Hybrid Cloud Computing Platform: The Next Generation IT Backbone for Smart Grid

Fengji Luo, Student Member, IEEE, Zhao Yang Dong, Senior Member, IEEE, Yingying Chen, Student Member, IEEE, Yan Xu, Student Member, IEEE, Ke Meng, Member, IEEE, Kit Po Wong, Fellow, IEEE

Abstract—This paper discusses the prospective of applying cloud computing technologies in the development of smart grid. Firstly, the conceptions of cloud computing are introduced, and then a hybrid cloud computing platform for smart grid is designed. After that, the distinguished characteristics of the proposed platform are explained in detail, following with the introduction of some potential power system applications. Finally, some notable state-of-the-art products that can be used to build the proposed platform are introduced.

Index Terms—Cloud Computing, Distributed Computing, Smart Grid

I. INTRODUCTION

MODERN power system is experiencing the rapid development and reorganization. One of the main trends is that the scale of the systems is becoming larger and larger. Along with the ongoing development of the long distance transmission systems, the super scale systems which can cover one or more countries are emerging; on the other hand, the structure of the system is becoming more complex, due to the transiting from vertically organization to deregulated power market. The rise of “Micropower” also increases the system complexity. Such large scale and complex environment imposes many challenges on design, analysis and control of the system.

Another major move is the smart grid development. With smart grid infrastructure in place in many nations, the information and data involved in the operation and planning of the power system, especially at distribution level, have increased in scale of tens of hundreds of times. For example, for distribution feeder monitoring, under the conventional system, a couple of measurement would be sufficient to estimate the feeder voltage at customer end and would meet the operation’s needs. However, with the smart grid in place, there could be tens of measurements for the same feeder. The processing power of the computing facilities will inevitably be facing big challenges with large scale networks.

(1) Computing Power. The smart grid needs more powerful computing platforms to handle large scale data analysis tasks and support complicated real-time applications. Currently, all these computing tasks are often independently scattered in each dispatch center with its own computing platforms (server, cluster, etc). Such centralized pattern has two disadvantages. Firstly, due to the limited processing capability, it is very difficult to support online analysis, unless simplifying the model, which will unavoidably affect the accuracy of analysis results. Secondly, the capital investment and maintenance cost of these computing facilities is very high, but the utilization is relatively low. For instance, the idle capability of the computing power of one dispatch center cannot be coordinated to help solving problems of another dispatch center.

(2) Data Storage. The smart grid facilitates interactive two-way communications and requires the powerful ability to collect and analyze tremendous data. This leads to the challenge of effectively integrating and seamlessly accessing the unprecedented data, which come from a variety of sources, advanced metering infrastructure (AMI), wide area measurement system (WAMS), and even embedded home compliances.

In the past, parallel computing has been widely adopted in power system analysis [1]-[4]. Furthermore, in literatures [5]-[8], the grid computing technology is proposed to solve power system problems. However, those approaches have their own limitations. Parallel computing often has the computing power bottleneck, and thus hard to meet the heavy and various online analysis tasks of the smart grid; grid computing platform can integrate distributed resource, however, it requires coordination among different parties and lots of funding, and it is difficult for end users to seamlessly access the grid resources due to lack of enough abstraction.

This paper designed a hybrid cloud computing platform as the next generation IT backbone for the development of smart grid. Cloud computing is a novel computing model that has drawn worldwide concerns in the recent years. However, there is still no standard definition for the cloud computing. Some anticipated features or characteristics of cloud computing by industry and academia can be envisaged and summarized as, (i) pay-per-use; (ii) elastic capacity and the illusion of infinite resources; (iii) self-service interface; (iv) resources that are abstracted or virtualized. In this paper, we will show that using the virtualization technology and serve-on-demand model, cloud computing can meet the IT requirements of smart grid in an economical and effective way.

This paper is organized as follows. In Section II, the conception of cloud computing is introduced, followed by the de-
sign and discussions of the designed hybrid cloud computing platform for smart grid in Section III; and then in Section IV some notable cloud computing products are introduced respectively; and conclusions are drawn in Section V.

