off controls, the latter also for non-zero currents. In the recent past, turn-off devices were more expensive and had higher losses than conventional thyristors. These drawbacks implied that the typical VSC capacity was smaller than HVDC ones. Therefore in practice, VSCs were mainly used for small to medium power applications such as FACTS devices, wind turbines and other distributed energy resources [62]. This is not strictly true anymore and VSCs have found vast applications today including applications in HVDC.

The typical VSC configuration includes a capacitor, a bi-directional inverter and a transformer. The capacitor allows maintaining constant dc voltage while the transformer provides Galvanic insulation. The ac/dc conversion is provided by the VSC, which also includes a series phase reactor and a shunt ac filter. Figure 14 shows the control scheme for voltage source converters [62].

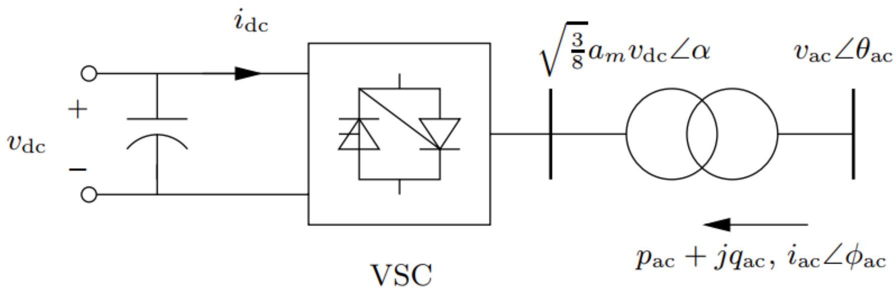

Figure 14 Voltage Source Converter Control Scheme [62]

Table 18 gives the introduction and brief description of typical VSC parameters.

Table 18 VSC Parameters [62]

Variable	Description	Unit
i_{ac}^{max}	Maximum ac current	pu
i_{dc}^{max}	Maximum dc current	pu
a_m^{max}	Maximum modulating amplitude	pu
a_m^{min}	Minimum modulating amplitude	pu
r_T	Transformer resistance	pu
x_T	Transformer reactance	pu
α^{max}	Maximum firing angle	rad
α^{min}	Minimum firing angle	rad

The VSC model is completed by the controls that regulates the modulating amplitude and the firing angle α. These controls depend on the application and should be defined separately to maintain the VSC model as general as possible. A special care has to be devoted to the operating limits of the VSC device. Both the firing angle and the modulating amplitude are limited:

$$\alpha^{max} \leq \alpha \leq \alpha^{min}$$

$$a^{max} \leq a_m \leq a^{min}$$

These limits are used for bounding the controls of the VSC device. Then the ac and dc current limits indicated in Table 17 have also to be checked during VSC operation. If a current limit is reached, then some of the VSC controls have to be locked. In other words, current limits impose indirect limits on the firing angle α and the modulating amplitude [62].

3.5.3.9. FACTS Class

Transmission networks equipped with high-speed electronic control devices are referred to as Flexible AC Transmission Systems while the electronic devices themselves are referred to as FACTS devices. At the heart of FACTS devices is a controlled semiconductor, the thyristor. Facts devices include static var compensators, thyristor controlled static compensators, static series compensators, static synchronous compensators and unified power flow controllers. FACTS devices are categorized into the following categories:

- The Static Var Compensator (SVC)
- The Thyristor Controlled Series Compensator (TCSC)
- The Static Synchronous Compensator (STATCOM)
- The Static Synchronous Series Compensator (SSSC)
- The Unified Power Flow Controller (UPFC)

i. **The Static Var Compensator (SVC)**

The Static Var Compensator (SVC) is a variable shunt capacitor that is varied to maintain a constant voltage at the bus to which it is connected. Static Var Compensators are categorized into

- Type 1 SVC Compensator
- Type 2 SVC Compensator

The admittance is varied using a thyristor-based switch. The firing angle α controls the turn-on period of the thyristor and hence varies the equivalent reactance of the SVC. SVC type II is approximate version of the type 1 SVC. Figure 15 shows the type 1 SVC scheme.

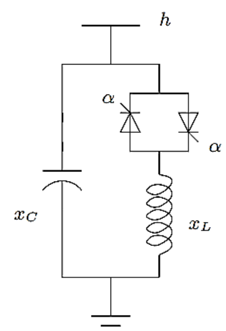

Figure 15 Type 1 SVC Scheme [62]

As discussed above, the controlled variable is the firing angle α. Thus, the regulator has to vary α in order to control the bus voltage. Figure 15 represents the SVC Type 1 control diagram.

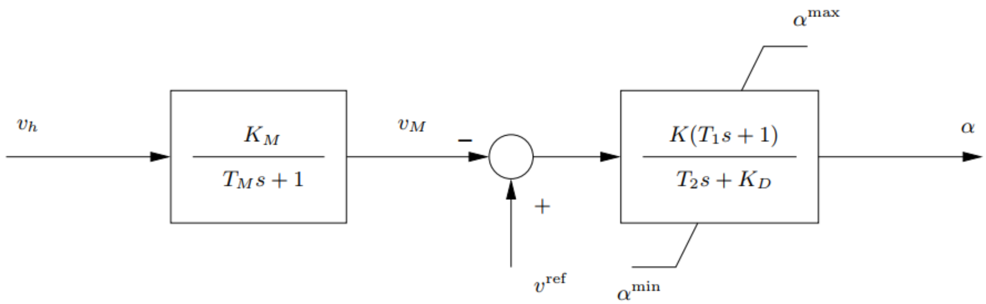

Figure 16 SVC Type 1 Control Block Diagram [62]

SVC Type 1 parameters are described in the table below.

Table 19 SVC Type 1 Parameters [62]

Variable	Description	Unit
K	Regulator gain	rad/pu
K_D	Integral deviation	-
K_M	Measure gain	pu/pu
T_1	Transient regulator time constant	s
T_2	Regulator time constant	s
T_M	Measure time delay	s
v^{ref}	Reference voltage	pu
x_L	Inductive reactance	pu
x_C	Capacitive reactance	pu
α^{max}	Maximum firing angle	rad
α^{min}	Minimum firing angle	rad

ii. The Thyristor Controlled Series Compensator (TCSC)

The Thyristor Controlled Series Compensator (TCSC) allows varying the series reactance of a transmission line and, thus, regulating the active flow through the transmission line itself. The functioning of the TCSC is similar to the SVC, but for the fact that the TCSC is a series device, as shown in Figure 17 below.

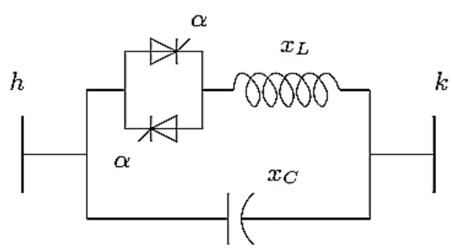

Figure 17 TCSC Scheme [62]

Figure 18 shows the TCSC control diagram.

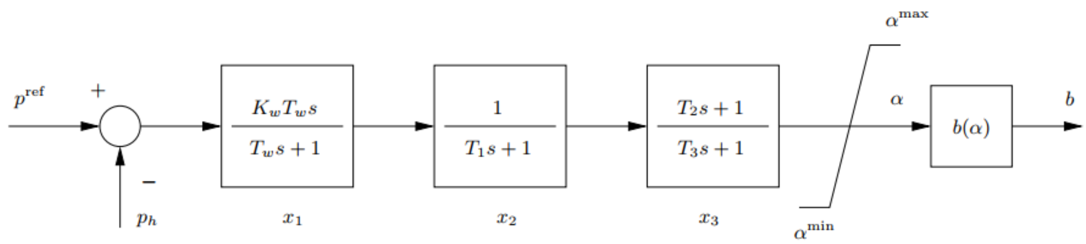

Figure 18 TSCS Control Diagram

Table 20 gives the description of TSCS parameters [52].

Table 20 TSCS Parameters

Variable	Description	Unit
K_w	Regulator gain	pu/pu
p^{ref}	Reference power	pu
T_w	Washout time constant	s
T_1	Low pass time constant	s
T_2	Lead time constant	s
T_3	Lag time constant	s
x_L	Inductive reactance	pu
x_C	Capacitive reactance	pu
α^{max}	Maximum firing angle	rad
α^{min}	Minimum firing angle	rad

iii. The Static Synchronous Compensator (STATCOM)

The Static Synchronous Compensator (STATCOM) is a shunt-connected VSC-based FACTS device that regulates the voltage of the ac bus to which it is connected. The detailed model consists of a shunt connected VSC device with a capacitor in the dc side. The model is composed of three parts, namely the dc network, the VSC and the controllers.

1. Dc network: The dc side is a parallel RC. Furthermore, the dc network must contain two nodes, one of which is connected to the ground. The VSC and the RC element are connected in parallel to the dc nodes.
2. Shunt-connected VSC model.
3. Regulators: The ac voltage control is obtained regulating the modulating amplitude.

Figure 19 shows the AC and DC STATCOM control diagrams.

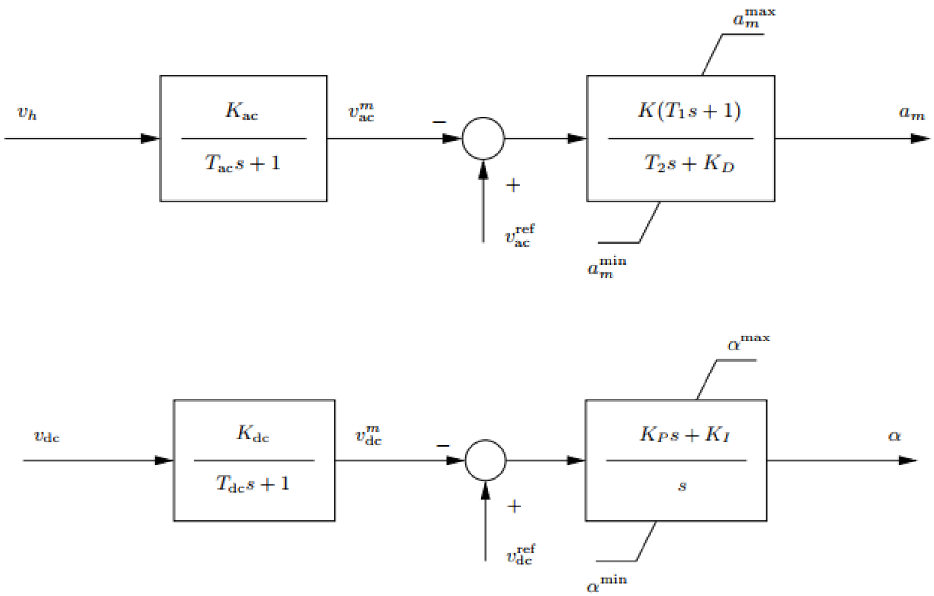

Figure 19 AC and DC STATCOM Voltage Control Diagrams [62]

Table 21 shows the brief description of STATCOM parameters.

Table 21 STATCOM Detailed Model Parameters [62]

Variable	Description	Unit
K	Gain of the ac voltage control	pu/pu
K_{ac}	Gain of the ac measurement	pu/pu
K_D	Integral deviation of the ac voltage control	-
K_{dc}	Gain of the dc measurement	pu/pu
K_I	Integral gain for the dc voltage control	rad/pu/s
K_P	Proportional gain for the dc voltage control	rad/pu
i^{max}	Maximum current	pu
i^{min}	Minimum current	pu
r_{dc}	Resistance of the dc circuit	pu
r_T	Resistance of the ac circuit	pu
T_1	Transient time constant of the ac voltage control	s
T_2	Time constant of the ac voltage control	s
T_{ac}	Time constant of the ac measurement	s
T_{dc}	Time constant of the dc measurement	s
v^{ref}	AC reference voltage	pu
v_{dc}^{ref}	DC reference voltage	pu
x_T	Reactance of the ac circuit	pu

iv. The Static Synchronous Series Compensator (SSSC)

The Static Synchronous Series Compensator (SSSC) is a series-connected VSC-based FACTS device that regulates the active power flow between the two ac buses to which it is connected. The detailed model consists of a series connected VSC device with a capacitor on the dc side. The model is composed of three parts, namely the dc network, the VSC and the controllers. Figure 20 shows the SSSC scheme.

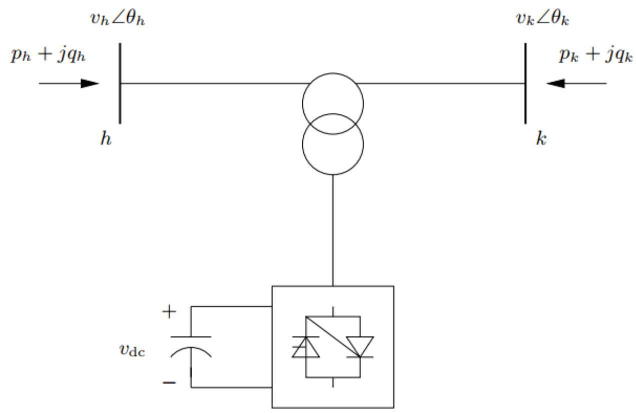

Figure 20 SSSC Scheme [62]

Figure 21 outlines the SSSC control schemes.

