Vaasan yliopisto
UNIVERSITY OF VAASA

Österbottens förbund
Pohjanmaan liitto

Regional Council
of Ostrobothnia

Technical Report TR 5.2

Information Security in Electricity Generation and Distribution Data Storage

Vaasa Energy Business Innovation Center
VEBIC

Smart Energy Systems Research Platform
SESP

Bahaa Eltahawy

Table of contents

Abbreviations

AAA	Authentication, Authorization, and Access Control (also Accounting)
ACS	Access Control System
CAN	Campus Area Network
CCTV	Closed Circuit TV
CPU	Central Processing Unit
DA	Distributed Automation
DAS	Direct Attached Storage
DCS	Distributed Control System
DDoS	Distributed Denial of Service
DER	Distributed Energy Resources
DMZ	Demilitarized Zone
DoS	Denial of Service
ESS	Electronics Security System
EU	European Union
FC	Fiber Channel
HTTPS	Hypertext Transfer Protocol-Secure
ICS	Industrial Control Systems
ICT	Information and Communication Technologies
IDPS	Intrusion Detection and Prevention System
IDS	Intrusion Detection System
IEA	International Energy Agency
IEC	International Electrotechnical Commission
IoT	Internet of Things
IP	Internet Protocol
IPS	Intrusion Prevention System
IPsec	Internet Protocol Security Protocol
IS	Information System
IT	Information Technology
LAN	Local Area Network

MAC	Media Access Control
NAP	Network Access Protection
NAS	Network Attached Storage
NAT	Network Access Translation
NIST	National Institute of Standards and Technology
OCED	Organization of Economic Cooperation and Development
OSI	Open System Interconnection Model
OT	Operation Technology
PMU	Phasor Measurement Unit
PPP	Point to Point Protocol
RADIUS	Remote Authentication Dial-In User Service Protocol
RAID	Redundant Array of Independent Disks
RAM	Random Access Memory
RoI	Return on Investment
SAN	Storage Area Network
SAS	Substation Automation System
SCADA	Supervisory Control and Data Acquisition
SFTP	Secure File Transfer Protocol
SG	Smart Grid
SM	Smart Meter
SSL	Secure Socket Layer Protocol
TACACS	Terminal Access Controller Access Control System Protocol
TLSTransport	Layer Security Protocol
VPN	Virtual Private Network
WAN	Wide Area Network

Executive Summary

With the present rapid developments the energy sector is witnessing, and with the transition from the yet dominant older power grid system to the more modernized, digitalized, and smarter grid systems, a new era is in close. An era that brings many opportunities into place, translated into optimization of use and integration of the different energy resources, as well as the ability for a better control of utilities and power systems by means of utilizing ICT technologies. In spite of this, it does not come without own flaws; the transition comes with a whole can of threats brought up by the computerized world. These are seen as with the window open to the cyber world and with the potential threats coming through it, benefits and added values could be overcome. Furthermore, in a worst case scenario critical and dangerous situations could be resulted.

In the same manner, earlier electricity utilities used to adapt and operate based on a centralized model that holds all control, operations, and data in one place. However, with the objectives changing, and with the introduced technologies supporting the new aims, the system is also transitioning to follow a more distributed decentralized model. With this introduced decentralization, the newer model requires distributed control, distributed data storage, and distributed processing, thus to be able to function properly and provide for the intended benefits. However, this remodeling comes similarly with its own threats and risks, as seen in energy-targeting attacks.

This report, Report 5.2, Information Security in Electricity Generation and Distribution Data Storing, intends to carefully study and address the pre-given issues. In this report, the main information security and/or cyber security challenges that face electricity generation in smart grids are discussed in favor of understanding the current situation and finding proper solution. As well, the issue of distributed data storage is well covered.

Key findings of this report include:

1. **An increase in the attack surface** due to the nature of the decentralized model. This in turn can result in increased number of intermittency and disruption.
2. **The need for distributed protected zones** that feature high level security, thus to provide protection for devices and sensors lacking it.
3. **SCADA and ICS upgrade requirement,** as many of these systems were not designed with the current scenarios taken into consideration, and thus their security measures are outdated.
4. **The need for parallel nonintegrated connections** to provide for resilience.
5. **Layered data storage approach** to facilitate data exchange, and to support secure operations.

1 Introduction to Power Grid Systems

The main goal of power grid systems is to provide consumers with their needs of electricity; this has been the goal for over a century, and systems were built upon that. To achieve this, systems were built simple without much of complexity. Typical power grid systems comprised of: power plants to generate electricity, utilities that manage the whole system and add functionalities, transmission links to carry electricity to the distribution stations, and finally distribution systems that transport electricity to consumers. The system could also include other sub-distribution systems, thus to form a higher level of transmission and distribution for consumers with special demands. In Figure 1 below, a typical power grid system is shown.

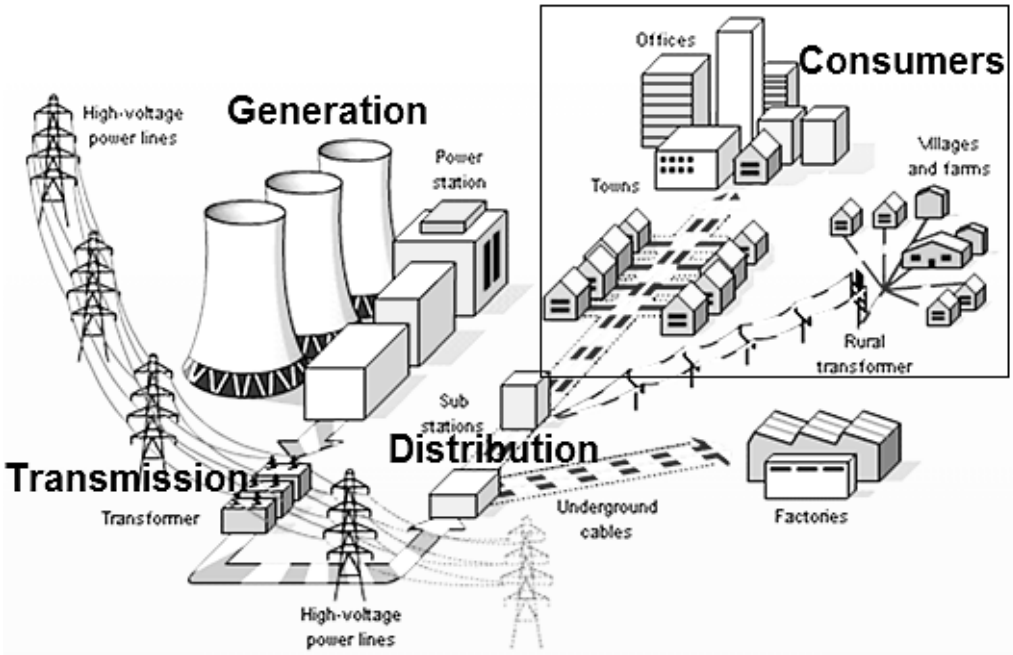


Figure 1: Typical power grid system.