II. CLOUD COMPUTING

Entering 21th century, the cloud computing has widespread concerns by various industries. Nowadays, several major IT enterprises have offered cloud services, such as Microsoft [10], Amazon [11], and Google [12]. In this section, we will give a brief introduction of cloud computing, in order to make power industry engineers better understand the cloud computing. This section consists of two parts. Firstly, the basic introduction of the cloud computing is given. Secondly, the service model of cloud computing will be introduced. The more detailed discussions of cloud computing can be found in [9], [13], [14].

A. Introduction of Cloud Computing

Currently, there is no strictly standard definition of cloud computing, however it has been widely accepted that it denotes a model on which a computing infrastructure is viewed as a ‘cloud’, from which businesses and individuals access applications from anywhere in the world on demand [9]. The main principle behind this model is offering computing, storage, and software “as a service”. For example, the infrastructures established by Amazon and Google respectively to offer on demand services are called two “clouds”.

Cloud computing is not a new fundamental technology. It is the evolution and integration of several technologies. This can be shown in Fig. 1.

![Fig. 1. Critical technologies of cloud computing [9]](image)

Each technology represents an aspect of cloud computing. Although each technology can be written as a whole book, here we just introduce the key features of each technology.

1) Web Service

Web service [15] represents a sort of standards which make the web applications interact with each other following a common mechanism. Using web service, each software or application is packaged as a “service”, which is self-described and self-contained to provide certain business function. A service is independent of the context of other services, and different services can interact with each other. Some standardized machine-interpretable interfaces and protocols are formulated to describe a service, publish a service, and transport messages among services. Relying on these standards, people can ‘glue’ together several services to create complex applications. Later we will show that in the cloud computing platform, the complex power system applications can be created by composing others services from different service providers.

2) Grid Computing

Grid computing [16] provides a way to aggregate geographical distributed computational resources and coordinately use them to accomplish certain tasks. Users submit tasks to the grid portal, and grid broker chooses suitable resources to execute the job, and finally return the result to the users. Cloud computing platform is backed by thousands of distributed resources, thus it often uses the idea of grid computing to aggregate thousands of resources to form the resource pool.

3) Utility Computing

Utility computing [17] is based the idea of “pay-per-use’. In the utility environment, users assign a “utility” value to their job, where a “utility” represents a certain measurable requirement (execution time, throughput, etc), and the value is the amount of money they are willing to pay the service providers to their demands. Thus this environment forms a marketplace, where services providers aim to maximize their profit while users compete for resources based on the utility value of their jobs.

4) Hardware Virtualization

Virtualization [18] is one of the core idea of cloud computing. In order to allow users easily access and use the distributed, heterogeneous resources in a uniform way, virtualization technology is necessary. The virtualization technology adds a software layer on the physical hardware layer, which is often called Virtual Machine Monitor (VMM). Based on the VMM, one or more virtual machines can be established, while each virtual machine runs the operation system and user’s applications. Here, the operation system and software stacks no longer bind to particular physical machines, but run on the VMM and the VMM maps the virtual machines to the physical hardware. This structure can be shown in Fig. 2.

![Fig. 2. Virtualization technology.](image)
protection, which can be achieved by using intelligent control and management systems.

According to the deployment model, a cloud can be classified as public cloud, private cloud, community cloud, and hybrid cloud [9]. The specifications of them are given as follows:

(a) Private Cloud. This is the cloud computing model run within a company’s own data center for internal and/or partner use;

(b) Public Cloud. This is the 3rd party, multi-tenant cloud infrastructure to serve the public/global users;

(c) Community Cloud. Community cloud shares infrastructure between several organizations from a specific community with common concerns, whether managed internally or by a third-party and hosted internally or externally.

(d) Hybrid Cloud. Hybrid cloud refers to the cloud computing infrastructure composed by two or more clouds (private, public, or community). By establishing hybrid cloud, enterprises can integrate public clouds to complement the internal resource shortage while keeping the cloud used internally.

B. Service Model

Using above technologies, cloud computing platform could provide different layers of services, and each layer represents a level of abstraction, as shown in Fig. 3.