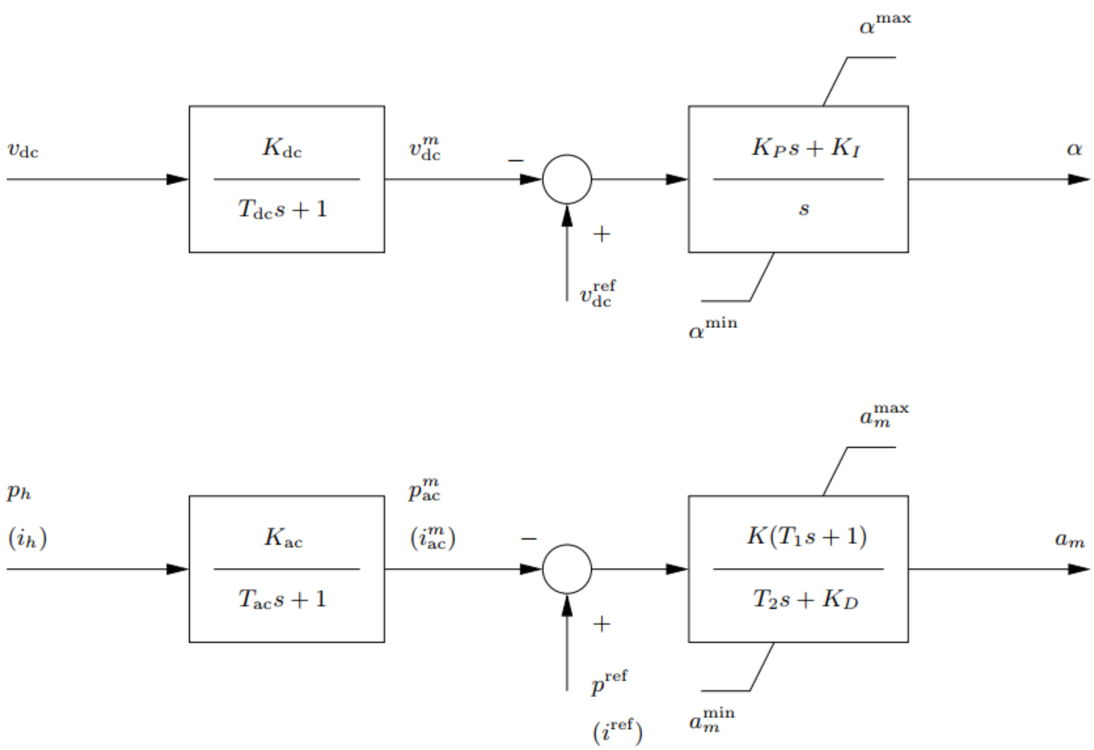

Figure 21 SSSC Control Schemes

Table 22 describes the SSSC parameters briefly.

Table 22 SSSC Parameters [62]

Variable	**Description**	**Unit**
K	Gain of the ac voltage control	pu/pu
K_{ac}	Gain of the ac measurement	pu/pu
K_D	Integral deviation of the ac voltage control	-
K_{dc}	Gain of the dc measurement	pu/pu
K_I	Integral gain for the α control	rad/pu/s
K_P	Proportional gain for the α control	rad/pu
$p^{ref}(i^{ref})$	AC reference power (current)	pu
T_1	Transient time constant of the ac voltage control	s
T_2	Time constant of the ac voltage control	s
T_{ac}	Time constant of the ac measurement	s
T_{dc}	Time constant of the dc measurement	s
v_{dc}^{ref}	DC reference voltage	pu

3.5.3.10. Load Class

The load class contains data about energy consumption and system load. Contains data about the conform load (Load that follows a daily and seasonal load variation pattern.), conform load group and schedule, Non Conform Load (loads that do not follow a daily load change pattern and changes are not correlated with the daily load change pattern), non-conform load group and schedule, customer load, load shedding data and day type data.

High amount of electricity is spent on lighting and heating. Traditional electric bulbs consume no reactive power and their power demand is frequency independent but, as the temperature of the filament depends on the voltage, the bulb cannot be treated as a constant impedance. Heating loads basically constitute a constant resistance. If a heater is equipped with a thermostat, the thermostat will maintain constant temperature and power output despite any variations in voltage. In such cases, the load can be modelled as a constant power rather than a constant resistance [63].

i. Load Models

Different types of load models are described below.

- **Constant Power/Current/Impedance**

The simplest load models assume one of the following features [63]:

- a constant power demand (P), used for load flow calculations.
- a constant current demand (I), represents a mix of resistive and motor devices.
  a constant impedance (Z), mainly represents lighting Loads.

- **Polynomial or ZIP Model**

It combines features of Z, I and P load models.

$$P = P_o\left[a_1\left(\frac{V}{V_o}\right)^2 + a_2\left(\frac{V}{V_o}\right) + a_3\right]$$

$$Q = Q_o\left[a_4\left(\frac{V}{V_o}\right)^2 + a_5\left(\frac{V}{V_o}\right) + a_6\right]$$

V_o, P_o, and Q_o are initial operating conditions. In the absence of any detailed information on the load composition, the real power is usually represented by the constant current model while the reactive power is represented by a constant impedance [63].

Table 23 shows the ZIP load parameters and their description.

Table 23 ZIP Load Parameters [62]

Parameter	Description	Unit
k_{pi}	Active current	% or pu
k_{pp}	Active power	% or pu
k_{pz}	Conductance	% or pu
k_{qi}	Reactive current	% or pu
k_{qp}	Reactive power	% or pu
k_{qz}	Susceptance	% or pu

- **Frequency Dependent Load**

Parameters for frequency dependent load and their description is shown in the Table 24.

Table 24 Frequency Dependent Load Parameters [62]

Parameter	Description	Unit
k_p	Active power	%
k_q	Reactive power	%
T_f	Filter time constant	-
α_p	Active power voltage exponent	-
α_q	Reactive power voltage exponent	-
β_p	Active power frequency exponent	-
β_q	Reactive power frequency exponent	

Table 25 shows the typical load exponent values for different types of loads.

Table 25 Typical Load Exponents [62]

Load	$\boldsymbol{\alpha_p}$	$\boldsymbol{\alpha_q}$	$\boldsymbol{\beta_p}$	$\boldsymbol{\beta_q}$
Filament lamp	1.6	0	0	0
Fluorescent lamp	1.2	3.0	-0.1	2.8
Heater	2.0	0	0	0
Induction motor (half load)	0.2	1.6	1.5	-0.3
Induction motor (full load)	0.1	0.6	2.8	1.8
Reduction furnace	1.9	2.1	-0.5	0
Aluminum plant	1.8	2.2	-0.3	0.6

Active and Reactive Power constraints on loads are defined below. Active power constraints:

$$p_D^{min} \leq p_D \leq p_D^{max}$$

The reactive power can be a function of the active power demand through a constant power factor:

$$q_D = p_D \frac{\sqrt{1 - cos^2\phi_D}}{cos\phi_D} = p_D tan\phi_D$$

ii. Induction Machine

Induction machine models can be formally formulated using the Park's approach. However, since induction machine rotors have no salient poles and since the rotor angular position is generally irrelevant, the Park's two-reaction approach (an approach according to which machine armature magnetomotive force can be split into direct and quadrature axes) is not strictly necessary. Induction machines can be modelled in a variety of ways some of which are single cage model and double cage model. Further classified into first order, second order, fifth-order, sixth-order, 14th-order and 15th order models [62]. A good proportion of electricity is consumed by Induction Motors. These types of motors are commonly found in industrial applications.

Detailed description of these models is beyond the scope of this thesis. Detailed list of parameters required to model induction machine is defined below. Figure 22 represents one type of the induction machine model equivalent circuit (IEC 61970-302).

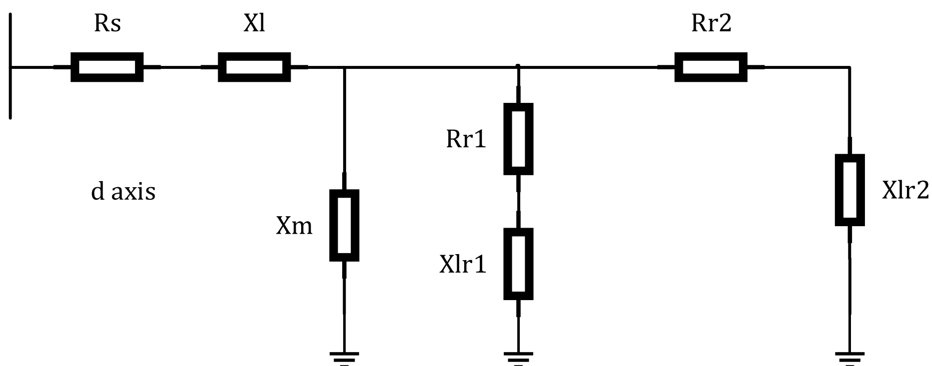

Figure 22 Induction Machine Equivalent Circuit [66]

$$X_s = X_m + X_l$$

Below table gives the description and typical values of induction machine dynamic parameters (IEC 61970-302).

Table 26 Induction Machine Dynamic Parameters and Typical Values [66]

Parameter	Description	Limit	Typical Value pu
D	Damping coefficient	≥0	0
H	Inertia constant	>0	3
S_1	Saturation factor	≥0	0.02
S_{12}	Saturation factor 120	$\geq S_1$	0.12
x_l	Stator leakage reactance	≥0	.15
r_s	Stator resistance	≥0	0.005
x_s	Synchronous reactance	$\geq x_p$	1.8
x'	Transient reactance	$\geq x''$	0.5
x''	Sub transient reactance	$>x_l$	0.2
$T'o$	Transient stator time constant	$\geq T''o$	5s
$T''o$	Sub transient stator time constant	>0	0.03s

3.5.3.11. Measurement Class

Measurement class contains dynamic measurement data. Dynamic data is provided by PMUs and other IEDs. Data from PMUs is received by Phasor Data Concentrators. Measurement class data consists of accumulator values (data accumulated in PDCs), accumulator limit, analog measurements, analog limit, command (Command is a discrete control used for supervisory control), control (Control is used for supervisory/device control. Control data represents control outputs that are used to change the state in a process, e.g. close or open breaker), control type, discrete measurements and measurements quality.

- **SCADA**

Contains information used by SCADA applications. Supervisory control supports operator control of equipment, such as opening or closing a breaker. Data consists of information about the communication link, remote control, remote source, remote unit and type of remote unit.

i. **Communication Links**

In the SCADA system, connection to remote units is through one or more communication links. Redundant links may exist. Communication Links class inherits its values such as name, mRID and description from the Power System Resource class (PSR).

ii. **Remote Controls**

Remote controls are outputs that are sent by the remote unit to actuators in the process. Remote controls inherit objects such as name, mRID, description from PSR class and has its own objects described below.

Name	Description
RemoteControlled	Set to 1 if actuator is remotely controlled.
ActuatorMaximum	The maximum set point value accepted by the remote-control point
ActuatorMinimum	The minimum set point value accepted by the remote-control point.

iii. **Remote Unit**

A remote unit can be an RTU, IED, substation control system, control center etc. The communication with the remote unit can be through various standard protocols (e.g. IEC 61850) or non-standard protocols (e.g. DNP, RP570 etc.). A remote unit contains remote data points that might be telemetered, collected or calculated. The RemoteUnit class inherits from Power System Resource Class. The intention is to allow remote units to include Measurements. These measurements can be used to model unit status as operational, out of service, unit failure etc.

iv. **Remote Point**

For an RTU, remote points correspond to telemetered values or control outputs. Other units (e.g. control centers) usually also contain calculated values.

v. **Remote Source**

Remote sources are state variables that are telemetered or calculated within the remote unit. Remote Source class inherits values from Identified object class and has its own values/objects defined below.

Name	**Description**
Deadband	The smallest change in value to be reported.
SensorMaximum	The maximum value the telemetry item can return.
SensorMinimum	The minimum value the telemetry item can return.
ScanInterval	The time interval between scans.

- **WAMS**

WAMS is the advanced and combined application of synchronized phasor measurement, communication engineering, and information technology in power systems. The aim of WAMS is to realize dynamic monitoring, analysis, and control for stable and efficient operation of the global power system. At present, the WAMS research mainly concerns the following two aspects [67]:

1. The construction and application of WAMS in grid.
2. The stability analysis and control based on wide area measurements.

The structure of WAMS mainly includes the following three parts:

1. PMU devices distributed in different areas for measuring local operating variables (e.g., voltage and current);
2. A monitoring and control system located in the power system dispatch center;
3. A digital communication network in charge of information exchange between PMU devices and monitoring system.

Regarding the advanced wide-area control and protection strategies based on WAMS, there are few applications in the practical power systems, and most of related researches only stay at the theoretical stages. In fact, the wide-area information provided by WAMS can be further utilized by different kinds of control devices (e.g., PSS, HVDC, and FACTS) to form a wide-area control strategy for the overall stability enhancement of large power systems [67].

3.6. Class Hierarchies and UML Class Diagrams

Within a system, a class represents a specific type of object being modelled. A class hierarchy is an abstract model of a system defining every type of component within a system as a separate class. A class hierarchy should reflect the real-world structure of the system. Unified Modelling Language (UML) class diagrams provide a useful means of visually representing object hierarchies [27].

Some more important definitions:

- **Inheritance** (also known as Generalization) defines a class as being a sub-class of another class. As a sub-class, it inherits all the attributes of its parent, but can also contain its own attributes.
- **Association** defines relationships other than the parent-child relationship.
- The **Aggregation** relationship defines a special kind of association between classes, indicating that one is a container class for the other.
- **Composition** is a specialized form of aggregation where the contained object is a fundamental part of the container object. This relationship implies that if the container is destroyed then all the objects related to it via composition are similarly destroyed.

3.7. CIM Equipment Model Parts (EQ)

Equipment Model Parts define the equipment components and their connectivity, which is the core description of the electrical grid and the foundation on which all CIM network modeling rests. In CIM network modeling, an EQ Model Part instance always represents just one state of the network, while the evolution of the network through time is captured as EQ Incremental Model Parts whose content describes changes.

There are two distinct aspects of a Model Part: Its description and content [68]. The description provides information about the Model Part as a whole. It contains labeling that says who is responsible for the Model Part, why it exists, when it was created and any other information relevant to business processes in which the Model Part plays a role. The Model Part content provides the actual network model or case information itself. The focus is on the content of an EQ Model Part, whose content is the description of power grid components and their connectivity.

The core of power grid modeling as a whole is describing the set of components (conducting equipment) that make up the network, along with the specification of what is connected to what electrically. CIM has three organizational concepts that help to navigate and understand this equipment data:

- An equipment containment hierarchy supports user navigation within Model Parts or assemblies of Model Parts down to individual equipment details.
- Model Parts are organized into frameworks according to what makes the best macro-modeling assembly process.
- Customized geographic and schematic diagrams may be created which facilitate visualization and human understanding.

Figure 23 outlines the CIM equipment containment hierarchy. Both Substation and Line inherit from the GeographicalRegion Class [68]. Geographical region may consist of specific control areas. It might represent countrywide control system and equipment. Similarly, geographical region is further divided into sub geo graphical regions.