Following the developments and advancements in the energy sector, the main goal remained clear, but rather more goals got also considered. These would focus on optimization of use, routing of electricity upon need, reduction of Carbon emissions, integrating renewable resources on large scale to replace fossil, and finally integrating energy produced by consumers to the main network. These goals combined with the other technological advancements were the spark for what is known now as smart grid systems. Smart Grid Systems are these ones that make use of the Information and Communication Technologies ICT for better delivery of energy. Unlike the older power grid systems that depend mostly on manpower for operations and configurations, smart grid systems depend on Industrial Control Systems ICS, and Supervisory Control and Data Acquisition SCADA systems to provide for near fully-automated operations. In Figure 2, the links of power and communication of a typical smart grid are shown.

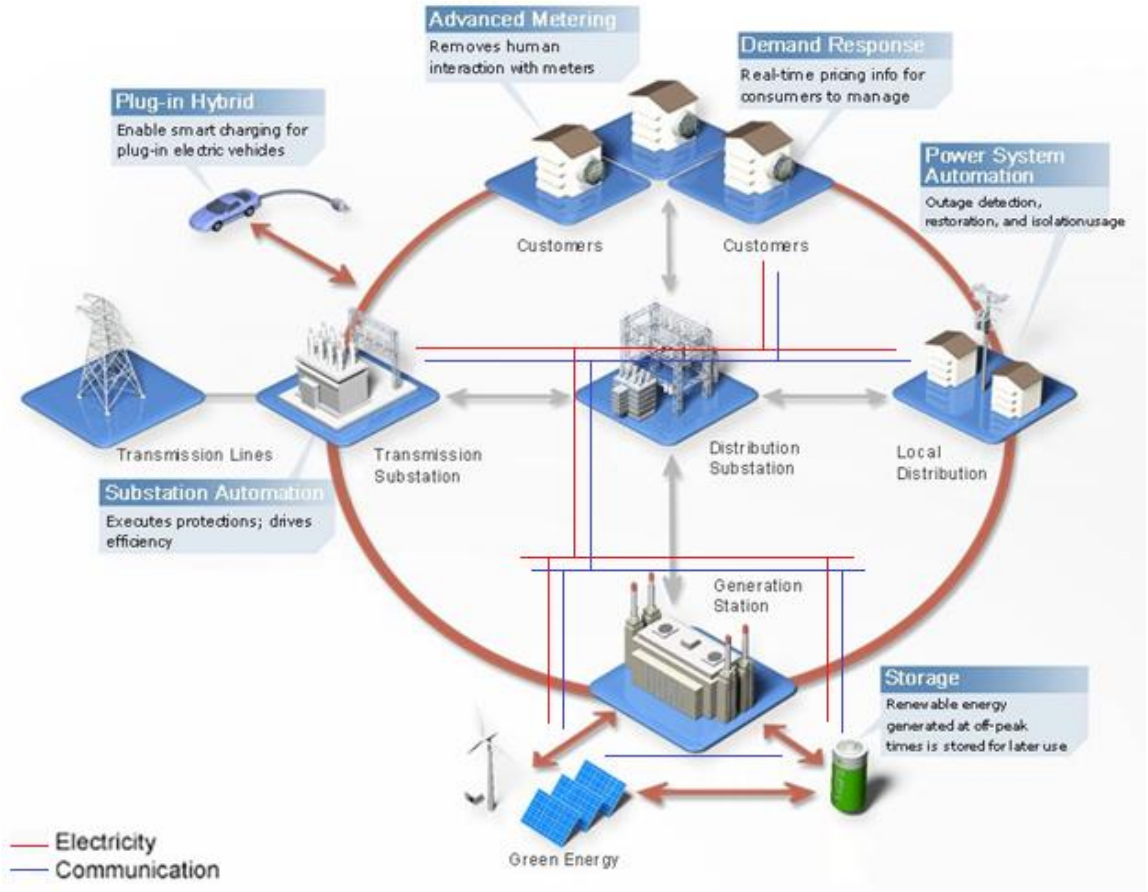


Figure 2: Smart Grid System.

At present, newer grids are being built following the theme of smart grid systems; however, older systems are still in place fully running. This could be seen as a proper utilization of existing and functioning deployments. On the contrary, due to the big differences between the older and newer deployments, older networks could affect efficiency and bring vulnerabilities and threats to the smart grid systems. Older power grids mainly lacked or had a minimum of intuitive control, so the first challenge seen is to integrate older grids with control systems. Secondly, these systems were totally independent of any communication systems, so one of the main challenges is about integrating and enabling ICT. Thirdly and more critically, in many cases the deployed infrastructure does not allow remedies as upgrades, e.g. costly, or at remote areas.

In the following sections, the different parts of the grid systems will be studied closely, in favor of understanding them, and revealing the associated vulnerabilities.

2 Components of Electricity Generation

The first and most fundamental component in the grid system is its generation facilities. Here, all electricity supply is generated, to be distributed later to end-users. In its older profile, these facilities were following utilities and service providers for all responsibilities, policies, technical details, and configurations. However, due to the changes happening to the structure of the grid, this is not the case anymore. In its current architecture, generation comprises of:

1. Generation Stations
2. Energy Storage
3. Electricity Utilities
4. End-users' Generation Facilities
5. National Cooperation

2.1. Generation Stations

According to the International Energy Agency IEA, the Organization of Economic Cooperation and Development OCED countries' shares of electricity generation by source as in 2016 is mainly combustible, then nuclear, then hydro and other resources. Shares are shown in Figure 3 below.

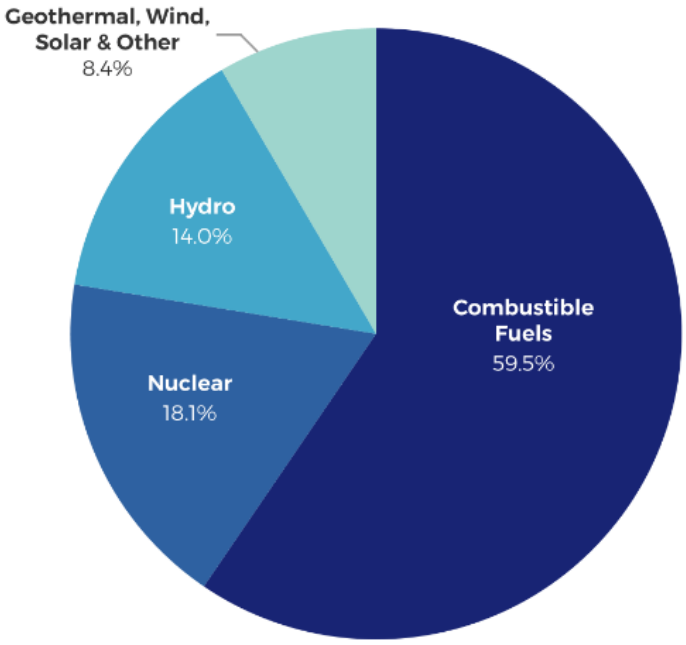


Figure 3: Electricity generation shares, as in 2016

Generation stations are the main component in the grid, and it is clear that the older sources as combustible fuel sources still play the most dominant role for generation. On the other hand, nuclear power, hydroelectric plants, and other renewables, e.g. solar and wind power, are getting more popularity seen as increased shared. This follows the European Union EU goals for decarbonization, and creation for an environmentally friendly energy sector. With this diversity of resources taken into account, also with other factors as: resources availability, population, transportation, pollution levels, noises, operation and maintenance costs, it is clear that the generation stations would be geographically located at different locations. However, with them distributed, still generation stations are well connected by means of transmission links.

