Re-Encryption for a Secure Cloud Computing Based Framework for Big Data Analysis of Smart Grid

#1 C. Deepika, #2 Dr. T. Hemalatha
#1 Research Scholar, School of Computer Science, Vels University
#2 Associate Professor, School of Computer Science, Vels University

Abstract:

Smart grid is a technological innovation that improves efficiency, reliability, economics, and sustainability of electricity services. It plays a crucial role in modern energy infrastructure. The main challenges of smart grids, however, are how to manage different types of front-end intelligent devices such as power assets and smart meters efficiently; and how to process a huge amount of data received from these devices. Cloud computing, a technology that provides computational resources on demands, is a good candidate to address these challenges since it has several good properties such as energy saving, cost saving, agility, scalability, and flexibility. In this paper, we propose a secure cloud computing based framework for big data information management in smart grids, which we call “Smart-Frame.”

The main idea of our framework is to build a hierarchical structure of cloud computing centre’s to provide different types of computing services for information management and big data analysis. In addition to this structural framework, we present a security solution based on identity-based encryption, signature and proxy re-encryption to address critical security issues of the proposed framework.

Keyword: Big data, Smart grid, Smart-frame, Cloud Computing, Re-Encryption

Introduction:

Power consumption is a very important terminology which makes India to be in bright. Power consumption refers to the electrical energy supplied over time to operate the electrical appliances like mobile, fridge, desktops, light, fan etc… where smart grid comes into existence.

smart grid is an electric gridwhich includes a variety of operational and energy measures including smart meters, smart appliances which is used to measure the power consumption of those devices, and it consists of renewable energy resources, and energy efficiency resources which can be used by those devices.

From these devices a huge amount of data are received. That information is very complex, and the data processing over those data is inadequate. It is not an easy task to manage these set of data, which includes selection, monitoring, and analysis of smart grid data.

The information, apart from users, it is also usable for the management services, distribution services etc…

There are many challenges while processing data in big data include analysis, capture, search, sharing, storage, transfer, visualization, and information privacy.

In real time, information processing is very difficult and it is required by smart grid. Delay in information processing may cause serious sequences to the whole system.

To make use of those data effectively and efficiently across the globe, we go for cloud computing technology where the information from those smart devices is maintained in cloud storage.

The information storage performs heavy tasks of distributing confidential data. Data which are processing over devices and cloud will be more secure. We can provide security indata processing by using encryption algorithms.

Let see over view of technologies that are used.

Big data:

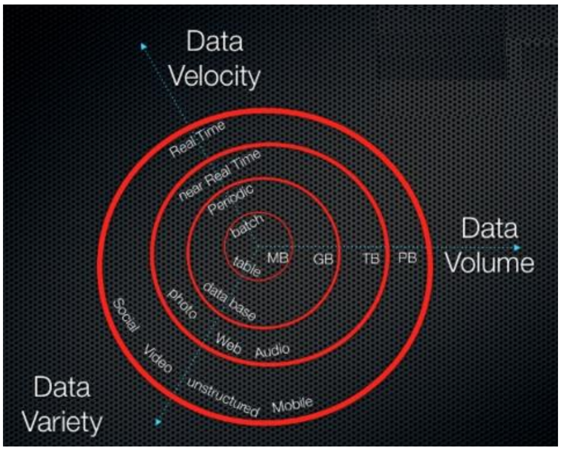
Big data is a concept which is used to describe a huge amount of data which is collected from various individuals, organizations etc… that may either be structured or unstructured. It becomes very difficult to process such data using traditional database models like (DBMS, RDMS) and software methodologies. A most important concern is that, if the volume of data is too big or it moves too fast or it exceeds current processing capacity, then it becomes a risky one.

Big data has the ability to provide, improve operations and it makes process faster, and take more intelligent decisions for the organizations.It gets origin from Web search companies who had the problem of querying very large distributed aggregations of loosely-structured data (XML,XHTMLand webbased document).