![Fig. 3. Service model of cloud computing platform.](image)

(a) Infrastructure as a service (IaaS). The virtualized physical devices can be delivered as the service in an on demand way. Users can access the virtualized devices through the VMM and do the operations on them, such as installing the software package, attaching the virtualized disks on the virtualized computer, etc;

(b) Platform as a service (PaaS). A higher level of abstraction is delivering the programming platform as a service. Users can access such development environment and develop cloud-enabled applications, and submit the applications to execute in the cloud platform;

(c) Software as a service (SaaS). On the top of the service model, mature cloud software can be delivered as the service in the on demand way. Relying on the web service technology, users can directly access and use them through web browsers.

III. CLOUD COMPUTING INFRASTRUCTURE FOR SMART GRID

Section II introduced the basic conception of cloud computing, the supported technologies, and the layered service model. In this section, the proposed cloud computing-based architecture for smart grid will be described, and then we will analyze how power system can benefit from this model.

A. Working Mechanism of the Platform

The hybrid cloud infrastructure for power system is shown as Fig.4. The model is called ‘hybrid cloud’ because it consists of the conception of community cloud and public cloud.

1) Power Community Cloud Data Center

Just as shown in Fig. 4, in power systems there are a large number of distributed IT resources, such as the storage devices and the computational devices. Those devices are distributed and owned by different organizations, such as different dispatching centers, market pools, GENCOs, etc. In the proposed cloud computing platform, those distributed data or computational resources will be aggregated by the power cloud data center. In the data center, the VMM provides the virtualization and physical accessing of those resources. After virtualization,
in the data center, the virtualized resources form a virtual resource pool managed by VMM, which can be delivered as the service to the end users. From the point view of users, they just apply for the virtualized resources from VMM, and operate those resources. They do not need to care about the details and the locations of the physical resources. Based on virtualization of power system resources, the power cloud data center provides higher level, power system-oriented services. Firstly, the cloud data center provides the PaaS level of service, which enable users develop cloud-based power system programs and run them in the virtualized resource pool. Those programs could be the offline planning programs, the data processing program, on line analysis applications, or scientific simulation programs. Unlike traditional power system applications development, the cloud-based applications often should be developed using some particular programming models. Once developed, the data center is responsible for applying on demand virtualized resources to execute.

The power system analysis is often data-intensive, thus the data collected from various kinds of physical data sources must finally be delivered to end users as the service. In the proposed model, this is implemented in the SaaS level. In the power cloud data center, the data providing service is encapsulated as an individual application and published to the end users. Users acquire the data by accessing this application interface.

In the level of SaaS, the cloud data center can publish different kinds of individual applications. These applications are encapsulated in the form of web service. Users can directly use these applications on demand, just through the web browsers. In the SaaS service layer, there could be many kinds of power system applications delivered as services, such as the power flow analysis program, the stability simulation software, the power market simulation tool, etc.

2) Integrating Public Clouds

One problem of the proposed power community cloud data center is the resource bottleneck problem. Proposed community cloud model uses the strategy of aggregating distributed computational/data storage devices of different power system organizations to coordinated work together. However, only relying on existing resources could not fully satisfy the computing and data processing requirements of future power system. There are 3 reasons. Firstly, power system applications are becoming more and more data-intensive and compute-intensive; secondly, there are wide ranges of power system applications; thirdly, the physical resources of power system were not originally established for constructing cloud data center. On the other hand, one notable feature of cloud computing is providing the ‘infinite’ resource view to end users. So, how to address this problem? There are 2 approaches. One is upgrading the physical resources, but it needs a lot of funding and requires lots of labor hood and time. Therefore, in the proposed model, we adopt another approach, that is integrating the public clouds to compensate the existing power system resources. The public clouds, such as Amazon EC2, own a huge number of resources and allow users to lease them. By utilizing the public clouds, power cloud data center can lease the resources to coordinated work if necessary, and finally provide infinite resource view to the users.

3) Interaction with Other Power Cloud Data Centers

Practically speaking, it is difficult to build a cloud data center covering all resources of the whole system. More practically situation is building data center for a certain area. Since power system is a large inter-connected system, it is unavoidable for different power cloud data centers to share information. In the proposed model, the data centers can exchange each other’s data in the SaaS level. One power cloud data center may also access the applications provided by other data centers to construct complex power applications.