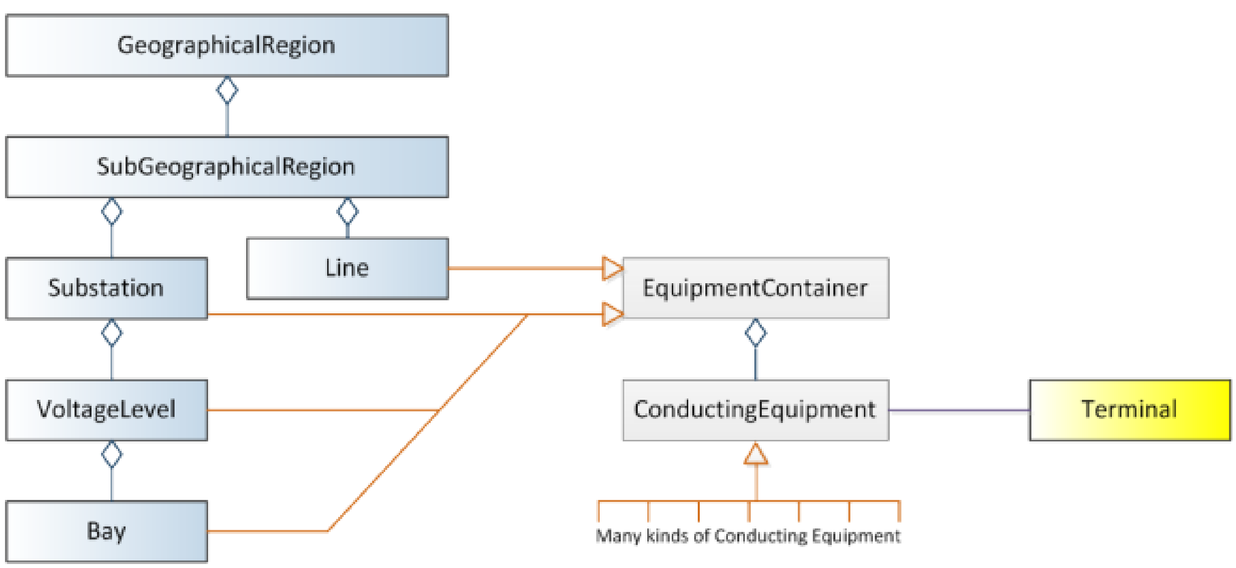

Figure 23 CIM Equipment Containment Hierarchy [27]

3.7.1. Additional Model Parts

All of the additional Model Part types are dependent on EQ Model Parts, meaning that they have objects in their content which have dangling references that point to objects in the content of EQ Model Parts [68].

3.7.1.1. SC Physical Network Model Part

Short circuit analysis requires the basic information supplied by EQ Model Parts plus additional construction detail necessary for accurately computing fault currents of various types. SC is considered a separate Physical Network Model Part in CIM, defined as one of several profiles in the IEC 61970-452 specification.

3.7.1.2. DY Physical Network Model Part

Dynamic studies require all steady-state properties, because they always use a steady-state as a starting condition, and then they need a large set of additional physical network properties. They do not introduce any new kinds of physical objects. Dynamic data is supplied as a separate kind of Model Part, called Dynamics (DY). Objects in the content of DY Model Parts reference corresponding objects in the content of EQ Model Parts.

3.7.1.3. DL Physical Network Model Part

Most network models and network analysis data are difficult to interpret without schematic and/or geographic diagrams. There are many kinds of diagrams. A partial list includes:

- Station schematic one-lines for SCADA operation. (Correspond to CIM Substation container.)
- Circuit schematic diagrams. (Correspond to CIM Line container.)
- Control room wallboard schematics.

CIM is designed on the premise that production and maintenance of these diagrams should be integrated with the production and maintenance of the EQ Model Parts that define the equipment shown on the diagrams.

3.7.1.4. GL Physical Network Model Part

The geographical location or geographical footprint of equipment is modeled in the CIM UML and is an important set of information for presentation of network analysis results.

3.8. XML

Extensible Markup Language (XML) is recognized as a universal format for structured data and documents. XML is quickly developing into a standard for storing structured machine-readable data which can be accessed all over the internet. XML allows users to design their own markup language according to their data structure because XML is basically a meta-language [69].

XML itself has not pre-defined semantics or tag syntax; however, it allows to define them according to user application in XML notation. Pre-defined syntax according to application needs must be provided to the application utilizing XML data otherwise it won't be able to interpret it. Therefore, in the power system applications because utilities have to exchange data, so, a standard semantic defined by CIM schema are followed which places a constraint on the content and structure of XML document.

The XML semantic based on CIM defines the relationships between power system classes such as the parent-child relationships. It also defines which power system element can have which data types, attributes and if the values must be fixed or changing.

3.8.1. Resource Document Framework (RDF)

In a basic XML document, the relationship between classes representing tables cannot be defined. XML Resource Document Framework or XML RDF schema solves this problem by introducing a framework that helps to define relationships [69]. The RDF Schema allows the user to describe the classes and properties themselves and indicate when they should be used together. The RDF combined with RDF Schema provides a mechanism for expressing a basic class hierarchy as an XML schema by specifying the basic relationship between classes and properties.

As has already been discussed that there is a need to exchanging power system model and data between the utilities due to reasons already identified, the CIM for Power Systems has the potential to maintain multiple copies of the same data in multiple formats while the XML, combined with RDF offers a means of storing the data in a format compatible with every piece of software, requiring the removal of application-specific data and a subsequent loss in precision [27].

4. IMPLEMENTATION OF THE DATABASE MANAGEMENT SYSTEM

The task of this thesis was to design and implement a DBMS for digital twinning of power system. This DT centered DBMS is an important part of power system control center. The core of DBMS is the database that stores all the information about power system operational, maintenance and planning activities. Owing to fast data rates of modern PMUs and other IEDs a proper database schema had to be chosen to cope up with the complex nature of data and be extendable in future with added advantages of robustness and security.

The database management system has been implemented using the PostgreSQL database schema. SQL is a power-packed language for manipulation and querying the data; however, it presents challenges when trying to integrate applications. Therefore, SQLAlchemy Object Relational Mapper (ORM) has been used for the implementation of database backend. SQLAlchemy ORM is the combination of Python-based SQL toolkit and Object Relational Mapper that gives application developers the full power and flexibility of SQL. SQLAlchemy helps in mapping Python objects to the tables of the database without complicating the existing Python code. It can be used for connecting the common databases such as MySQL, Postgres, Oracle and SQLite. SQLAlchemy is most famous for its object-relational mapper (ORM) which is an optional component that provides the data mapper pattern, where classes can be mapped to the database in open-ended, multiple ways - allowing the object model and database schema to develop in a cleanly decoupled way from the beginning. The SQLAlchemy Object Relational Mapper presents a method of associating user-defined Python classes with database tables and instances of those classes (objects) with rows in their corresponding tables [70]. Specifically, Object Relational Mapper (ORM) of SQLAlchemy toolkit is used for the implementation of the digital twin database. Matlab acts as database front end in this DBMS and OPC Servers have been used to copy SCADA/WAMS behaviors.

Initially, IEEE modified 9 bus system is used for the understanding and realization of the database. It is a very simple system consisting of 9 buses, 3 generators, 3 transformers, 3 loads and 6 transmission lines. These steps are followed in order to design and evaluate the database.

1. Determine the fields that make up the table.
2. Determine relationships between the tables.
3. Test the database.
4. Input data.
5. Test data integrity.

Three distinct codes are written utilizing the SQLAlchemy ORM mapper. One code defines the database classes, the second one initializes these classes into the defined database schema. Third code reads data from the excel sheets using the pandas library in python and writes the values into the database tables. Figure 24 gives the circuit diagram of IEEE 9 bus system [71].

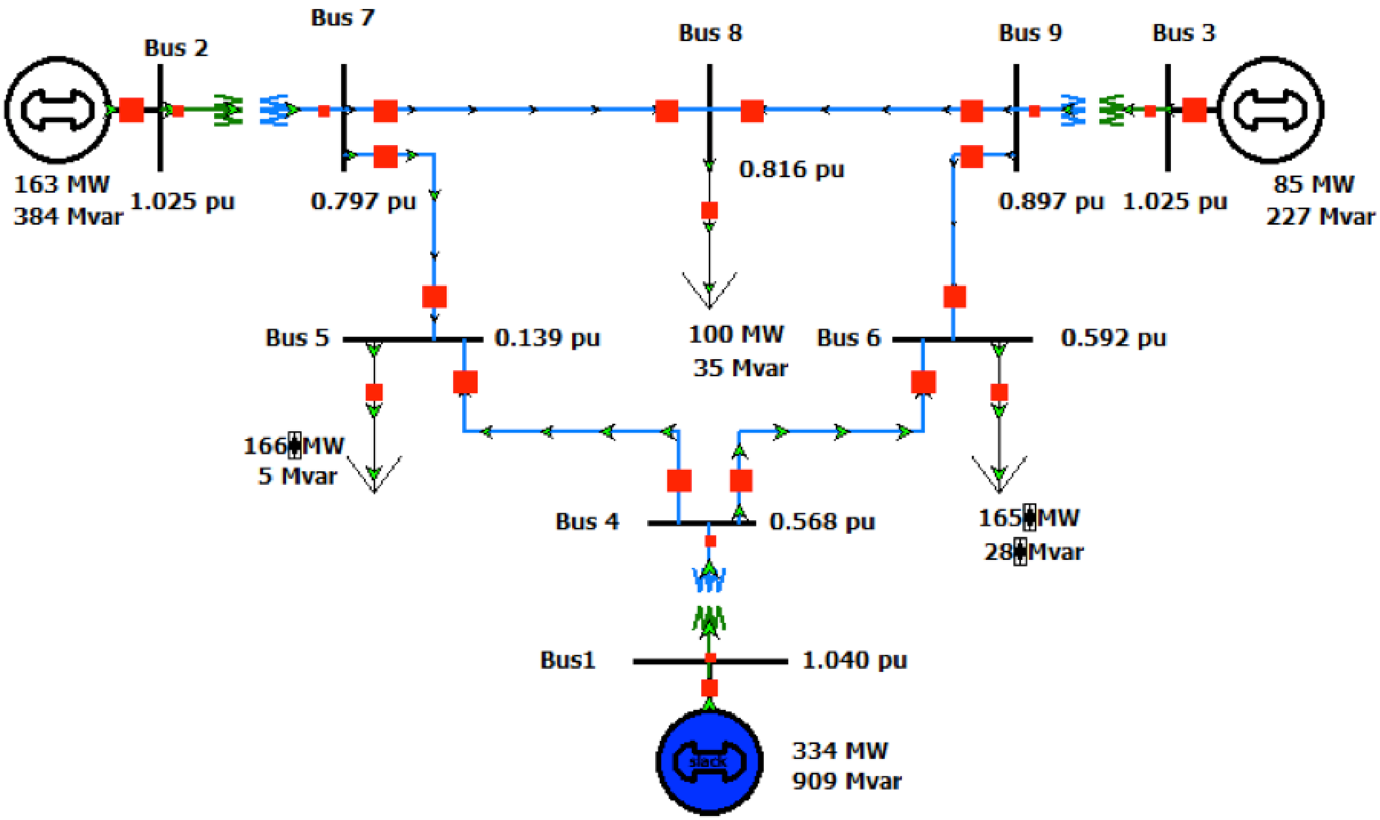

Figure 24 IEEE 9 Bus System [71]

4.1. Conversion of Power Network into CIM Model

A methodology for conversion of power system model/network or circuit into CIM model is discussed in this section. A conversion of 9 bus system into CIM model has been done and the datasets have been attached in the Appendix 1.

The CIM defines a comprehensive logical view of the power system elements and represents their attributes as well as the relationships using a set of Unified Modeling Language (UML) Class Diagrams. Figure 25 below gives the simplified CIM topological model for conversion of any power system model [23].

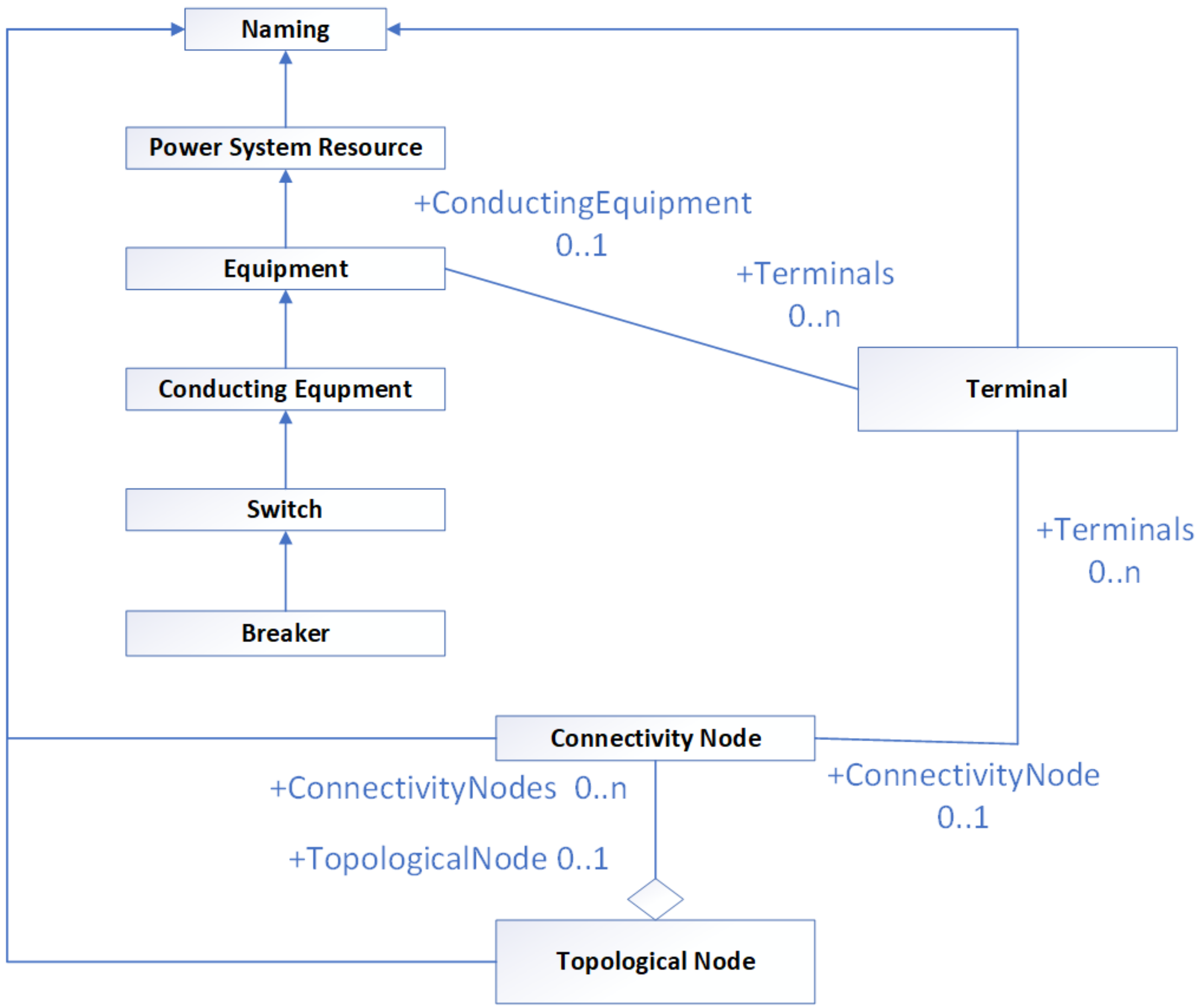

Figure 25 CIM Topological Model [23]

For a general power system this CIM topological model can be used to transform power system network into UML Diagram. These UML diagrams can be translated into relational database syntax. Then the power system network model can implement on a Relational Database. Based on the CIM topological model, equipment in a power system and its topology can be represented as suggested below in Figure 26.