2.2. Energy Storage

Energy storage facility plays an important role in supporting the electricity generation cycle. Earlier, with the dependence mainly on fossils and combustibles to provide for electricity, it was easier to control the supply according to the needs, and thus reducing generation once not needed. However, with the current integration of energy resources, and with the fact that some resources cannot be controlled, e.g. wind and solar power, a proper method for storing excess energy production is required. This will prevent excess generation from rendering wasted. Luckily, at present there are many methods that are used for this purpose, and they can store energy at its generation location or at another site by other means. Many technologies exist to serve for this purpose, e.g. batteries, super capacitors, hydrometrics storage, etc… However, the technology to be utilized depends solely on the location of the generation station, costs, efficiency required, and the amount of excess energy that would need storage.

2.3. Electricity Utility

Electricity utility is the corporation that directly interfaces between the electricity service and consumers, holding full responsibility over the electricity service from early-stage generation to its final distribution. Here, many functions are included, as: transmission, communication, and control of the generation stations and storage facilities. Previously, utility companies used to part of the public sector, and/or services. This made it easier for utility companies to create their own rules, to set own regulations, and to establish cooperation with other organizations. Currently, the structure of utility companies has changed, as many of them became either privately owned, or public profitable companies. This nonetheless increases the challenges utility companies face, since alternatives became available. Utilities accordingly are required to include other services; provide for reliability, efficiency, and security measures; and follow national or international regulations according to their size and growth.

2.4. Consumers' Generation Facilities

Electricity supply used to be unidirectional from utility to consumers according to their needs; however, with current deployments this has changed. Earlier, customers unable to connect to the supply network used to implement their own electricity generation solutions. Later on, with transmission links spanning for more coverage, this resulted in an excess of electricity. In an advanced stage, customers could integrate their electricity generation facilities to the main grid to provide the main grid with this excess of electricity. With this integration, many benefits could be shown. For instance, this helped to reduce the load on the supply network since the need is less; excess generation could be routed to where electricity is in need; costs and operations overhead are reduced; and on top it promotes for green/environmental-friendly electricity. On the other hand, the model is non-uniform as it has many structures, and is associated with own complexities, as seen later on.

2.5. National Cooperation

The difference in the generation capacity and needs' gaps across the different neighborhood is clear, and it could be solved by means of cross-national cooperation. Cross-national cooperation aims to fill the gaps in need and supply by expanding networks across borders, thus to serve the excess supply of one country to another. This model features acquiring inexpensive energy supply, the promotion and continuous use of renewable resources, and on top reducing/eliminating the need for energy storage since excess electricity could be directly served somewhere else than storing it. Again, though the many benefits it could offer, the model needs a careful consideration as it is associated with many complexities.

Distributed Systems

As discussed above, smart grid systems have many components and players involved, that do not need to be collocated. With different goals taken into account, smart grid systems favor for distributed systems with distributed control, and distributed data acquisition nodes. This brings to the grid many advantages, e.g. optimization of electricity generation; optimization of distribution and transmission; generation and load routing from a location to another; and the involvement and control of customers generated electricity. With this infrastructure distributed, more services are added, along with achieving for better sustainability, efficiency, and flexibility.

Sustainability is one of the main key values added with the introduction of distributed infrastructure. Here, systems can be prolonged and run smoothly for longer times, as the need for them to be up and running around the hour is no longer required. This would save resources,

reduce operation costs, provide for cheaper services, and allow for integration and development of coexisting solutions. Efficiency is one of the targeted goals of smart grid systems, as the older deployments suffered from inadequate generation, since utilization could not meet with the needs. Previously, this was addressed by either having less generation, or excess of electricity, which in either way results with inefficiency and waste or resources. With distributed systems in place currently, efficiency is met at its peak since generation could match with demands. Finally flexibility is well achieved with distributed systems, since the system could interface and integrate with many services and resources upon need. Additionally, the system includes many alternatives, thus does not suffer from shortage or scarcity.

On the other hand, despite the proven advantages of implementing the grid in a distributed fashion, the move from centralized systems brings many complexities to the system, as follows:

1. The involvement of many players and the arisen conflict with policies and regulations.
2. The use of many nodes and the complexities associated with design, management, and maintenance of these nodes.
3. The security paradox that arises with the increase of the attack surface by means of increasing the entry points to the system.
4. Data generation, data transportation, processing, and the burden due to the increase of resources.
5. Data storage and the proliferation of data.
6. Etc…

In Figure 4, the older system and the distributed system are compared. It is clear the nodes of distribution on the right side, and the complexity they bring from the design point of view, not to mention yet the associated data issues.

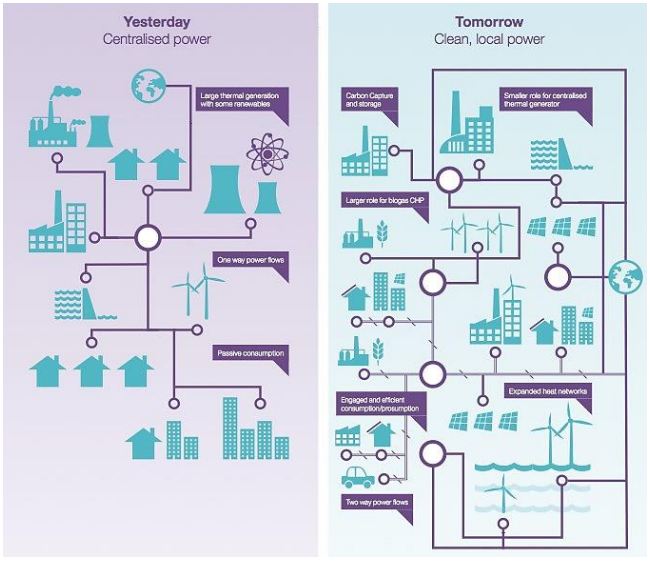


Figure 4: Comparison of centralized and decentralized (distributed) electricity systems

3 Smart Grids and Risks

As shown earlier in Figure 2, smart grids comprise links for electricity transmission/distribution that go in parallel with communication links that carry data necessary for operation and control. Earlier, control was done in a mechanical manner, thus networks could not be compromised without physical access to the utilities and control facilities. At present, control is rather automated and handled via ICS. Industrial Control System ICS is an integrated system that encompasses hardware, software, communication technologies, and computational algorithms to provide for efficient control. ICS mainly comprises of two subsystems, namely SCADA, and Distributed Control System DCS. These systems provide for a supervised control rather than a fully automated one, in addition to distributed control loops among the system. ICS systems introduce wide range of functions and configurations, and among the benefits they offer is remote configuration of remote utilities. This however puts the ICS in exposure to the cyber world, being vulnerable to all of its excising or coming threats. This happens due to compromising the communication networks, and thus careful consideration should be given to networks and network security. In ***Report 5.1, Cabled and Wireless Communication Security in Electricity Generation and Distribution Systems***, the issue of network security was discussed thoroughly.