B. Characteristics of the Platform

As discussed above, cloud computing model aggregates distributed resources and provide different levels of service to users. So, compared with parallel computing and grid computing which have been already successfully applied in power system, what are the distinguish features of the proposed cloud model? Following are 7 unique characteristics which cannot be achieved by other parallel/distributed technologies.

1) Infinite Resource

Parallel computing and grid computing has the resource bottleneck problem. However, the proposed cloud computing platform has the ability to integrate infinite resources. This is achieved by utilizing the public clouds. The public cloud providers are often the leading IT enterprises of IT industry, and declare that they can deliver infinite resources to public [9]. Therefore, by leasing the resources from public cloud providers if necessary, the proposed cloud platform can easily obtain any number of resources as more as needed.

2) Service Oriented Architecture

Other distributed technologies can also aggregate various kinds of resources, and provide interfaces to let users access them. But proposed cloud platform use the service oriented architecture (SOA) to deliver everything as web services. The difference lie in that each web service is encapsulated as a self-contained functional module, which is described and published to public using standard protocols and languages. Other programs can access this service and communicate with this service using standard protocols. This means that based on the standardized service framework, people can ‘glue’ several services freely to create complex application. For power systems, for example, the power flow analysis can be encapsulated as a web service while a stability analysis can be encapsulated as another web service. The users can write glue program to combine them to create a stability study application.

3) Improve the Power System Resource Utilizing

The resources of power system are geographically distributed and owned by different organizations. The IT resources are often established for their expected maximum load [9], this leads to the utility of the resources are often very low if there lacks of an effective mechanism to make them coordinate work together. For example, a power company owns a 20-node Linux PC cluster, but this equipment is only used to handle the large scale analysis 3-5 times each month. Then, in most of the
time, the cluster is just ether idle or handles some lightweight computing tasks. Obviously, this leads to the low degree utilizing of the cluster and unnecessary cost to that power company. In the proposed model, the resources are abstracted to form a virtualized resource pool. The end users just operate the virtualized resource pool, and VMM is responsible for mapping the virtualized resources to the physical resources. In this process, VMM will use the advanced load balance algorithm to maintain the load balance of the system and maximize the resource utilization. For example, the VMM may close some servers if they do not participant in executing the computing tasks for a long time, in order to save the energy.

4) **Distributed Mass Power System Data Management**

One feature which cannot be achieved by traditional parallel/distributed technologies is that they do not have the ability to store and management of the distributed mass data. In the proposed model, the data is delivered as the service in the SaaS layer, which means users can just access the service to get data. However, in the background, the data center stores the data distributed in the form of blocks, and the data blocks are often duplicated as mirrors in different servers. When a data query demand coming, the data center can firstly check the location of the user and then choose the nearest server which contains the copy of the required data, and transfer data from that server to the user. In this way, cloud computing can reduce the data transferring time a lot to provide efficient data management service.

5) **On Demand Elastic Service Delivery**

The elastic service delivery is one of the notable features of cloud computing. In the traditionally parallel computing and grid computing technologies, for a particular application, the amount of resources used to execute it is determined ahead of execution (this process is called scheduling). Such scheduling process is static. In cloud computing, since all the applications are running in the virtualized resource pool, there is no strong binding between applications and physical resources. Therefore, VMM can assign more/less resources to the applications adaptively according to the runtime requirements of the applications. Such elastic, flexible scheduling mechanism can improve the performance and utilization of the system a lot.

6) **Lightweight Accessing Media**

Traditionally, power engineers often install EMS system in their PC or work stations to solve power problems. Cloud computing changes such computing model fundamentally. In the proposed model, power operators, engineers, and researchers can access the power cloud data center using various kinds of media, as long as those media support web browser. Those media may include PC, work station, mobile phone, PDA, etc. Since everything is delivered as the services and can be accessed through the HTTP/XML protocols, power engineers no longer need to install any client software. They can connect the power cloud data center at anywhere of the world using any media, and use web browser to acquire data, develop applications, use power software to perform power system study.