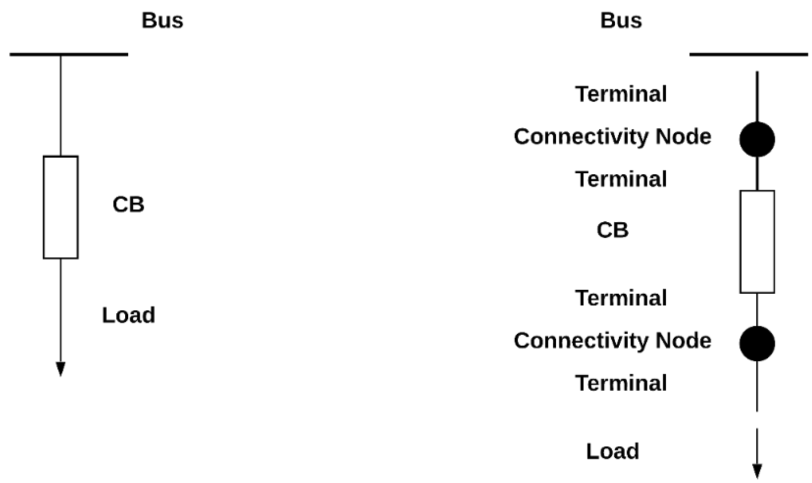

Figure 26 CIM representation of an Object/Component

All the power system objects with assigned attributes can be assembled into a standard XML/RDF file as exchangeable data and stored into a corresponding database for common access. The CIM based database should reflect the relationship, such as aggregation and association in CIM/UML diagram. For power system applications, EPRI has decided to use XML/RDF schema as the format for exchanging data. Therefore, when two components or applications within a utility or between utilities want to exchange information, the actual message has to be implemented based on the CIM tag convention in XML/RDF format first. The CIM XML/RDF Schema specifies the format of how these messages are constructed [23]. Figure 27 shows the entity-relationship/UML diagram of the relational database schema based on the CIM topology.

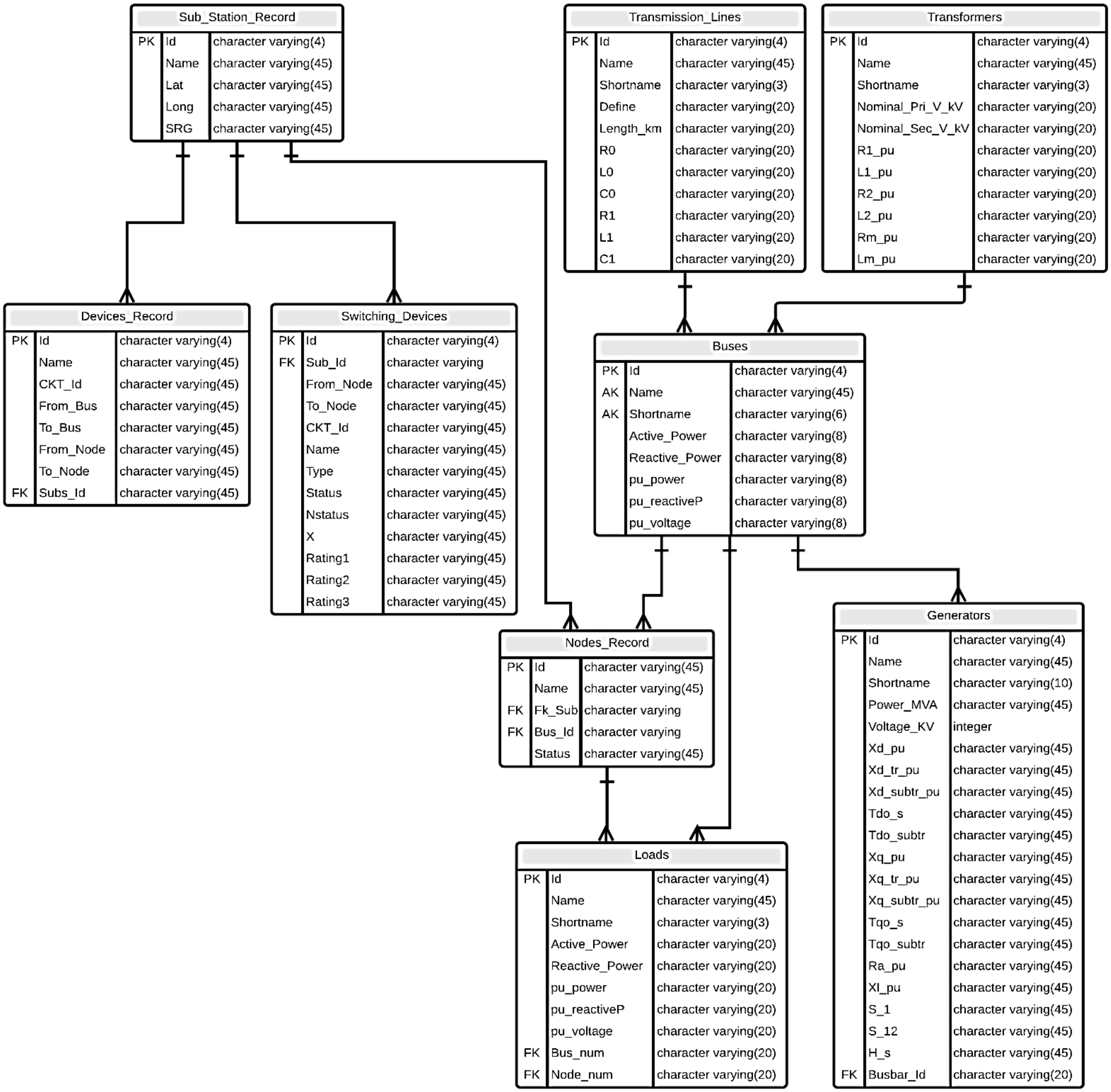

Figure 27 Entity Relation Diagram

Entity relation/UML diagram shows the relationship between different tables in the database. For example, there is one to many relationships between the Sub_Station_Record and Devices_Record table. This ERD maps the CIM database model to a relational database.

Next step in this work is conversion of CIM schema into CIM RDF XML format because it has already been discussed that RDF schema address the general problem of representing entities and relationships and RDF is an abstract model that defines relationships between entities (called resources in RDF). RDF is a foundation for processing metadata. It provides interoperability between applications that exchange machine-understandable information. Based on the CIM/RDF schema, a CIM conformed data file can be developed. This CIM XML file using RDF schema could be used as a message file among applications.

4.2. CIM UML Diagrams for Database Implementation

Below a simplified CIM UML diagram for DT database management system application is shown. This CIM UML diagram has been generated using IBM Rational Software Architect Designer version. The advantage with using IBM software is that it can generate UML to XML files. Figure 28 outlines the UML Diagram based on the CIM recommendations for any power system network not specific to the test cases (IEC 61970-301).

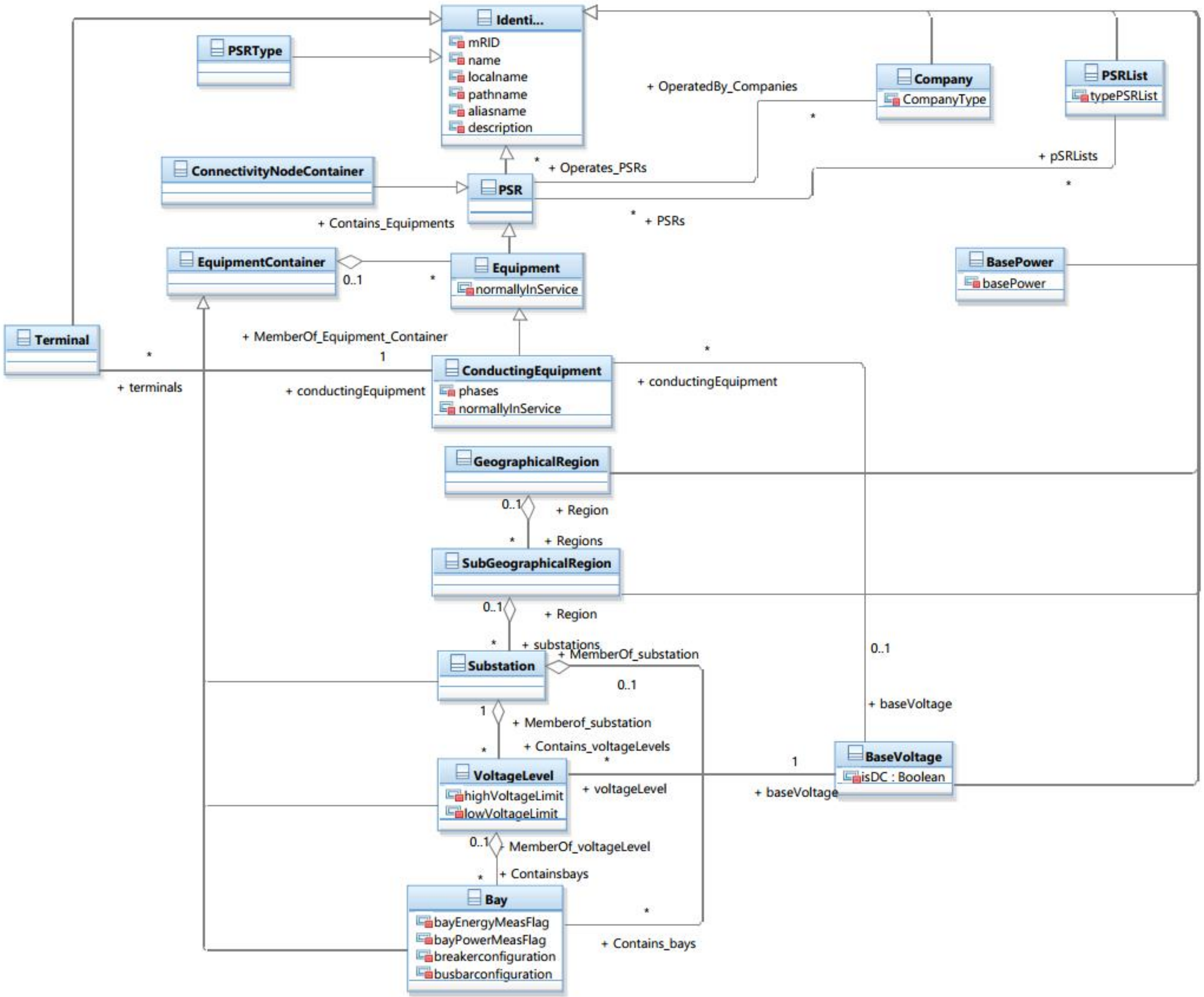

Figure 28 Power System CIM UML Diagram [27]

Figure 29 represents the power system classes inheritance structure and hierarchy.