Another important concern to pay attention to is the deployment of ICS native systems with the older ones lacking communication and control functions. This shortage is overcome by means of implementing mediator systems that are capable of handling the transition, by handling communication signals, and providing control in a way that the system can deal with. Such systems depend solely on Internet Protocol IP and Medium Access Control MAC functions to perform their tasks. However, these systems mainly are designed to provide for functionality and operability rather than security. As well, though the fact that the mediator systems integrate many access points to the system, they are not security native by design. This in turn introduces a potential risk to the whole grid system, since an easily manipulated and compromised system can result in compromising the whole grid. A third factor that requires careful consideration here is the change and configuration management. Since networks might need to get reconfigured at certain instances for adding functionalities, fixing existing issues, or for general maintenance criteria, a tight tracing and/or logging system should be in place to hold all necessary information for the change and configuration.

It is a matter of fact that the involvement of Information Technology IT to the field of control and Operation Technology OT is not going to be seamless without its own flaws. The main challenges here are: the threats that the cyber world possess; and how to balance for a functioning yet robust system. These are discussed in the subsequent sections.

3.1. Sources of Threats and Vulnerabilities

Firstly, it is important to learn about sources of threats, and the associated vulnerabilities and risks they bring. Typically, threats can be framed into three categories, namely, natural, human, and technology. At a higher and rarer level, natural threats are *"Naturally occurring events that may cause disruptions to an organization or have other negative effect on people or the environment."* Though this type of threats happening infrequently, countermeasures should be ready to face them upon incidents. In the second category of threats comes human threats, which can be defined as the "*the caused disruptions in operations or breach of security controls resulting from intentional or unintentional human actions".* In this category, the human factor is the main cause that can bring the system down, either by random incidents, or by performing actions with malicious intent, e.g. data theft or preventing legitimate services. Finally, technology threats are these ones that are man-made, and target exploitation of systems in place. In this context, systems can include all sorts of equipment and supporting networks required for operation.

From another perspective, threats could be classified into two main categories, namely, physical threats, and information threats. In the former category, physical threats concern all sorts of damage or harm caused by physical access to the facility, and/or causes physical damage to the service. In contrast, information threats mainly concern information systems and manipulation of data without the need for physical access or physical damage. However, information threats have the tendency to render systems useless without causing any damage, by simply preventing them from operation, or by means of malfunction by manipulation and modification of services.

The given introductory make it clear that the threats the energy sector mostly face are related to human and technology, also falling within physical and information threats. In more details, smart grid systems' components are susceptible to human and technology threats due to their dependence on the human factor for operation, and the heavy involvement of technology for control and automation. Threats that can occur at these levels result mainly from the lack of security measures in the system, and that the systems/services would not perform as designed upon incidents of disruption. It mainly depends on the organization and the criticality of the given services, but tight security measures are required to provide for robustness, and a backup resilience measures are required at these levels to provide for availability measures.

In this report, we will mainly cover these threats associated with electricity generation. These are discussed in the next section.

3.2. Risks

Many risks exist in the current electricity generation system due to the changes in its structure as explained earlier. The most risk concerns here are:

1. Physical Security
2. Distributed systems and the attack surface
3. Data Exchange
4. Internet of Things IoT
5. Legacy systems and the integration of OT
6. Regulations across different parties
7. International standards.

3.2.1. Physical security

Though most current systems are digitalized, physical security remains the foundation of security practices that need to be assured. However, physical security is not well cared of currently. Physical security concerns the practices that keep devices, systems, and communication links secure in their physical form from any direct unauthorized access, or modification of the system values by intruders. These forms of actions when being unwatched, they can render damage or deviation of functions and/or services from their norms. Not only limited to these, but physical security can in some cases lead to more dangerous attacks by means of manipulating information systems, and accessing information stored within, for a later usage. It is clear that with the increase in the number of utilities, generation facilities, and the associated increase in the number of links between them, the risk also increases. And though it is significantly challenging how to provide the means to protect all these devices from physical access, it is not negotiable.

3.2.2. Distributed systems and attack surface

As mentioned earlier, systems are no more centralized, but rather distributed. Despite the benefits this change offers, security tends to be the main challenge. Firstly, Distributed systems can be located at many locations that differ in their set up and the implemented security measures, thus not easy to maintain the desired level of protection. Secondly, systems might include older systems that are up and well-functioning, while lacking security measures. These legacy systems tend to be one of the main cyber security challenges that face smart grids, and generally speaking it affects all modern systems across all sectors. The main solution for this shortage is by deploying separate security solutions that are separate and isolated from the

modern grid. However, these solutions do not provide for full integration and security measures compared to the native ones, which makes them still vulnerable for advanced attack scenarios. The third issue arises from the nature of distributed systems, since their networks span across long distances, which requires measures for surveillance, protection, and maintenance. Finally, distributed systems in many cases do not have their control systems at the same location; rather control is done remotely. This latter issue can possess a threat of manipulating control links and taking over functioning resources.

On a larger scale, the main threat that turns out from transforming centralized systems into distributed ones is the attack surface. Now with the system distributed over different sites, the number of devices almost folds with the number of sites, as well as inter/intra -communication links in between. This set up allows for many entry points that could be used to gain unauthorized access, or at least intercepting links between the different sites.

3.3.3. Data Exchange

The newer smart grid system heavily depends on data exchange to perform its functions in contrast to the older system that had less or no intelligence required for control; which would arise many concerns. Firstly, the need for separate data links, or the adoption of data over power lines, and/or other media. Secondly, the inclusion of legacy systems that were not designed to provide data, and thus the need for mediator devices to provide for such functionality. Thirdly, integrity measures are required to make sure data is not altered on its way to control facilities. Finally, measures to protect against manipulation of measurement devices at end users providing supply to the network, as well as smart meters to provide for consistent data.

Data exchange concerns extend beyond the typical information security practices, as large number of devices are incapable of processing power for advanced functions, e.g. remote sensors, and Internet of Things IoT devices. These will need careful consideration to provide for the required quality.

3.3.4. Internet of Things and Remote Sensors

Internet of Things IoT devices and remote sensors could be installed remotely to collect data and send it back to the central node for further processing or mining. These devices can also perform minor pre-set functions upon certain conditions triggered. Typically IoT devices and remote sensors feature the least of processing power, limited functionalities, and have minimum to none of security features. As a result, these devices can run for few years without the need for power supply or battery replacement. On the other hand, due to their limitations, these devices

possess increased risks to the system, as they introduce weak links, increase the chance of device manipulation, and with no data protection arises the risk of data modification. Typically, security in these systems would depend on central modules that interface with sensor devices, perform some level of data protection, and finally send data in their behalf.

3.3.5. Legacy Systems and the Integration of OT

Legacy systems have their own effect in increasing systems' complexity. These systems are still capable of performing; however, they are not designed for the current demands of the advanced functionalities and services of the nowadays' demands. From one perspective, these systems are assets that are worth utilization; from another perspective, they are threats that can bring the whole system down. Currently, with the implementation of OT, these systems come to be a real challenge, as they cannot integrate with OT systems natively without being interfaced with other systems that can handle control on their behalf. Even so, since the mediator systems try to combine both legacy and current measures, they themselves tend to have own security flaws. On a higher complexity level, legacy links are the most concern when comes to legacy systems. These links are hard to replace, as they span for long distances across the nation, and the fact that replacing them would be too costly, time consuming, and with slow Return on Investment RoI.