Characteristics:

Big data can be characterized by 3Vs:

- Volume: Big data is just a large amount of data. It simply observes and tracks the on-going process.
- Velocity: Big data is available in real time scenarios.
- Variety: Big data is a mixed data that can be drawn from text, images, audio, video etc…



Importance of Big Data:

When big data is effectively and efficiently captured, processed, and analysed products, competitors, which can lead to efficiency improvements, increased sales, lower costs,better customer service, and/or improved products and services. Companies are able to gain a more complete understanding of their business, customers,

Effective use of big data exists in the following areas:

- Using information technology (IT) logs to improve IT troubleshooting and security breach detection, speed, effectiveness, and future occurrence prevention.
- Use of voluminous historical calls centre information more quickly, in order to improve customer interaction and satisfaction.
- Use of social media content in order to better and more quickly understand customer sentiment about you/your customers, and improve products, services, and customer interaction.
- Fraud detection and prevention in any industry that processes financial transactions on-line, such as shopping, banking, investing, insurance and health care claims.
- Use of financial market transaction information to more quickly assess risk and take corrective action.

Evaluation of Big data:

Column-Oriented databases:

Traditional, row-oriented databases are excellent for online transaction processing with high update speeds, but they fall short on query performance as the data volumes grow and as data become more unstructured. Column-oriented databases store data with a focus on columns, instead of rows, allowing for huge data compression and very fast query times.

Schema-less databases or NoSQL databases:

There are several database types that fit into this category, such as key-value stores and document stores, which focus on the storage and retrieval of large volumes of unstructured, semi-structured, or even structured data. They achieve performance gains by doing away with some (or all) of the restrictions traditionally associated with conventional databases, such as read-write consistency, in exchange for scalability and distributed processing.

Map Reduce:

This is a programming paradigm that allows for massive job execution scalability against thousands of servers or clusters of servers. Any Map Reduce implementation consists of two tasks: The "Map" task, where an input dataset is converted into a different set of key/value pairs. The "Reduce" task, where several of the outputs of the "Map" task are combined to form a reduced set of tuples.

Cloud computing:

Cloud computing is a technology to access the resources available in the servers through Internet. Cloud computing technology becomes popular in the recent years due to its several advantages over traditional methods, like flexibility, scalability, agility, elasticity, energy efficiency, transparency, and cost saving. Cloud resources are shared resources which can be accessed by any one, anytime and anywhere. It is accessible through any devices like mobile, desktops, laptops, tablets etc... The resources and information are provided for the users based on on-demand services. It allows the users to pay only for the resources and workloads they use.

Cloud is nothing but a server and a number of servers interconnected through it. Cloud providers are the one who own large data centers with massive computation and storage capacities. They sell these capacities on-demand to the cloud users who can be software, service, or content providers for the users over the internet. In the recent years the major cloud providers are Google, Microsoft,and Amazon etc...

These clouds provide different types of Services:

Infrastructureas a Service:

Infrastructure as a Service is a form of cloud computing service which provides virtualized resources which are required over the Internet. Among many services it is an important one because, it provides, server spaces, bandwidth requirement, internet connections, load balancing etc…

Platformas a Service:

Platform as a service is a form of cloud computing services which provides a platform which allows customers to develop, run, and manage their web applications without the necessity of developing and maintaining the infrastructure which is required for developing and launching an application.

Softwareas a Service:

Software as a Service is a form of cloud computing services which provides the software's in which the developed applications are hosted by the service provider. Further, a service provider gives access for those applications to the customers through Internet by terms of pay per use.

Network as a Service:

Network as a Service is a type of business model which allows us to access the network functionalities directly and securely.A Service provider allows us to access the Internet virtually by terms of pay per use or for monthly basis.

Virtualization:

Virtualization is the key concept in sharing the resources. It allows the single instance of resources among multiple customers or among different organizations.Creating a virtual machine over existing operating system and hardware is referred as Hardware Virtualization. Virtual Machines provide an environment that is logically separated from the existing hardware.