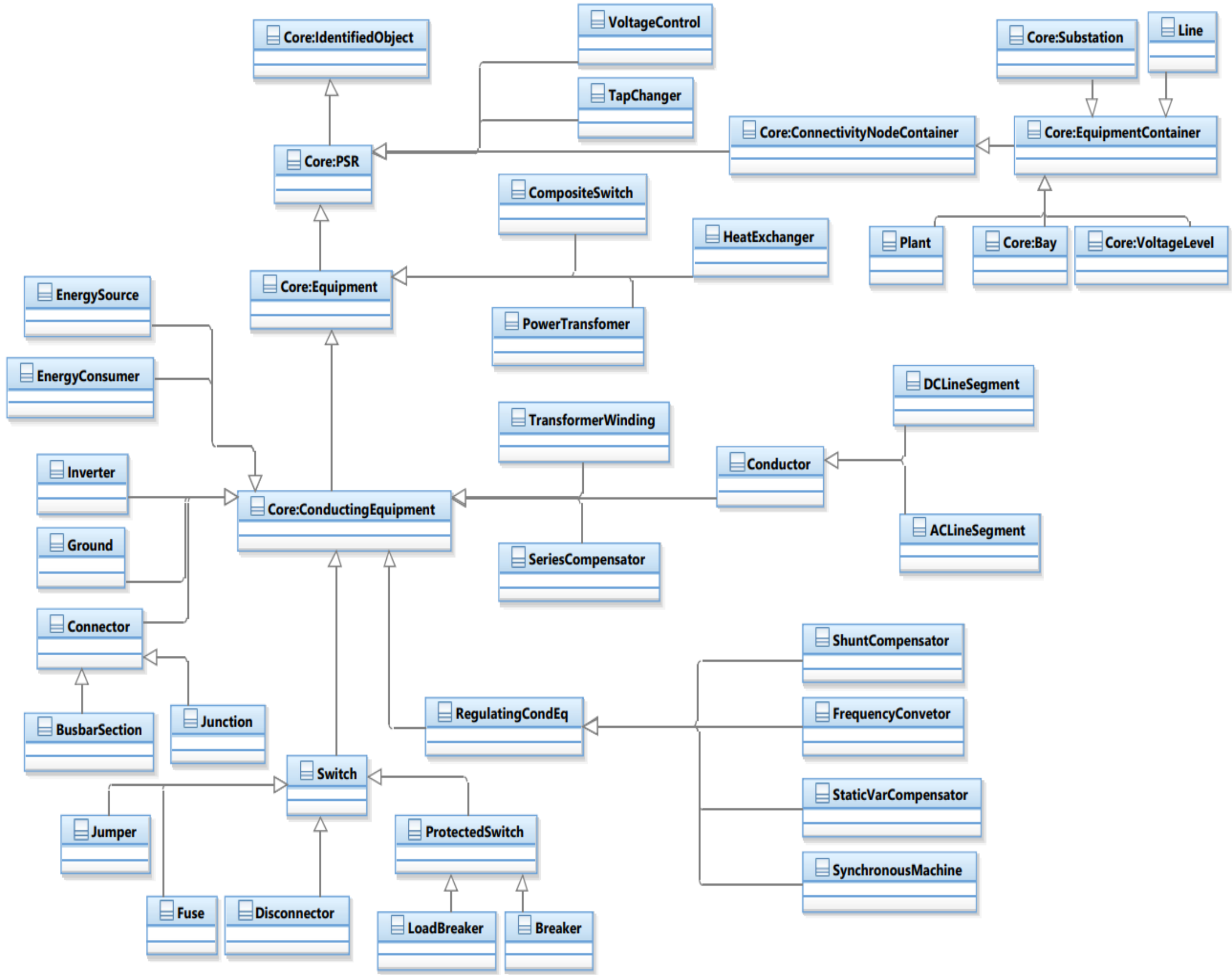

Figure 29 CIM Power System Inheritance Structure [27]

Figure 30 outlines the power system CIM Topological model which is based on the IEC 61970-301.

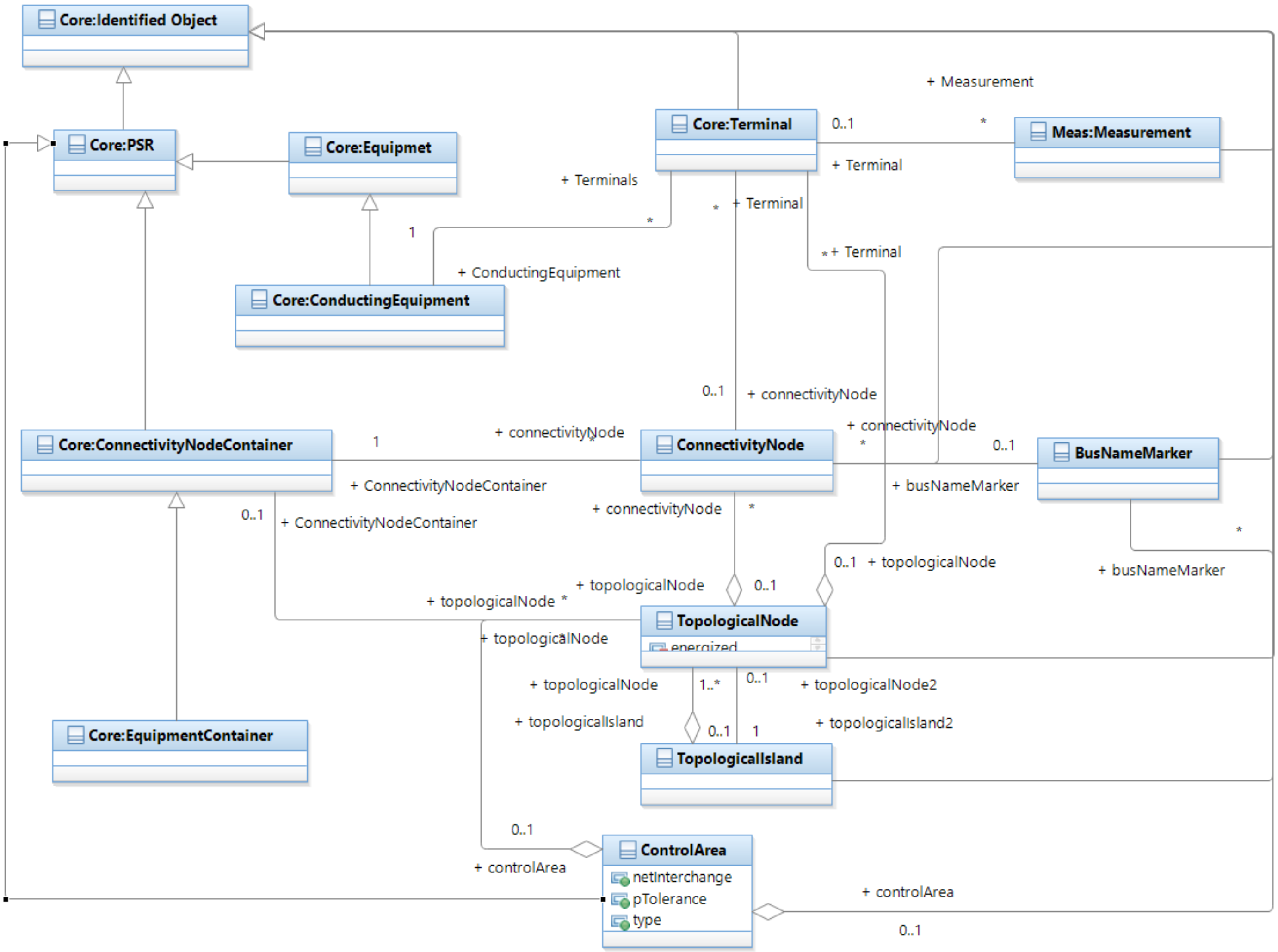

Figure 30 Power System CIM Topology [27]

Figure 31 below gives the operational limits of the power system in a single UML package. It defines the relationship between different classes which deal with the power system operational limits. This diagram is very important from the point of view of data cross check validation algorithm because devices operational limits are defined in this UML package.

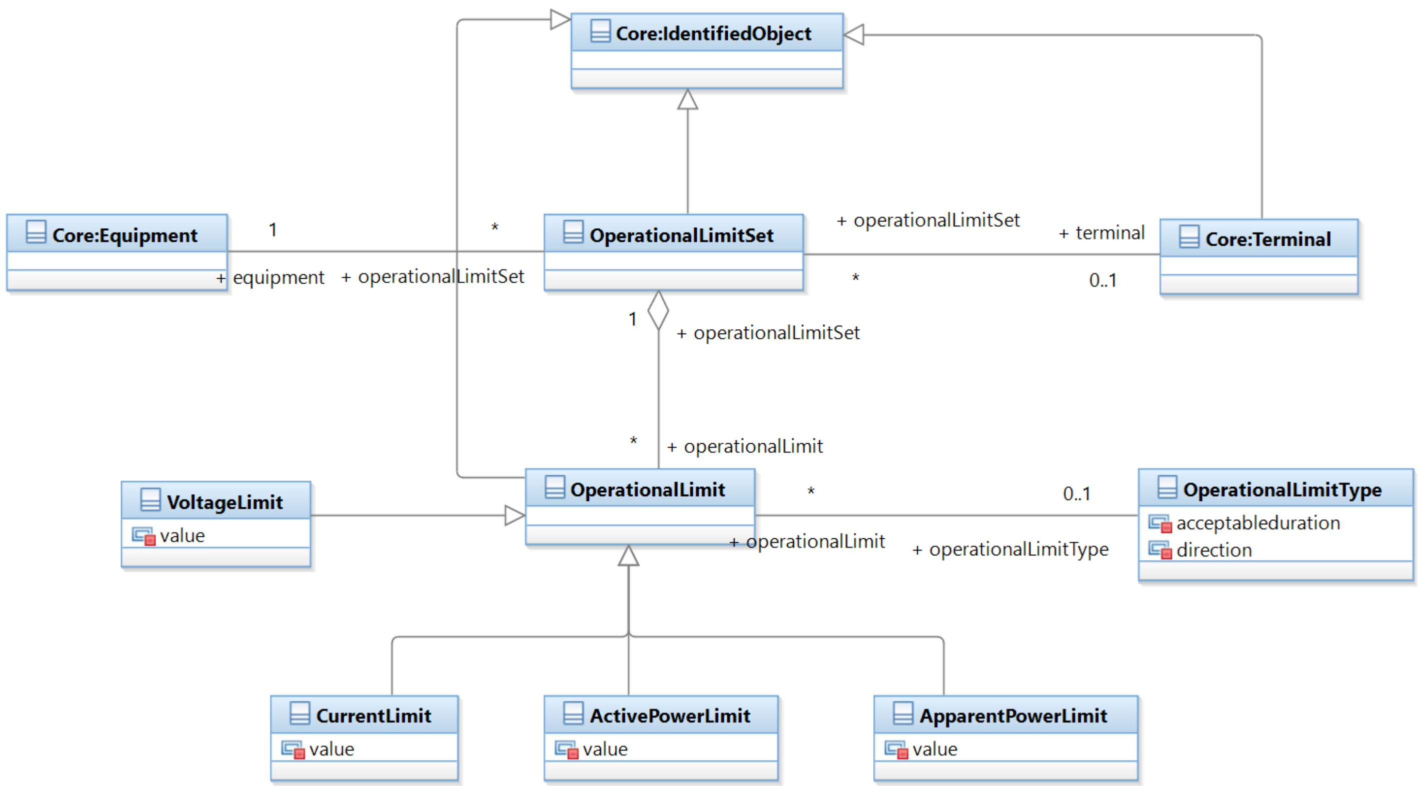

Figure 31 Operational Limits [27]

Again, these diagrams have been drawn with the help of IBM software keeping in mind the CIM UML specification in the standard IEC 61970-301. Above four diagrams provide us the basic architecture for implementing the CIM UML schema in a relational database for complex power system networks. Detailed models for power system components like transformer, generator and transmission line etc. can be embedded in these packages and then we can expand these CIM UML diagrams to give us the full picture of relational database implementation in the Database Management System. For the purpose of exchange of data, UML to XML files have been generated with the help of IBM Rational software.

4.3. DBMS Schematic

Figure 32 represents the schematic diagram for database management system. Here the database block is the PostgreSQL database based on the Object Rational Mapper (ORM) schema. MATLAB is used as a database frontend; it can read and write data into the database using JDBC driver. Different analysis like load flow analysis, transient analysis and stability analysis can be performed utilizing the PowerFactory and store results into an excel file which can read into the database using python code.

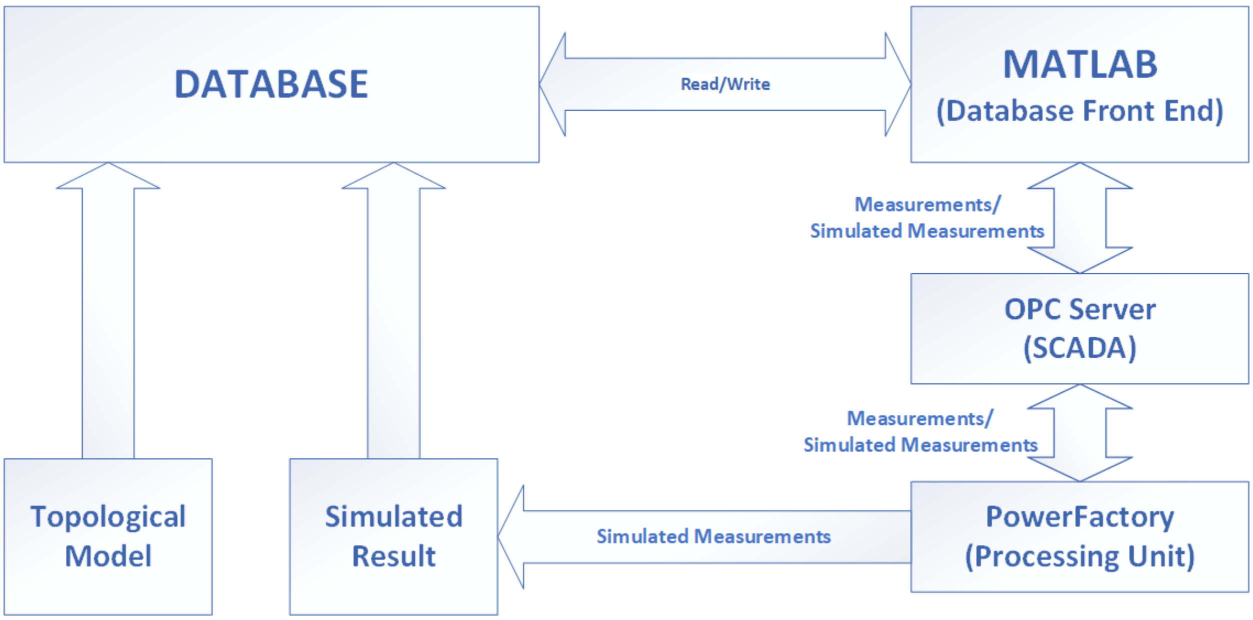

Figure 32 DBMS Schematic

Figure 32 describes the overall database management system structure. Because of lack of access to a practical power system where real-time measurements can be read and stored in the database, a test setup has been designed as shown in above figure. Here database is the PostgreSQL database which is based on the power system topological model, for the test cases, IEEE 39 and 9 bus system models are used. Matlab is acting as the database front end for the test cases in the real-world application of this DBMS a different proprietary software can be used as database front end. As database front end Matlab gives user interface for accessing data from the database. OPC Server acts as the SCADA/WAMS system. PowerFactory acts as the processing unit where time domain, load flow and state estimation simulations can be run. PowerFactory needs to be configured so it can receive external measurements and send back to the OPC Server as simulated measurements.