3.3.6. Regulations across Different Parties

Now with the cooperative generation, electricity is generated and managed by multiple parties at the same time. Earlier, utility companies were public organizations; however, now many private utilities as long as end-customers are all in the process equally. Regulations play a vital role here, either from the legal considerations, or the technical side. Different parties typically have different interests, which in many cases might result in conflict of interests. For instance: w a public utility would request to collect data on consumption for generation optimization; a private customer would prefer to keep data shared to minimal; the commercial entity would want to tune costs by means of analyzing data about generation and customers. In many situations, regulations and agreements therefore would need to get revised, thus to match with services' requirements, and to fill in their existing gaps. Another issue to consider here is the cross-industry regulations in which regulations span across many sectors. In this situation, an agreement/standard is applied to multiple sectors as is, while it cannot fit with their individual requirements or level they seek. This latter could result in some compromise in security, or affect the given services.

3.3.7. International Standards

Related to the previous section: standards are made typically as zone specific according to the regional interests, and the local forms of authority. With the current scenario of spanning electricity generation across multiple zones or countries, there might be a conflict between the applied standards. For instance, within the EU, some standards and agreements target the Nordic countries, while others target Central and Western Europe. Conflicts in standards, goals, and regulations thus would be seen an obstacle for seamless smooth integration, or in a worse case might prevent cooperation.

3.3 Mitigation Scheme

3.3.1. Physical security

Firstly, safety measures should be the foundation for protection, and are followed prior to security implementation. Here, systems would include the basic protection as fire suppression, shielding, temperature control systems, and isolation of servers and control rooms. Secondly, the fundamental practices of physical security are applicable at this stage. In Report 5.1, these were addressed. Briefly, the Electronic Security Systems ESS concerns all means for providing physical security, including systems to record events, detect activities, prevent access, and send alarms upon preset conditions being triggered. Typically, systems that can provide for these functions include but not limited to: fences, Scanning Systems SS, Access Control Systems ACS, Intrusion Detection and Prevention Systems IDPS, Surveillance Systems SS with Closed-Circuit TV CCTV, and Intercommunications Systems IS.

3.3.2. Distributed systems

Though distributed systems can possess risks themselves, they also can provide for an adequate level of protection. Distributed systems as a fact provide for an easy access integrable security scheme that matches better with the local services, facilities, and utilities. The most benefits seen from implementing distributed systems is the creation of protected and isolated zones. Distributing data to different locations can also provide for data retrieval means at instances of attack or data damage at a certain location. Additionally, with implementing some of the higher schemes as distributed RAID Network Storage, data could be always available to all locations despite failure at a one or more sites (depending on the RAID scheme). Furthermore, decreasing the burden on a specific site is one of the main advantages of distributed systems, as

well as the ability to implement different homogenous systems at different locations without any drawbacks or tradeoffs.

3.3.3. Data exchange

Since not every device is capable of handling security functions, other measures should be in place to protect devices with limitations. Isolation is one of the recommended solutions that can be implemented here. In this criteria, remote devices will be protected by some means of Demilitarized Zones DMZ or firewalls, thus acting as completely hidden from the public network. The other criteria is by implementation of centralized zones, in which a central native security devices would only interface to the other devices, and thus all communication will be bypassed securely. Solutions that can fit here typically include implementation of proxy servers, and network traffic encryption mechanisms by means of Virtual Private Network VPN solutions.

3.3.4. Parallel links, Resilience, and Fallback measures

Parallel links are always a good solution to implement for various reasons. At this point, links will provide for an adequate level of availability, but mostly they will provide for a higher level of security. Upon failure or incident of disruption to one of the links, other links could still function properly and keep services running. For availability, resilience and fallback services/measures should be also considered in the early design phase, not as a response *(revise risk assessment measures in earlier reports)*. Resilience cares about how to recover failure and get services back in shape, with minimum tradeoff, and within the shortest period. Consequently, fallback measures should be implemented to provide for resilience. Here, replications of critical services will be isolated, thus to resist attacks or incidents of disruption. Then, upon certain criteria triggered, standby services would replace the infected ones until they are fully functioning again, which is when they switch back to standby mode. Standby services, fallback measures and resilience however do not need to provide for full functionality of the running services, but rather for the acceptable level for operations and service maintenance.

3.3.5. Reduction of Attack Surface

Reduction of the attack surface is critical for maintaining the network well secured. As a fact, the less devices seen by the network, the less the issues and security breaches that could result. In practice, this could be achieved by means of grouping of services, an adequate level of isolation, and creation of secure zones. Implementations to serve this include: proxy servers, data concentrators, and other Network Address Translation NAT solutions. Another solution is by

utilizing other communication protocols, media, and schemes for the internal communication than the external ones. With these configurations, the number of active addresses visible will be less, and thus higher protection scheme.

3.3.6. Broader regulations

The current regulations and laws need to be revised as they were not meant to address the independency of customers, the variety of sources, and the proliferation of information flows that might result. Privacy is one of the major concerns that need considered, as well as the regulations of customers contributing to the grid. On a higher level, cross-border cooperation needs further and more detailed regulation as it is of many benefits, but yet not implemented as it should.

Information Security in smart grid systems

Like other digital systems, information security plays a vital role here to protect resources and networks from harm, either intentional, or accidental. Here, the foundational information security practices will be still followed, i.e. authentication, confidentiality, and availability. Under this umbrella, other practices also exist, e.g. authorization and access control; utility; possession; etc.. As a matter of fact, these practices solely depend on encryption and logging mechanisms, to protect data, and to assure its availability to the right personnel at the right time without any sort of manipulation. However, in general many issues would arise at this point, including:

1. The capability of remote or sensor devices to handle the encryption algorithms.
2. The increased burden on the network due to continuous encryption.
3. Proliferation of logs records due to the increased number of devices.
4. The integration of legacy devices to incorporate security functions.

Some specific areas that suffer from these issues include but not limited to the following:

1. Distributed Energy Resources DER which include generators and energy devices might be located anywhere, including the customer’s side. These devices might not be securely configured or operated under the right conditions. Moreover, DERs consequently could be manipulated to provide for wrong data, or in a larger scale they could be used to initiate a Distributed Denial of Service DDoS attack.
2. Distributed Automation DA refers to that fact that automation, monitoring and control devices are no longer located at a central location. This will provide for remote capabilities for example, or to perform some level of protection by delegating some of the operations to local devices than centralized one. On the other hand this comes with many vulnerabilities; for instance, some of the automation devices might reside at customers or substations who do not follow some regulations, also it is not guaranteed the level of

protection communication links would have. Another critical scenario resembles the one mentioned before with DER, which is by launching a DDoS attack by manipulation of the DA control.