Big Data in the cloud:

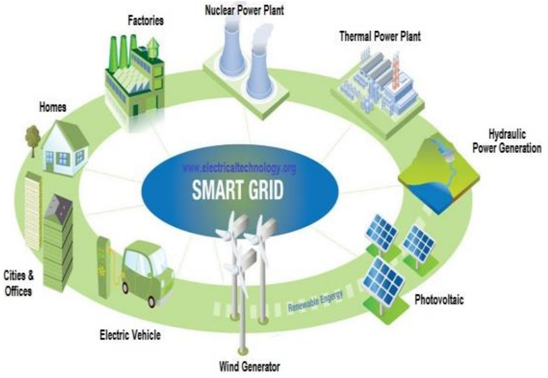
Most of the technologies are closely associated with the cloud. The products and platforms mentioned are either entirely cloud-based or have cloud versions themselves. Big Data and cloud computing go hand-in-hand. Cloud computing allows organizations of all sizes to get more value for their data than ever before, by enabling fast analytics at a minute of previous costs. This, in turn drives companies to acquire and store even more data, creating more need for processing power and driving a virtuous circle.

Smart grid:

Smart grid is an information management technique and involves three basic tasks: Information gathering, processing and storing.

Information gathering:

Smart grids are those which gathersinformation from different devices at different locations. The main research challenge is to build efficient communication architecture. Several solutions have been proposed to address this challenge for processing the data.



This proposal for standardization of data structures used in smart grid applications has recently addressed this issue. The Cloud computing appears to meet this demand and also satisfy challenges of information storing. The properties of smart grid and cloud computing were analysed to prove that cloud computing is a good candidate for information management in smart grids. Due to their large-scale deployment, smart grids suffer fromseveral security vulnerabilities. Since any securitybreach in smart grids may lead to a big loss there are initiatives to address security challenges in this type of systems.

Existing system and functions:

Security for the data is the main concern while transmitting or receiving the data between end user devices and the cloud. We can provide security for the data by means of algorithms by which secure transmission is possible. While providing security, the important is that, it will degrade the efficiency and performance of the system. Algorithms provide security by means of data encryption and re-encryption. If the smart grid store data in cloud, data is encrypted and transmitted and it is re-encrypted when data is processed.

Algorithm:

Identity based scheme is the existing algorithm used for security purpose. The idea of this algorithm is that, the cloud centres and the end devices are to be represented by their identities which can be used as encryption keys. By employing an identity-based re-encryption scheme, the information storages, which are components of regional clouds, can re-encrypt the received confidential data from cloud to devices. So that the services requested will decrypt the confidential data without compromising the information storage private keys.

Function:

Identity based scheme works as a two-step process. First, the identity of the data along with the identities of the high level entities are encrypted, and then, the output of the encrypted process is again sent as an input for further encryption to provide more security. In an identity-based encryption scheme, the private key generator (PKG), a trusted party, first generates secret master key mk and public parameter params. Note that params, which is long-term, will be given to every party that is involved.

Once a receiver submits their identity, denoted by IDrec, the PKG computes the private key KIDrec associated with IDrec by running the private key extraction algorithm Extract providing its master secret key mk as input. Here, the identity IDrec can be any string such as an email address, a telephone number, etc. Note that the distribution of the private keys can be done in a similar way as digital certificates are issued in normal public key cryptography.

Users wouldauthenticate themselves to the PKG and obtain private keysassociated with their identities. Secure channel may have to be established between the PKG and the users depending on the situation to prevent eavesdropping. Now any sender, who is in the possession of IDrec, encrypts a plaintext message M into a cipher text C by running the Encrypt algorithm. Upon receiving C, the receiver decrypts it by running the Decrypt algorithm providing the private key KIDrec obtained from the PKG previously as input.

Problems:

The main problem is that, it is a two-step process, where the number of thread requirement is more. So it is suitable only for less number data processing. If number of smart grid increased, data resources utilization will be increased. In parallel, the efficiency and the performance of the system is highly affected. The processing of huge amount of data efficiently still remains as a big challenge.