Open Client Communication (OPC) Servers are utilized to provide communication between the database, Matlab and Power Factory [72]. Matlab provides the database frontend. So, database parameters can be controlled from within the Matlab Interface. For the data processing between PowerFactory and Matlab, OPC Servers are utilized. PowerFactory can be controlled utilizing Matlab with the help of OPC server as a bridge between them. MatrikonOPC Simulation servers are utilized for this particular task because it is a freeware and easy to use.

First, the OPC Server is setup then pre-configured OPC tags are imported from the csv file. The csv file contains all the elements that are to be written or read from the OPC Server. The csv file is saved in a proper format and it contains element name and its data type along with other information if required. Next step is

PowerFactory configuration so that it can be used to communicate with OPC Server. After that, the network model (i.e. 9 bus system model) is imported into PowerFactory. OPC link must be properly configured in PowerFactory; only then a link will be established between PowerFactory and OPC Server. The link must be properly setup for Online State Estimation and Time Domain Simulation (TDS). With the educational version of PowerFactory, Online State Estimation and Load Flow calculations are not possible. Only Time Domain Simulation can be run. After proper configuration of OPC Link now it is ready to send and receive data.

4.3.1. Time Domain Simulation

Time Domain Simulation (TDS) is useful for a thorough simulation of a network dynamic behavior. After running the TDS in PowerFactory, results of simulation will be sent to the OPC Server and if a change is made in the parameter values inside the OPC server it will affect the running TDS in PowerFactory. The results of the simulation which are called simulated results in Figure 30 can be stored in the database after writing them into an excel file. Plots can be generated to visualize changing behavior. Values of system parameters can be easily visualized with the help MatrikonOPC explorer which is installed during the installation of MatrikonOPC Simulation Server. Parameter values can not only be visualized but changed from within the MatrikonOPC Explorer. Before running time-domain simulation in PowerFactory initial conditions for simulation need to be calculated within the PowerFactory.

After establishing a connection between the OPC Server and PowerFactory now if a connection can be established between the OPC server and Matlab then Figure 30 functionality is achieved. Matlab has a dedicated OPC Toolbox [73] which provides access to live and historical OPC data directly from Matlab and Simulink. With the help of this Toolbox, OPC data can be logged, read and written from a number of sources including the SCADA and PLC systems. OPC Toolbox establishes connection between the OPC server and Matlab without the need of any Python code which was previously used for the purpose. Once a connection between the OPC data access server and Matlab is established, then the next step is to create data access group objects which correspond to the collection of OPC server data access items. After data access group creation data access items can now be imported into it. This enables writing parameter values to the OPC server and monitoring of the OPC server items.

When state estimation simulations are run in PowerFactory, PowerFactory can help in detecting and eliminating bad measurements. If we change value of an external measurement to an unrealistic extent using MatrikonOPC Explorer, PowerFactory will detect this bad measurement and ignore it while performing state estimation.

To establish connection between Matlab and PostgreSQL database a simple code is presented below.

```
datasource = 'DynTest'; % DB_name
username = 'postgres';
password = 'ilmchnswN1778';
conn = database(datasource,username,password);

if isopen(conn)
    disp('Connection to SQL established')
else
    disp('Connection to SQL failed')
end
```

Connection between Matlab and database is established with the help of JDBC driver. JDBC data source needs to be properly configured to make database connection. Using JDBC driver data can be imported from database into Matlab and data analysis can be performed.

Following simple code can be used to import all data from a table named SubStation into Matlab.

```
tablename = "SubStation";
data = sqlread(conn,tablename);
head(data)
```

Database Explorer app from within the Matlab can also be used to achieve the functionality described above. Database Explorer app lists all the database tables in a Data Browser panel after correct connection with the database. Database queries can now be inserted to fetch data from database and import into Matlab.

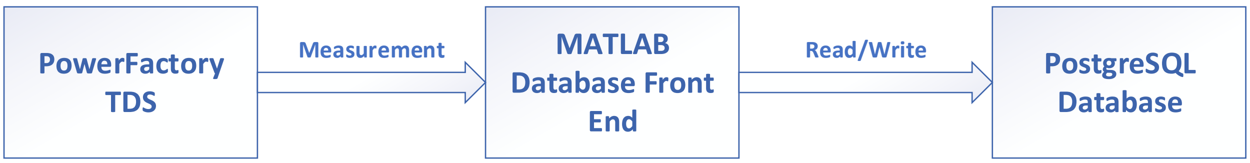

Figure 33 Measurement Schematic

Figure 33 displays simple connection scheme between PowerFactory, Matlab and the database. Connection between PowerFactory and Matlab is established indirectly using the OPC server as also explained above. The connection between Matlab and database can be established in more than couple of ways with JDBC driver and Database Explorer app few of them. PowerFactory has the network model i.e. IEEE Nine Bus System or any other complex network model of any power system. OPC server receives network parameter values from PowerFactory and changes in values are sent back to PowerFactory. As a connection is present between OPC server and Matlab parameter values can be read and written in the Matlab. As explained before, Matlab can communicate with database so it can send and receive data.

4.3.2. PgAdmin Query Tool

To access or manage the database pgAdmin software can also be utilized in place of Matlab. Queries can be inserted into the pgAdmin software provided with the PostgreSQL package. With the help of pgAdmin data management is quite easy. Queries can be written using the Query Tool [74] to view and store data in the database. PgAdmin 4 provides dialogs that allow to modify all table properties and attributes. Use the Rule dialog to define or modify a rule for a specified table or view. A PostgreSQL rule allows the definition of an additional action that will be performed when a SELECT, INSERT, UPDATE, or DELETE is performed against a table.

Here is simple query to view data in the table named Loads:

SELECT * FROM public."Loads"

Output of query is displayed below. Query time is approximately 40 milliseconds.

Data Output Explain Messages Notifications

	Id [PK] character varying (4)	Name character varying (45)	Shortname character varying (3)	Active_Power character varying (20)	Reactive_Power character varying (20)	pu_power character varying (20)
1	1	Load1	LD1	125	50	1.25
2	2	Load2	LD2	100	35	1.0
3	3	Load3	LD3	90	30	0.9

Figure 34 PgAdmin Data Management

Here is a simple query to view all the tables in a database.

```
SELECT
  *
FROM
  pg_catalog.pg_tables
WHERE
  schemaname != 'pg_catalog'
AND schemaname != 'information_schema';
```

To modify the content of a table, each row in the table must be uniquely identifiable. If the table definition does not include a primary key, the displayed data is read only. To modify the displayed data:

- To change a numeric value within the grid, double-click the value to select the field. Modify the content in the square in which it is displayed.
- To change a non-numeric value within the grid, double-click the content to access the edit bubble. After modifying the content of the edit bubble, click the Save button to display the changes in the data grid, or Cancel to exit the edit bubble without saving.

To add a new row to the table, enter data into the last (unnumbered) row of the table. As soon as the data is stored, the row is assigned a row number, and a fresh empty line is added to the data grid. To write a SQL NULL to the table, simply leave the field empty. If change in existing row is stored, the value NULL will explicitly be written. To write an empty string to the table, enter the special string ‘’ (two single quotes) in the field.

4.4. Cross-Check Validation Algorithm for Checking Data Integrity

Data integrity checking is very important for the proper operation of our data management system. A cross-check validation algorithm is developed that checks if the data input from different sources is indeed as we expect it to be. Most of the time, the power system data is based on some calculations and formulas. Power System component models such as synchronous machine dynamic models or transformer models are based on standard power system formulas. These models specify the power system component parameters based on typical formulas. So, these formulas help a lot when we have few real-time parameter measurements than the number of variables required to describe the whole model. Sometimes, utilizing these formulas unknown measurements can be calculated from the known measurement values. This gives an idea and a way to check if sensor measurements are correct as typical parameter values are usually known. Comparing typical parameter values with ones calculated in the data integrity algorithm; if there is a huge difference between the two, it can be safely assumed that there is an error in the measurement. Such measurements having too much deviation from the typical parameter values are usually not reliable to be stored in a database. It is not good to store wrong data in the database without first checking its integrity for several reasons. One reason is because nowadays data is shared among different groups within one utility and between several utilities. So, sharing of wrong data can have devastating effects on the control of power system equipment especially and may even lead to blackouts. A python-based data integrity checking algorithm is developed to achieve objectives defined above.

Some databases such as PostgreSQL can enforce relationships at the systems level, using foreign keys, constraints, and other similar concepts, but these built-in checks and constraints represent only a small set of possible characteristics that may be modeled in any given system. Therefore, there is a need to implement unit test and integration tests to make sure that the data entered into the database is valid. The type of tests to be performed depend upon the type of bugs occurring most frequently in the code and the nature of data.

4.4.1. Constraints

Data types are a way to limit the kind of data that can be stored in a table. For many applications, however, the constraint they provide is too simple. For example, a column containing a status of switching devices table should probably only accept positive values. But there is no standard data type that accepts only positive numbers. PostgreSQL allows to define constraints on columns and tables. Constraints give as much control over the data in tables as wanted. If a user attempts to store data in a column that would violate a constraint, an error is raised [75].

4.4.1.1. Check Constraints

A check constraint is the most generic constraint type. It allows specification that the value in a certain column must satisfy a Boolean (truth-value) expression. For instance, to require positive status values:

```
CREATE TABLE SwitchingDevices (
    Id integer,
    Name text,
    Status numeric CONSTRAINT positive_only CHECK (Status > 0)
);
```

A check constraint can also refer to several columns. For example, in the Generators table it is required that the sub transient reactance be always less than the transient reactance. Below is a way to achieve it:

```
CREATE TABLE Gen (
    Id integer,
    Name text,
    Xd_Tr numeric CHECK (Xd_Tr > 0),
    Xd_SubTr numeric CHECK (Xd_SubTr > 0),
    CHECK (Xd_Tr > Xd_SubTr)
);
```

Check constraint is usually implemented at the time of database tables creation. However, if tables are already present and check constraint has to inserted. ALTER TABLE statement can be used to do so. It should be noted that a check constraint is satisfied if the check expression evaluates to true or the null value.

```
ALTER TABLE Gen ADD CONSTRAINT TransientReactance CHECK (
    Xd_Tr > 0
    AND Xd_SubTr > 0
    AND Xd_Tr > Xd_SubTr
);
```

The CHECK constraints are very useful to place additional logic to restrict values that the columns can accept at the database layer. By using the CHECK constraint, it can be made sure that data is updated to the database correctly.

4.4.1.2. Not-Null Constraints

A not-null constraint simply specifies that a column must not assume the null value. A not-null constraint is always written as a column constraint. A simple example:

```
CREATE TABLE SwitchingDevices (
    Id integer Not Null,
    Name text,
    Status numeric CONSTRAINT positive_only CHECK (Status > 0)
);
```

4.4.1.3. Unique, Primary Key and Foreign Key Constraints

Unique constraints ensure that the data contained in a column, or a group of columns, is unique among all the rows in the table. If Unique constraint is combined with not Null constrain it becomes equivalent to Primary Key constraint. Primary Key, Foreign key and Unique constraints have been already used in the implementation of database.

```
CREATE TABLE SwitchingDevices (
    Id integer Unique,
    Name text,
    Status numeric CONSTRAINT positive_only CHECK (Status > 0)
);
```

Unique constraint may be defined for more than one column in a table.

```
CREATE TABLE Gen (
    Id integer,
    Name text,
    Xd_Tr numeric,
    Xd_SubTr numeric,
    UNIQUE (Xd_Tr,Xd_Subtr)
);
```

A primary key constraint indicates that a column, or group of columns, can be used as a unique identifier for rows in the table. Relational database theory dictates that every table must have a primary key. This rule is not enforced by PostgreSQL, but it is usually best to follow it. A foreign key constraint specifies that the values in a column (or a group of columns) must match the values appearing in some row of another table. Primary Key, Foreign key and Unique constraints have been already used in the implementation of database.

All of these constraints can be inserted using the pgAdmin Query tool also if not already implemented into the database schema. The Query Tool is a powerful, feature-rich environment that allows execution of arbitrary SQL commands and review the result set. The Query Tool allows to [74]:

- Issue ad-hoc SQL queries.
- Execute arbitrary SQL commands.
- Displays current connection and transaction status as configured by the user.
- Save the data displayed in the output panel to a CSV file.
- Review the execution plan of a SQL statement in either a text or a graphical format.
- View analytical information about a SQL statement.

After entering a query, select the Execute/Refresh icon from the toolbar. The complete contents of the SQL editor panel will be sent to the database server for execution. Trigger function can also be implemented in the pgAdmin which helps to create logs. A trigger function defines the action that will be invoked when a trigger fire. A trigger executes a specified function when certain events occur.

4.4.2. Unit Test

The *unittest* module provides a rich set of tools for constructing and running tests. A testcase is created by subclassing *unittest.TestCase*. The three individual tests are defined with methods whose names start with the letters test. The crux of each test is a call to *assertEqual()* to check for an expected result; *assertTrue()* or *assertFalse()* to verify a condition; or *assertRaises()* to verify that a specific exception gets raised. The *setUp()* and *tearDown()* methods allow defining instructions that will be executed before and after each test method. Method is called to prepare the test fixture. and is called immediately before calling the test method; other than *AssertionError* or *SkipTest*, any exception raised by this method will be considered an error rather than a test failure. Method called immediately after the test method has been called, is called even if the test method raised an exception, so the implementation in subclasses needs to be particularly careful about checking internal state. The basic building blocks of unit testing are test cases which are single scenarios that must be set up and checked for correctness. In *unittest*, test cases are represented by *unittest.TestCase* instances. The testing code of a TestCase instance should be entirely self-contained, such that it can be run either in isolation or in arbitrary combination with any number of other test cases. If the test fails, an exception will be raised with an explanatory message, and unittest will identify the test case as a failure. Any other exceptions will be treated as errors.

A testing unit should focus on one tiny bit of functionality and prove it correct. Each test unit must be fully independent. Each test must be able to run alone, and also within the test suite, regardless of the order that they are called. The implication of this rule is that each test must be loaded with a fresh dataset and may have to do some cleanup afterwards. This is usually handled by *setUp()* and *tearDown()* methods.

Unit Test method has been used in the implementation of cross-check validation algorithm for checking data integrity. A python code has been written using the Unit Test procedures in combination with constraints to check if the measurement values and load flow analysis data are within the pre-defined limits for the 9-bus system before storing in the database. Same code can be extended for 39 bus and larger systems.