3. Data security practices for utilization of data, as data requires an adequate level of protection, as well as measures to facilitate extracting valuable information for optimization of use. This case is currently regulated by the European General Data Protection Regulations GDPR which states how to handle data, thus to keep the right for users, and to protect their privacy. Data security as well concerns data in all of its three forms, i.e. at rest, in transit, and in use.
4. Communication protocols that are used to set configurations and secure operations might not work as desired, since there are many sub-systems included, also the conflict between devices' capabilities according to manufacturers and their legacy. For instance, SCADA, Substation Automation Systems SAS, DER, DA, Smart Meter SM, Phasor Measurement Unit PMU, etc…
5. The conflict that arises from installations of analogue and digital signals side by side, as the analogue ones fail to meet the security requirements of their deployments. For instance, latency mismatching, and interoperability criteria.
6. Some standards' messages do not include security feature by default, for instance the case with IEC 61850. These messages can then be spoofed, modified, or used as a base for replay, injection, or generation attacks.
7. The use of free licensed bands in some operations brings the networks jamming and DoS attacks.
8. Coordinated attacks, as the system is generally designed for certain functionality than security features. Thus, with a series of attacks in a coordinated manner, the system might fall to respond, since availability is critical.

The aforementioned issues so far emphasize on the general picture of the distributed and smart grid systems. Remedies for these issues would include: Firewalls, IDS, IPS, Isolation, and Fragmentation; these have been extensively covered in report 5.1.

4 Distributed Data Storage

With the abovementioned deployments, it is clear that data no more resides at a single place, rather it is distributed among several locations by design, as seen in Figure 5. With this taken into consideration, data protection and data storage are of vital role. As declared earlier, decentralization of data provides for immediate and more feasible results with critical decisions required. However, with the need to transfer data from a location to another, data is exposed to many threats. It is important here to understand that data can take one of three forms, namely: Data at rest, Data in use, and Data in Motion. In brief, data at rest refers to the data in its inactive state, being stored in a form of digital media. Data at rest would be retrieved from storage media for further actions at some point when needed. Data in transit is when data is transferred from a place to another for processing, or for storage. Data in transit in reality is all data flowing within a network. Finally, data in use is when data is being processed and operated, and in this case it resides in Random Access Memory RAM modules than storage media itself. Data in use typically would last for short periods, then being transferred to storage media again for storage or for further transfer.

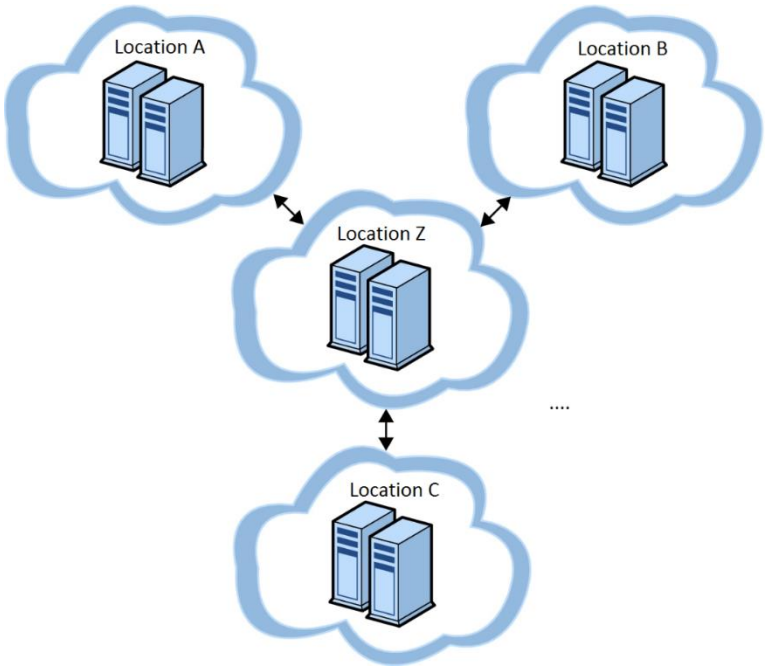


Figure 5: Data storage distributed

Data in motion/transit typically follows all the procedures and practices of network security. This merely focuses on integrity and encryption mechanisms to keep data intact without any modifications or confidentiality breaches. Many protocols are used at this level, for instance: SSL, and IPsec. Data at rest can be well protected by means of encryption mechanisms as well. Here, encryption mechanisms can be different from the ones used with network as there will be

enough processing power, and thus high-level encryption mechanisms will not cause the same burden as with networked data. Finally, data in use protection comes to be the most critical and challenging, as data is continuously changing at this stage, so protection mechanisms should avoid adding any burden or causing any functionality tradeoffs.

4.1. IBM Data Protection Practices

IBM has their own implementation to keep data protected across all stages. In their deployment, nearly all aspects of data protection methods are covered, as shown in Figure 6 below.

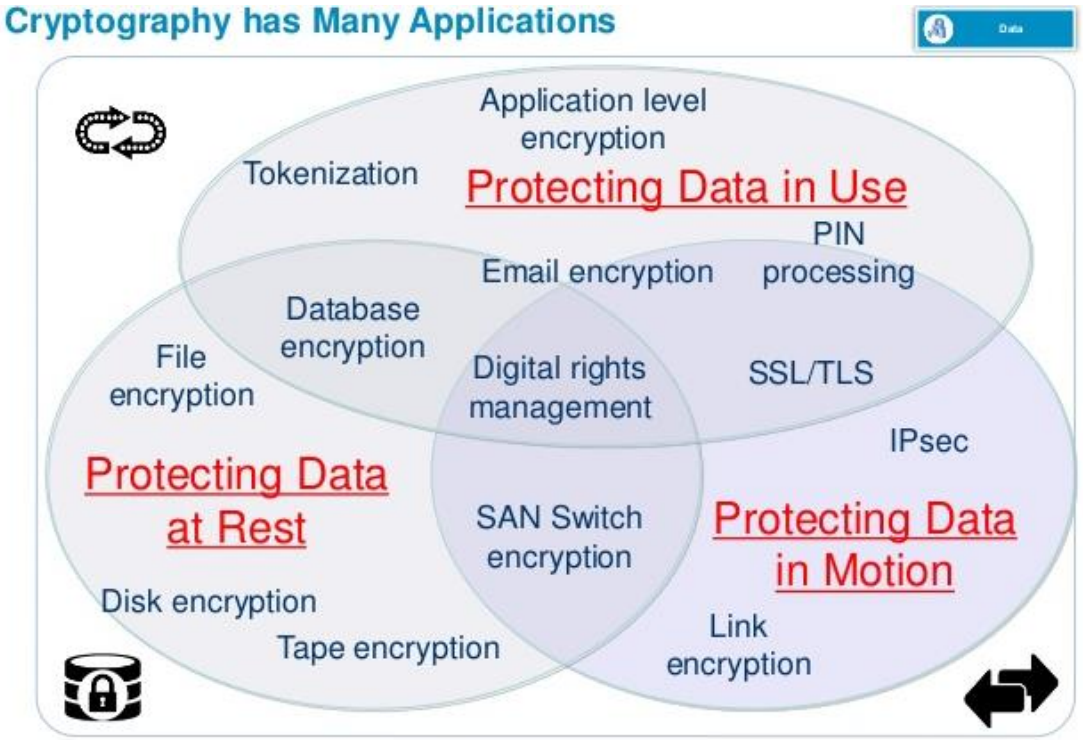


Figure 6: IBM Deployments on data protection

As clear, Data at rest is treated with several levels of encryption, including full disk encryption, file encryption, database encryption, and in case of utilizing Storage Area Network SAN then SAN Switch would be also encrypted. It is very important to notice that although encryption is mandated here, digital rights or authentication, authorization, and access control AAA measures are well taken care of. AAA concept is one of the fundamental roles when it comes to security in general, and practically should never be omitted. Secondly, with data in motion/transit between networks, network security practices are followed. Here, these would include protocols of Transport; Network; and Data Link layers of the OSI Internetworking

model; and can include but not limited to TLS/SSL; IPsec; tunneling and VPN techniques; PPP; RADIUS; 802.1X; TACACS; and so on.