Solution:

To process huge amount of data effectively along with security, the solution is that, instead of using identity based scheme we can use triple-DES which requires less number of threads when compared with identity based scheme. It provides triple time more secure and increases the efficiency of the system.

Triple-DES Algorithm:

Triple DES (3DES) is the common name for the Triple Data Encryption Algorithm (TDEA or Triple DEA) symmetric-key block cipher, which applies the Data Encryption Standard (DES) cipher algorithm three times to each data block.

The original DES cipher's key size of 56 bits was generally sufficient when that algorithm was designed, but the availability of increasing computational power made brute-force attacks feasible. Triple DES provides a relatively simple method of increasing the key size of DES to protect againstMeet-in-the-middle attacks that are effective against double DES encryption. In cryptography, Triple DES is a block cipher created from the Data Encryption Standard (DES) cipher by using it three times. In general TDES with three different keys (3-key {k1, k2, k3} TDES) has a key length of 168 bits: three 56-bit DES keys (with parity bits 3-key TDES has the total storage length of 192 bits), but due to the meet-in-the-middle attack the effective security it provides is only 112 bits. Another version, called two-key TDES (2-key TDES), uses k1 = k3, thus reducing the key size to 112 bits and the storage length to 128 bits. However, this mode can be taken advantage of through certain chosen-plaintext or known-plaintext attacks and so TDES is treated by NIST to have only 80 bits of security. By design, DES and therefore TDES, suffer from slow performance in software. TDES is better suited to hardware implementations, which are many of the places it is still used.

Conclusion:

We have introduced the Smart-Frame, a general framework for big data information management in smart grids based on cloud computing technology. The secure aggregation protocols followed the bottom-up traffic model (i.e., device-to-centre), which is spread widely in power systems in earlier system. We focused specifically on providing our Smart-Frame with security framework based on identity-based encryption/signature and identity-based proxy re-encryption schemes. Already, the proxy re-encryption technique is applied to provide mobile applications in clouds with security. New we specifically apply identity-based cryptographic techniques to address the scalability issues of smart grid applications. One of the obvious benefits we can gain from applying identity-based cryptography to the Smart-Frame isthat through using identities rather than digital certificates which depend on traditional public key infrastructure (PKI),

Future enhancement:

From this proposal we identified the few limitation while increase the number of user. If top level data centre handled all the device information & user data, the performance will weaken. So we built the regional and zone level data centre for maintaining the data. The top cloud level provides a global view of the framework and other will provide the information to parent cloud.

From the above 3DES algorithm, we provided a solution based on "identity-based cryptography and identity-based proxy re-encryption"which provides secure communication services with the Smart-Frame. This will achieve not only scalability and flexibility but also security features.

References:

1. Chou, Timothy. Introduction to cloud computing business and technology.

2. "Realization of Interoperability and Portability among Open Clouds by Using Agent's Mobility and Intelligence" – TechRepublic.

3. Magoulas, Roger, Lorica, Ben – "Introduction to Big Dat.

4. M.Shargal and D.Houseman, "The big picture of your coming smart grid," Smart Grid News, Mar.

5. F.Li, B.Luo, and P.Liu "Secure information aggregation for smartgrids using homomorphic encryption," in Proc. IEEE Conf. Smart Grid communication.

6. Webster, John. "MapReduce: Simplified Data Processing on Large Clusters", "Search storage".

7. Boja.C: Pocovnicu "Distributed Parallel Architecture for Big Data".

8. http://www.hcltech.com/sites/default/files/solving_key_businesschallenges_with_big_data_lake_0.pdf

9. H. Li, Y. Dai, L. Tian, and H. Yang, "Identity-based authentication for cloud computing," in Proc.

10. H. Khurana, M. Hadley, N. Lu, and D. Frincke, "Smart-grid security issues," IEEE Security Privacy, vol. 8, no. 1, pp. 81–85, Jan./ Feb. 2010.