5. CONCLUSION

The main aim of this thesis was to design and develop a database management system capable of handling requirements of power system digital twin data. Digital twin is a fairly new concept and not a lot of power system control applications of digital twin are found in practice at the moment let alone applications of digital twin in the power system control centers. Technologies to implement control centers based on digital twin are available which include communication standards, WAMS, big data storage and analysis platforms.

In the first chapter, introduction of the thesis topic and motivation behind it was presented. A brief overview of evolution of power system control centers was given and the outlook of modern control centers was also provided. It was illustrated that the traditional control centers which were mainly centralized in nature are not adequate for meeting the challenges of fast and huge amount of data provided by the IEDs and PMUs. Therefore, future control centers have to distributed in nature and able to merge and store data from both SCADA and WAMS. Distributed intelligence might help in realizing distributed control centers. It is discussed that the IoT, (enabler of digital twin) if coupled with AI, can be quite helpful in the realization of power system digital twin. Importance of utilizing CIM for data exchanges inter and intra utilities is also discussed. Emphasis has been made to build DT database management system based on the CIM guidelines. Role of database management system in control center is presented. Along with requirements of a DBMS in a power system some important attributes of a modern DBMS are also presented. Emphasis has also been made on the advantages and utilities of power system digital twin. It is discussed that the DT will enable utilities to increase the power system efficiency and help in understanding grid behavior and optimize the performance of installed equipment.

In the second chapter, power system control architecture is discussed. The common components of control center such as EMS, SCADA, WAMS have been described in detail. Then the concept of Digital Twins is described in detail which suggests that the digital twin can be thought as a replica of physical asset in digital form. Applications of DT in the field of power system are described. The different names of the DT concept used by different companies are discussed briefly. It is understood that SIEMENS and GE are contributing a lot in the implementation of DT in power systems.

In the chapter three, a database structure suitable for implementing the concept of DT is discussed. A comparison has been made between the relational and graph database architectures. In the comparison of graph and relational databases, subjective and objective measures to choose the better database structure for this application are discussed. Relational databases owing to fast query processing for integer loads, maturity, scalability, flexibility and suitability for implementation of structured data are preferred over the graph databases. Then a further comparison between different relational databases such as SQLite, MySQL and PostgreSQL have been made to make the decision easier to choose most suitable among them for this work. PostgreSQL has been found to meet the requirements of modern database management system such as data integrity, fast parallel processing, extensibility and availability of NoSQL features.

In chapter 4, power system data, models and parameters are discussed in detail. Power system equipment has been classified into different classes in order to easily implement whole power system model into the database schema. Along with the discussion of power system equipment models, parameters associated with them are also described and for some parameters their typical values and range is also discussed. CIM is also discussed in detail. XML is also discussed briefly because the CIM based database schema has to be translated into XML files so that it can be shared between different utilities.

In chapter 5, implementation of PostgreSQL database management system is discussed. IEEE 9 bus system is used as a test case for the implementation of the database. Codes were written using the SQLAlchemy ORM. A process to convert power system model into common information model has be discussed. Then CIM based database schema has been converted into XDF message. DBMS schematic has been drawn and discussed in detail which tells how PowerFactory, Matlab and OPC Server have been utilized to make the DBMS working. It has been demonstrated that checking data integrity is quite important for the efficient working of DBMS. A cross-check validation algorithm for checking data integrity has been implemented using Unit Test procedure. Future recommendations are discussed in chapter 6.

A database management system using PostgreSQL is implemented due to various advantages of PostgreSQL as described in this thesis. Matlab has been used as database front end which gives full user access to the database. PowerFactory has been used as the data processing unit on which time domain, load flow and state estimation analysis can be performed. OPC servers have been used to mimic the functionality of SCADA/WAMS system.

6. FUTURE RECOMMENDATIONS

At this moment, PowerFactory simulated results are not stored automatically in an excel file so that they can be stored in the database automatically. Results of PowerFactory should be stored in an Excel file automatically which can be achieved by installing Python for Windows Extensions PyWin32 package. This package includes Pythonwin extensions Win32 API and COM support. A python script will be written which imports PowerFactory simulated results into an Excel file ready to be imported into the database.

Relational databases Instead of using relational database, a graph database may be implemented to achieve fast results because during the last 8 months when this work was started until now Graph databases have improved and there is a lot more community support if you run into a problem. Also, it has been learned from this work that the relational databases impose difficulties when data is received at fast rates from PMU devices. Also, as the structure of dataset becomes more complex relational model becomes excessively loaded with large table joins which affects database performance.

Another improvement can be made in the translation of database schema to XML RDF format which until now could be only translated into XML XDF format. XML XDF format is not CIM compliant so it would be difficult to exchange data between the utilities. Converting database schema into XML RDF format is very time-consuming and error-prone if done by hand or manually. Therefore, a suitable translation software must be identified and utilized.

This DT DBMS test setup needs to be tested on a real power system only then, it can be said that this DBMS is adequate for DT needs.

Appendix I IEEE 9 Bus Datasets (Example Only)

IEEE 9 bus system represents a small power system consisting of 3 generators, 3 loads, 9 buses, 3 transformers and 6 transmission lines. It was first introduced in [76].

The 9-bus system has been divides into 6 sub geographical regions with each region having one Substation. Data in tables below is organized according to CIM compliant classes of the DB schema.

Table 27 Substation Data

Id	Name	Lat	Long	SRG
1	Sub1	30	40	0.001
2	Sub2	60	70	0.001
3	Sub3	50	60	0.001
4	Sub4	20	40	0.001
5	Sub5	60	30	0.001
6	Sub6	50	70	0.001

FK meaning foreign keys provide link between the tables.

Table 28 Nodes Data

Id	Name	Fk_Sub	Bus_Id	Status
1	N1	1	1	1
2	N2	1	1	1
3	N3	1	4	1
4	N4	1	4	1
5	N5	2	5	1
6	N6	2	5	1
7	N7	2	5	1
8	N8	3	7	1
9	N9	3	7	1
10	N10	3	2	1
11	N11	3	2	1
12	N12	3	7	1
13	N13	4	8	1
14	N14	4	8	1
15	N15	4	8	1
16	N16	5	9	1
17	N17	5	9	1
18	N18	5	3	1

19	N19	5	3	1
20	N20	5	9	1
21	N21	6	6	1
22	N22	6	6	1
23	N23	6	6	1
24	N24	1	4	1

Table 29 gives the switching devices data. Status tells the status of switching device and Nstatus represents the normal state of switching device i.e. whether it is normally closed or open etc.

Table 29 Switches Data

Id	Sub_Id	From_Node	To_Node	CKT_Id	Name	Type	Status	Nstatus	X	Rating1	Rating2	Rating3
1	1	1	2	1	CB1	CB	1	1	0.1	arb	arb	arb
2	3	11	12	1	CB2	CB	1	1	0.1	arb	arb	arb
3	6	19	20	1	CB3	CB	1	1	0.1	arb	arb	arb

Table 30 provides data of power system devices such as transformers, generators and transmission lines.

Table 30 Devices Data

Id	Name	CKT_Id	From_Bus	To_Bus	From_Node	To_Node	Subs_Id
1	Gen1	1	N/A	1	N/A	1	1
2	Gen2	1	N/A	2	N/A	2	3
3	Gen3	1	N/A	2	N/A	3	5
4	Load1	1	N/A	5	N/A	6	2
5	Load2	1	N/A	8	N/A	14	4
6	Load3	1	N/A	6	N/A	22	6
7	Transf1	1	1	4	2	3	1
8	Transf2	1	2	7	9	10	3
9	Transf3	1	3	9	17	18	5
10	Transm1	1	4	5	4	5	1
11	Transm2	1	5	7	7	8	2
12	Transm3	1	7	8	12	13	3
13	Transm4	1	8	9	15	16	4
14	Transm5	1	6	9	20	21	5
15	Transm6	1	4	6	23	24	6

Table 31 provides loads data. During load flow calculation, the loads of the Nine bus System have constant active and reactive power demand as described in table below. Rest of the tables data is self-explanatory.

Table 31 Load Data

Id	Name	Shortname	Active_Power	Reactive_Power	pu_power	pu_reactiveP	pu_voltage	Bus_num	Node_num
1	Load1	LD1	125	50	1.25	0.5	1	5	6
2	Load2	LD2	100	35	1	0.35	1.017	8	14
3	Load3	LD3	90	30	0.9	0.3	1.012	6	22

Table 32 Busbar Data

Id	Name	Shortname	Active_Power	Reactive_Power	pu_power	pu_reactiveP	pu_voltage
1	Bus bar 1	B1	N/A	N/A	N/A	N/A	1.04
2	Bus bar 2	B2	163	N/A	N/A	N/A	1.025
3	Bus bar 3	B3	85	N/A	N/A	N/A	1.025
4	Bus bar 4	B4	N/A	N/A	N/A	N/A	1.025
5	Bus bar 5	B5	125	50	1.25	0.5	1
6	Bus bar 6	B6	90	30	0.9	0.3	1.012
7	Bus bar 7	B7	N/A	N/A	N/A	N/A	1.027
8	Bus bar 8	B8	100	35	1	0.35	1.017
9	Bus bar 9	B9	N/A	N/A	N/A	N/A	1.033

Table 33 Generators Data

Shortname	GEN1	GEN2	GEN3
Power_MVA	512	270	125
Voltage_KV	24	18	15.5
Xd_pu	1.7	1.7	1.22
Xd_tr_pu	0.27	0.256	0.174
Xd_subtr_pu	0.2	0.185	0.134
Tdo_s	3.8	4.8	8.97
Tdo_subtr	0.01	0.01	0.033
Xq_pu	1.65	1.62	1.16
Xq_tr_pu	0.47	0.245	0.25
Xq_subtr_pu	0.2	0.185	0.134
Tqo_s	0.48	0.5	0.5
Tqo_subtr	0.0007	0.0007	0.07
Ra_pu	0.004	0.0016	0.004
Xl_pu	0.16	0.155	0.0078
S_1	0.09	0.125	0.1026
S_12	0.4	0.45	0.432
H_s	2.6312	4.1296	4.768
Busbar_Id	1	2	3

Table 34 Transformers Data

Id	Name	Shortname	Nominal_Pri_V_kV	Nominal_Sec_V_kV	R1_pu	L1_pu	R2_pu	L2_pu	Rm_pu	Lm_pu
1	Transformer1	TR1	24	230	1.00E-10	2.88E-02	1.00E-10	2.88E-02	5.00E+03	5.00E+03
2	Transformer2	TR2	18	230	1.00E-10	3.13E-02	1.00E-10	3.13E-02	5.00E+03	5.00E+03
3	Transformer3	TR3	15.5	230	1.00E-10	2.93E-02	1.00E-10	2.93E-02	5.00E+03	5.00E+00

Table 35 Transmission Lines Data

Id	Name	Shortname	Define	Length_km	R0	L0	C0	R1	L1	C1
1	Transmission Line1	TL1	4 to 5	89.93	5.88E-01	3.98E-03	5.89E-09	5.88E-02	1.33E-03	9.81E-09
2	Transmission Line2	TL2	5 to 7	170.338	9.94E-01	3.98E-03	5.41E-09	9.94E-02	1.33E-03	9.01E-09
3	Transmission Line3	TL3	7 to 8	76.176	5.90E-01	3.98E-03	5.89E-09	5.90E-02	1.33E-03	9.81E-09
4	Transmission Line4	TL4	8 to 9	106.646	5.90E-01	3.98E-03	5.90E-09	5.90E-02	1.33E-03	9.83E-09
5	Transmission Line5	TL5	6 to 9	179.86	1.15E+00	3.98E-03	5.99E-09	1.15E-01	1.33E-03	9.98E-09
6	Transmission Line6	TL6	4 to 6	97.336	9.24E-01	3.98E-03	4.88E-09	9.24E-02	1.33E-03	8.14E-09

Appendix II Transmission Line Parameters

Table 36 3 Phase High Voltage Transmission Line Parameters

Nominal Voltage	kV	110	380	380	750	750	1150	1150
R1	Ω/km	0,033	0,027	0,009	0,013	0,009	0,007	0,006
X1	Ω/km	0,266	0,260	0,242	0,276	0,272	0,291	0,290
C1	nF/km	13,65	14,08	14,99	13,13	13,29	12,68	12,72
R0	Ω/km	21	0,23	0,17	0,20	0,18	0,19	0,18
X0	Ω/km	1,46	1,32	1,34	1,31	1,31	1,19	1,19
C0	nF/km	5,48	6,47	6,46	6,08	6,11	5,60	5,63

7. BIBLIOGRAPHY

[1] Michael Grieves, “Origins of the Digital Twin Concept,” 2016, https://www.researchgate.net/publication/307509727_Origins_of_the_Digital_Twin_Concept (2019, Jan. 07).

[2] F. Fusco and U. Fischer, “Data Management System for Energy Analytics and its Application to Forecasting,” http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.741.2112&rep=rep1&type=pdf.

[3] C. Brosinsky, D. Westermann and R. Krebs, “Recent and prospective developments in power system control centers: Adapting the digital twin technology for application in power system control centers,” in *2018 IEEE International Energy Conference (ENERGYCON): 3-7 June 2018*, [Piscataway, NJ]: IEEE, 2018.

[4] F. F. Wu, K. Moslehi and A. Bose, “Power System Control Centers: Past, Present, and Future,” *Proc. IEEE*, vol. 93, no. 11, pp. 1890–1908, 2005.

[5] J. Taft and P. D. Martini, “Ultra Large-Scale Power System,” https://www.gridwiseac.org/pdfs/cisco_control_architecture_white_paper.pdf, 2012.

[6] S. K. Khaitan, J. D. McCalley and C. C. Liu, *Cyber Physical Systems Approach to Smart Electric Power Grid*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2015.

[7] C. Warren, *Can Artificial Intelligence Transform the Power System? - EPRI Journal.* Available: http://eprijournal.com/can-artificial-intelligence-transform-the-power-system/ (2019, May. 07).