At a higher level comes data in use protection. Protection at this level needs to be done using: lightweight mechanisms to not affect operations or cause burden; fast enough, thus to be able to handle the change in the stream of inputs and outputs of an operation; and finally inclusive, to further include the part of the memory where operations are withheld. Similarly as previous solutions, encryption is the major technology used here; this would include application encryption, email, and database encryption. Moreover, the solution also focuses on AAA procedures and monitoring techniques, data anonymization, RAM encryption with the possibility to store keys at CPU storage, and finally tokenization solutions for small deployments.

4.2. Off-site Storage Solutions

Currently many of the off-site storage solutions are more preferred than the typical on-site on-premises ones, as they could solve for the distributed data storage issue that is arising. Typically, these solutions incorporate Cloud Storage, Network Attached Storage NAS, and Storage Area Network SAN. These solutions benefit from the reduced costs and the minimized configurations, and/or maintenance required compared to the traditional ones; however, they also come with their own flaws.

4.2.1. Cloud Storage

Cloud storage is when data is stored off-premise at a remote location, thus storage and operations can be maintained by another party than the owner. Cloud storage features many flexibilities, e.g. the ability to control the storage size on demand, and the option to process data more powerfully by means of the shared infrastructure. Cloud storage is classified according to the mode of operation into: Private, Public, Hybrid, and Community Cloud. These modes only differ with how resources are shared. For instance: with private cloud, resources are shared exclusively among one party; for community, they are rather shared among several parties simultaneously; and for public cloud, resources are shared among all participants. Still, in all modes storage resources are remote and managed by a third party.

Cloud computing and cloud storage provide for good and low-priced solutions that solve for the proliferation of data resulted with the current deployments. However, some factors should be taken into consideration when adopting cloud solutions, as follows:

a. Cloud storage does not provide for the highest speeds to store and retrieve data in contrast to the Direct Attached Storage DAS implementations. Cloud storage also needs Internet connection consistently to function, so an ICT infrastructure should be in place.
b. Data is operated and location by third party, which highlights the issue of privacy and regulations with data handling.
c. Regulations and laws, as these might be different from a place to another.

It is recommended when implementing cloud storage solution to follow the best practices to protect data in remote. These include but not limited to the following:

a. Network and Internet security practices.
b. Anonymization of data, and tokenization of sensitive data.
c. Encryption of raw data, thus to prevent data mining or operations by the third party.
d. Data federation, thus to prevent data from being stored at a storage facility out of the agreed region.

4.2.2. Network Attached Storage

Network Attached Storage NAS is a model in which external storage media are attached and accessible through the network. The model resembles an external storage media that is rather a centralized one. In this model, connected devices can offload their data to NAS directly instead of storing data locally, thus devices do not need extra storage than the minimal required. NAS features the capability of large storage capacity, and being accessible as other network devices by means of Ethernet and Wireless media utilizing the standard IP protocols. Figure 7 shows a typical NAS.

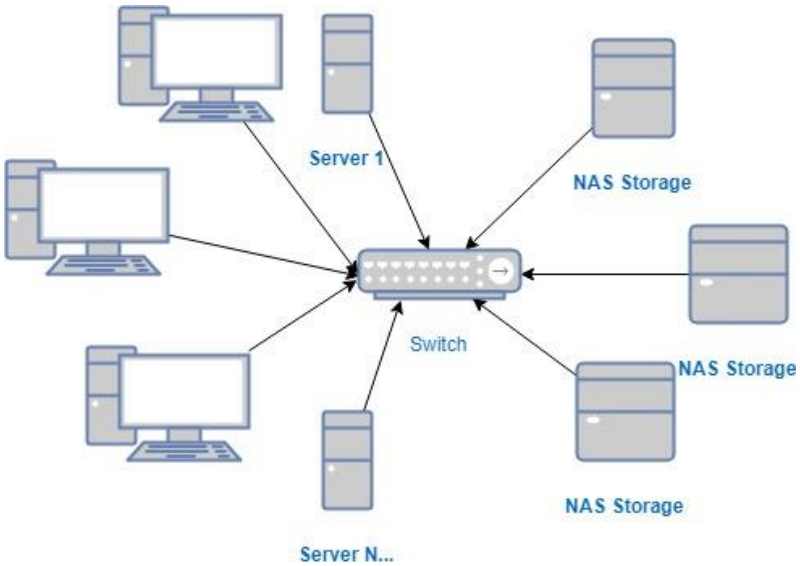


Fig 7: Typical NAS

The main advantage of NAS is seen when devices are in near proximity, forming a Local Area Network LAN, or Campus Area Network CAN. In these networking models, NAS would provide for the highest data transfer rates, and storage capacity. NAS also can be configured for remote access, which means that it also can work with WAN networks via Internet. This latter promotes flexibility and mobility of devices as they can be totally independent of storage location. NAS solutions are highly scalable, as they allow storage expansion by means of replacing the storage media with ones of higher capacity, or by addition of more NAS systems with a NAS aggregator solution. NAS has many advantages to give, these include:

a. Distributed data storage support.
b. Fully online data storage.
c. Failover and Redundant Array of Intendent Disks RAID techniques to prevent failure and data loss.
d. Security features implementation by means of encryption.
e. Privacy Protection,.

Unlike cloud, NAS is fully managed on-premises; thus it does not suffer from the related privacy and third-party management issues. On the other hand, NAS is a network implementation, thus all measures of network security should be considered here by the data owners. NAS security recommendations include:

a. Use of Virtual Private Network VPN to protect data on the go.
b. Full Disk Encryption.
c. Secure Data Transfer, by means of HTTPS, SSL, and SFTP protocols.
d. Multi-factor authentication.
e. Configurations allowed only locally not remotely.
f. Antivirus implementation.
g. Network Access Protection NAP along with Intrusion Prevention System IPS.

4.2.3. Storage Area Network

Storage Area Network SAN is another technology for data storage remotely as NAS. However, unlike NAS which operates as one independent storage, SAN operates as a network of connected storage devices. SAN also differs with the way data is dealt with, as data is considered as blocks with SAN rather than files with NAS. SAN finally uses different technologies for access, as it uses Fiber Channel FC protocol to access data, which provides for higher speeds compared to NAS. With this difference in technology, SAN brings many advantages; however, SAN is for large implementations only due to its complexity, the operation, and management effort.

Like NAS, SAN follows the same security concerns and recommendations as it operates at the network level.

4.2.4. Summary of Distributed Data Storage Technologies Solutions

With the proliferation, generation, and integration of data at distributed locations, distributed data storage solutions are a must, and the right implementation should take all models into consideration. With the previous models explained, a layered storage is suggested below:

1. Minimal DAS of own devices.
2. For services in vicinity, NAS to be implemented.
3. SAN to be implemented remotely.
4. Cloud Solution avoidance or minimal usage, with the recommended practices of data protection and encryption.