[8] J. Morgan, *A Simple Explanation Of 'The Internet Of Things'.* Available: https://www.forbes.com/sites/jacobmorgan/2014/05/13/simple-explanation-internet-things-that-anyone-can-understand/#13a0db171d09 (2019, Jul. 02).

[9] R. Rosen, G. von Wichert, G. Lo and K. d. Bettenhausen, “About The Importance of Autonomy and Digital Twins for the Future of Manufacturing,” *IFAC-PapersOnLine*, vol. 48, no. 3, pp. 567–572, 2015.

[10] Pradeep Neelam, Venkatesan Natarajan, Vivek Diwanji, “Is Your Organization Ready to Embrace a Digital Twin?,” https://www.cognizant.com/whitepapers/is-your-organization-ready-to-embrace-a-digital-twin-codex3636.pdf, 2018.

[11] Luigi Vanfretti and Joe H. Chow, “Synchrophasor Data Applications for Wide-Area Systems: PMU/WAMS Invited Session 17th Power Systems Computation Conference,” 2011.

[12] D. S. Markovic, D. Zivkovic, I. Branovic, R. Popovic and D. Cvetkovic, “Smart power grid and cloud computing,” *Renewable and Sustainable Energy Reviews*, vol. 24, pp. 566–577, 2013.

[13] EPRI, *EPRI Grid Transformation Phase II: Seamless Geospatial Power Systems Model.* Available: https://www.epri.com/#/pages/product/3002006508/?lang=en-US (2019, May. 07).

[14] SIEMENS, *Electrical Digital Twin Brochure: Siemens Electrical Digital Twin.* Available: https://assets.new.siemens.com/siemens/assets/public.1535057572.66c9013092b493265e091c154a33f9dd38d36c20.electricaldigitaltwin-brochure-final-intl-version-singlepages-no.pdf (2019, May. 07).

[15] *Internet of things guide.* Available: https://www.i-scoop.eu/internet-of-things-guide/industrial-internet-things-iiot-saving-costs-innovation/digital-twins/ (2019, Apr. 29).

[16] N.J. Balu, M.G. Lauby, and P.S. Kundur, Eds, *Power system stability and control*: Mc Graw Hill Education (India) Private Limited, 1994.

[17] "SIEMENS Spectrum Power Information Model Manager," https://w3.usa.siemens.com/smartgrid/us/en/transmission-grid/products/ems-applications-components/ems-applications-components-tab/Documents/IMM_10-25-2010.pdf, 2010.

[18] J. Liu, X. Li, D. Liu, H. Liu and P. Mao, "Study on Data Management of Fundamental Model in Control Center for Smart Grid Operation," *IEEE Trans. Smart Grid*, vol. 2, no. 4, pp. 573–579, 2011.

[19] Oxford, *Oxford English Dictionary,* 2nd ed.: Oxford University Press 1992.

[20] R. Bond, "European Union Directive on the Legal Protection of Databases," *Computer Law & Security Review: The International Journal of Technology Law and Practice*, vol. 6, no. 13, pp. 401–412, 1997.

[21] P. Rob and C. Coronel, *Database systems: Design, implementation, and management,* 7th ed. Boston Mass.: Thomson/Course Technology, 2007.

[22] M. Amin, "The Case for Smart Grid," http://www.ourenergypolicy.org/wp-content/uploads/2015/06/20150604091846-Amin-Maaterials-PUF-1503.pdf, 2015.

[23] J. Wu and N. N. Schulz, *Overview of CIM-oriented database design and data exchanging in power system applications.: Proceedings / 37th North American Power Symposium*. Piscataway, NJ: IEEE Service Center, 2005.

[24] M. Kezunovic, G. Gurrala, A. Bose and P. Yemula, *The Next Generation Energy Management System (EMS) Design: Final Project Report*, 2013.

[25] *The common information model CIM: IEC 619681970 and 62325 - a practical introduction to the CIM*. Berlin: Springer, 2012.

[26] ENTSO-E, *Common Information Model.* Available: https://www.entsoe.eu/digital/cim/#common-information-model-cim-for-energy-markets (2019, May. 07).

[27] A. McMorran, "An Introduction to IEC 61970-301 & 61968-11: The Common Information Model," https://pdfs.semanticscholar.org/aef1/00f8d98d839a1b467195bd12e44447daada7.pdf?_ga=2.121224460.567316071.1557240601-390624284.1557240601, 2007.

[28] S.C. Savulescu, Ed, *Real-time stability assessment in modern power system control centers*. Piscataway, New Jersey, Hoboken, New Jersey, Piscataway, New Jersey: Wiley IEEE Press; IEEE Xplore, 2009.

[29] U. Häger, C. Rehtanz and N. Voropai, *Monitoring, Control and Protection of Interconnected Power Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.

[30] Dirk Westermann, "EES3 Lecture Notes," 2018.

[31] J. D. McCalley, "Energy Control Centers," http://home.engineering.iastate.edu/~jdm/ee457/ECC.pdf.

[32] J. Jayasamraj, "SCADA Communication & Protocols," http://srldc.in/var/NRC/SRLDC%20Fees%20Feb%202013/Trg%20Psti%20System%20Operator%20Trg/PSO%20PSTI%20%205-17%20Sep-2011%20Adisesha/ppts/Protocols/SCADA%20Communications%20and%20Protocols.pdf.

[33] E. S. Karapidakis, "Operation Management on Autonomous Power System," in *Lecture Notes in Electrical Engineering, Proceedings of the European Computing Conference*, N. Mastorakis, V. Mladenov, and V. T. Kontargyri, Eds, Boston, MA: Springer US, 2009, pp. 15–24.

[34] U. Kerin, C. Heyde, R. Krebs and E. Lerch, "Real-time dynamic security assessment of power grids," *Eur. Phys. J. Spec. Top,* vol. 223, no. 12, pp. 2503–2516, 2014.

[35] M. Scherer, "Frequency Control in the European Power System Considering the Organisational Structure and Division of Responsibilities," Doctoral Thesis, ETH Zürich, Zürich, 2016.

[36] European Commission Press Release, "Clean Energy for All Europeans," http://europa.eu/rapid/press-release_IP-19-1836_en.htm, 2019.

[37] F. Cleveland, "Data Acquisition and Control DAC," http://smartgrid.epri.com/UseCases/DataAcquisitionandControlDAC.pdf, 2004.

[38] *Wide Area Monitoring System (WAMS) | PMU OPC Server | Phasor data viewing - Substation Automation Systems & Solutions | Power System Substations | ABB.* Available: https://new.abb.com/substation-automation/systems/wide-area-monitoring-system (2019, Jun. 28).

[39] A. Pal, "PMU-Based Applications for Improved Monitoring and Protection of Power Systems," PhD, Virginia Tech, 2014.

[40] A. G. PHADKE and T. BI, "Phasor measurement units, WAMS, and their applications in protection and control of power systems," *J. Mod. Power Syst. Clean Energy*, vol. 6, no. 4, pp. 619–629, 2018.

[41] S. Haag and R. Anderl, "Digital twin – Proof of concept," *Manufacturing Letters*, vol. 15, pp. 64–66, 2018.

[42] Cornelius Baur, *Manufacturing's next act Industry 4.0.* Available: https://www.mckinsey.com/business-functions/operations/our-insights/manufacturings-next-act (2019, Jul. 04).

[43] Aaron Parrott, *Industry 4.0 and the digital twin technology | Deloitte Insights.* Available: https://www2.deloitte.com/insights/us/en/focus/industry-4-0/digital-twin-technology-smart-factory.html (2019, Jul. 04).

[44] *Digital Twin - We See | Emerson CA.* Available: https://www.emerson.com/en-ca/about-us/featured-stories/digital-twin (2019, May. 07).

[45] Ian Skerrat, *The Reality of Digital Twins for IoT.* Available: https://dzone.com/articles/the-reality-of-digital-twins-for-iot (2019, Apr. 29).

[46] G. Bansal, *Why is Digital Twin Technology Important in The Age of Big Data | Analytics Insight.* Available: https://www.analyticsinsight.net/digital-twin-technology-important-big-data-age/ (2019, May. 07).

[47] Fedem Technology, *WHY are Digital Twins useful? | Digital Twin | Fedem.* Available: http://www.fedem.com/digital-twin/why-are-digital-twins-useful/ (2019, May. 07).

[48] GE Power Digital Solutions, *GE Digital Twin: Analytic Engine for the Digital Power Plant.* Available: https://www.ge.com/digital/sites/default/files/download_assets/Digital-Twin-for-the-digital-power-plant-.pdf (2019, May. 07).

[49] Christoph Brosinsky, Dirk Westermann, Rainer Krebs, Tom Sennewald and Florian Sass, *AC/HVDC Control Center - Test and Demonstrator System*. Frankfurt am Main: VDE, 2017. Available: https://ieeexplore.ieee.org/xpl/mostRecentIssue.jsp?punumber=8278713.

[50] C. Vicknair, M. Macias, Z. Zhao and X. Nan, et al, "A comparison of a graph database and a relational database," in *ACM-SE 2010: Proceedings of the 48th annual Southeast Regional Conference, April 15-17, 2010, Oxford, MS, USA*, New York, N.Y.: ACM Press, 2010.

[51] B. M. Sasaki, *Graph Databases for Beginners: Why We Need NoSQL Databases.* Available: https://neo4j.com/blog/why-nosql-databases/ (2019, May. 07).

[52] Alexandra Martinez and Rodrigo Mora, "A Comparison between a Relational Database in the context of a Personalized Cancer Treatment Application," https://pdfs.semanticscholar.org/8e22/ef2b6869a008712625ffb98a1428220a8288.pdf?_ga=2.59822126.229762400.1562495211-1292940239.1561374905, 2016.

[53] Y. Chen, "Comparison of Graph Databases and Relational Databases When Handling Large-Scale Social Data," Master Thesis, University of Saskatchewan, 2016.

[54] B. Rund, *The Good, The Bad, and the Hype about Graph Databases for MDM | Transforming Data with Intelligence.* Available: https://tdwi.org/articles/2017/03/14/good-bad-and-hype-about-graph-databases-for-mdm.aspx (2019, May. 07).

[55] G. Ravikumar and S. A. Khaparde, "A Common Information Model Oriented Graph Database Framework for Power Systems," *IEEE Trans. Power Syst,* vol. 32, no. 4, pp. 2560–2569, 2017.

[56] M. Drake, *SQLite vs MySQL vs PostgreSQL: A Comparison Of Relational Database Management Systems | DigitalOcean.* Available: https://www.digitalocean.com/community/tutorials/sqlite-vs-mysql-vs-postgresql-a-comparison-of-relational-database-management-systems (2019, May. 07).

[57] S. Riggs, G. Ciolli, H. Krosing and G. Bartolini, *PostgreSQL 9 Administration Cookbook,* 2nd ed. Birmingham: Packt Publishing, 2015. Available: http://gbv.eblib.com/patron/FullRecord.aspx?p=2040610.

[58] S. S. Latif and W. Osborn, "Database system design for energy economization of smart home energy services," in *2017 the 5th IEEE International Conference on Smart Energy Grid Engineering (SEGE): August 14-17, 2017, UOIT, Oshawa, Canada*, Piscataway, NJ: IEEE, 2017.

[59] D. Povh, D. Retzmann and J. Rittiger, "Benefits of Simulation for Operation of Large Power Systems and System Interconnections: The 4th IERE General meeting & IERE central & Easterm Europe Forum," http://www.ptd.siemens.de/Simulation_IERE.pdf, 2004.

[60] X. Y. Wang, J. H. Zheng, S. Z. Zhu and S. D. Shen, "Parameter identification and database management system in power systems," in *PowerCon 2002: 2002 international conference on power system technology*: IEEE, 2002.

[61] M. K, "POWER SYSTEM DYNAMICS MODELING: 10th International Conference CONTROL OF POWER SYSTEMS 2012," https://www.researchgate.net/profile/Karel_Maslo/publication/272825196_POWER_SYSTEM_DYNAMICS_MODELING/links/54f0655a0cf2432ba65ad592/POWER-SYSTEM-DYNAMICS-MODELING.pdf.

[62] F. Milano, *Power System Modelling and Scripting*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2010.

[63] Jan Machowski, Janusz W.Bialek and James R.Bumby, *Power System Dynamics: Stability and Control*: John Wiley & Sons, Ltd, 2008.

[64] D. Oeding and B. R. Oswald, *Elektrische Kraftwerke und Netze*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2011.

[65] Shaarbafi and Karim, "Transformer Modelling Guide," https://www.aeso.ca/assets/linkfiles/4040.002-Rev02-Transformer-Modelling-Guide.pdf, 2014.

[66] *IEC 61970-302: 2018 - IEC standards - VDE VERLAG.* Available: https://www.vde-verlag.de/iec-normen/225522/iec-61970-302-2018.html (2019, May. 07).

[67] Y. Li, D. Yang, F. Liu, Y. Cao and C. Rehtanz, *Interconnected Power Systems*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2016.

[68] EPRI, *Using the Common Information Model for Network Analysis Data Management: A CIM Primer Series Guide. Product Id: 3002002587.* Available: https://www.epri.com/#/pages/product/3002002587/?lang=en-US (2019, May. 07).

[69] W3C, *RDF 1.1 XML Syntax: W3C Recommendation 25 February 2014.* Available: https://www.w3.org/TR/rdf-syntax-grammar/ (2019, May. 07).

[70] *The Python SQL Toolkit and Object Relational Mapper.* Available: https://www.sqlalchemy.org/ (2019, May. 07).

[71] *WSCC 9-Bus System: Illinois Center for a Smarter Electric Grid (ICSEG).* Available: https://icseg.iti.illinois.edu/wscc-9-bus-system/ (2019, May. 07).

[72] *PowerFactory OPC Guide.*

[73] MathWorks, “OPC Toolbox™ User's Guide,”

[74] *Query Tool — pgAdmin 4 4.6 documentation.* Available: https://www.pgadmin.org/docs/pgadmin4/dev/query_tool.html (2019, May. 16).

[75] *PostgreSQL: Documentation: 9.4: Constraints.* Available: https://www.postgresql.org/docs/9.4/ddl-constraints.html (2019, May. 16).

[76] P. M. Anderson and A. A. Fouad, *Power system control and stability*. Ames, Iowa: Iowa State University Press, 1977.