This is shown as in Figure 8 below.

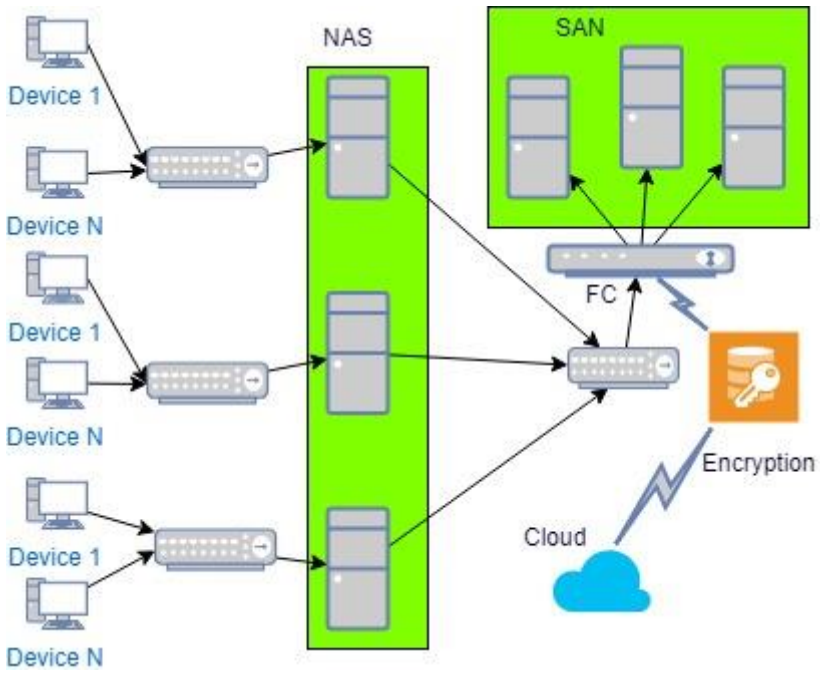
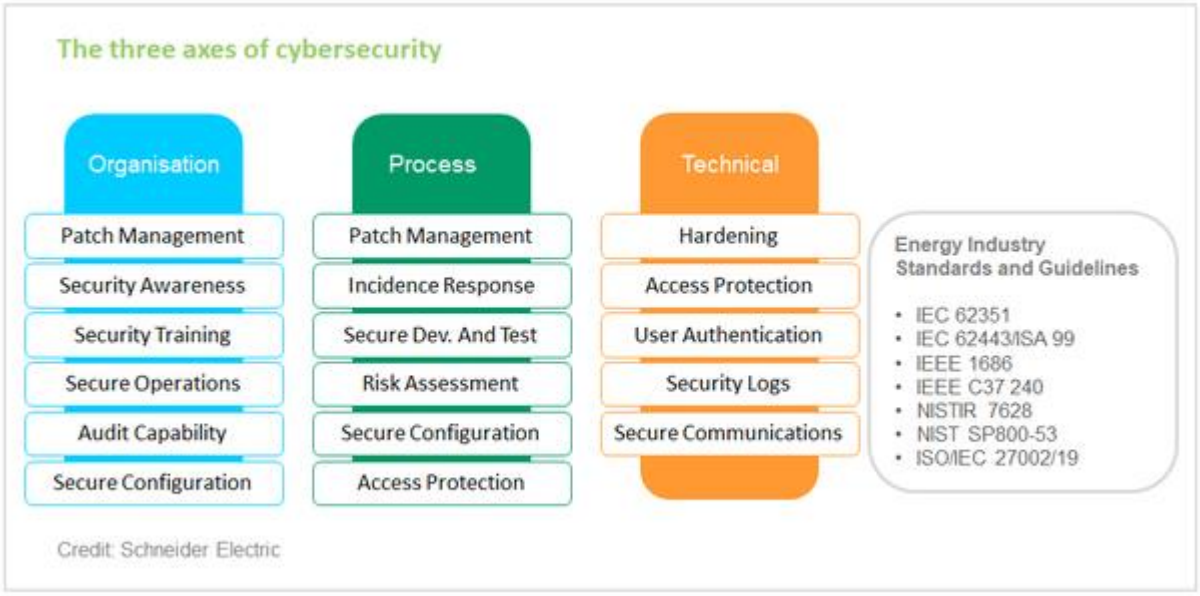


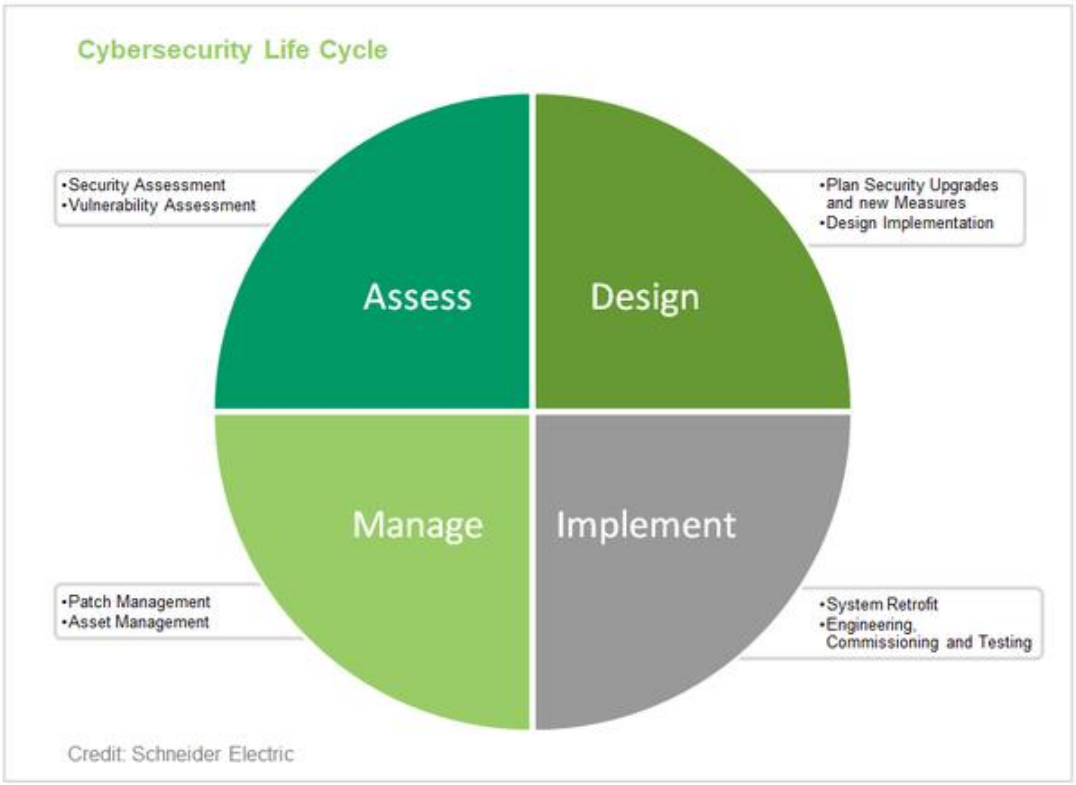
Figure 8: Hierarchy of Distributed Data Storage.

Extra Material

1.



2.



3. Best practices for cyber security in the electric power sector, IBM report, attached.