7) **Robust, Self-healing and Fault-tolerance**

Comparing traditionally technologies, the cloud computing is more robust and fault-tolerance, and it is self-healing. In parallel computing, applications are divided into several sub-tasks to paralleled run in the multi-core machines or clusters. Once the physical machine breakdown, the applications cannot be accomplished. In grid computing, the computing tasks are running in the geographical distributed resources. Once some resources breakdown, the tasks on those resources must be migrated. Since those tasks are bind to physical resources, they are very hard to be migrated into other heterogeneous resources.

On the contrary, in our proposed model, applications are only bind to the virtualized resource pool. The migrating of the applications thus becomes very easy. Once some breakdowns of the physical resources happen, the VMM encapsulates the state of the applications and migrates them into other virtual machines automatically. The state of the applications is only bind to the virtual machine, rather than physical resources. In this way, power cloud data center could operate in a robust and self-healing way.

IV. **POTENTIAL APPLICATIONS**

In above sections, the hybrid cloud computing platform for smart grid is proposed and its working mechanism is explained. This section illustrates some potential power system applications which can be developed based on the cloud computing platform. However, it is notable that this section just gives a brief introduction, while many new kinds of power system applications could be developed on the cloud computing platform rather than just limited to the applications list in this section.

A. **Generation Scheduling**

In the operation of the power system, the generation scheduling mainly consists of unit commitment (UC) and economical dispatch (ED). Traditional methods, such as the dynamical programming, Lagrange relaxing method, have the computing time or convergence problems. In order to find the global optimization solution while shorten the computing time, many heuristic methods are proposed [19, 20, 21]. Those algorithms search the global optimal solution in the high dimension space, and often can be easily paralleled implemented in the cloud computing platform, relying on the aggregated virtualized resource pool provided by the cloud computing platform.

B. **Transient Stability Simulation and Analysis**

Literature [4] accelerated the PSS/E dynamic simulations by using the EnFuzion-based distributed computing method. The author used EnFuzion software to dispatch the PSS/E time-domain simulations which incorporate different contingency cases to multiple machines, and solve them simultaneously. This approach released the computational burden greatly and improved the simulation speed effectively.

Recently, Enfuzion has published the cloud computing version, which is called CloudFuzion [22]. CloudFuzion can lease the resources of Amazon EC2 and dispatch the jobs to the resource pool of Amazon EC2. Therefore, it can be expectable that by employing the CloudFuzion, above transient stability analysis could be easily migrated to cloud computing platform and obtain better performance.
C. Power Flow, Probabilistic Power Flow and Distributed Power Flow

In order to enhance the performance of power flow algorithm, several literatures proposed different parallel power flow algorithm, such as Newton-based parallel power flow algorithm [23] and the Differential Evolution algorithm based optimal power flow algorithm [24]. On the other hand, Monte-Carlo probabilistic power flow algorithm is a useful method to simulate the uncertain of the power system, while due to its high computational cost, Monte-Carlo probabilistic power flow algorithm is often paralleled to run on multiple CPUs. The virtualized resource pool of cloud computing platform is very suitable for running above power flow algorithms, and because of the massive computing resources provided by cloud computing, those algorithms can get more precise result comparing with traditional parallel computing.

Literature [25] proposed a kind of Distributed Power Flow method, which incorporates the boundary data of the neighborhood areas to perform power flow calculation. The web service-oriented cloud data center can satisfy the information exchange of such kinds of data-intensive application very well. Through the web service interface of the data center, different areas can easily exchange data with each other and then perform the Distributed Power Flow using the virtualized resource pool.

D. System Restoration

The power system restoration is a complex, nonlinear optimization problem. The characteristics of the modern power system, such as the deregulation, long distance power transmission, and the large number of distributed power sources, impose new challenges to the system restoration problem. The data center of the proposed cloud computing platform enhances the information sharing among different participants. The participants can transparently access the information through the data center and coordinated work with each other. At the same time, the virtualized resource pool and the participation of the public cloud provides the powerful computing ability to deploy the parallel/distributed system restoration algorithm to find the optimization system restoration solution.

V. IMPLEMENTATION TECHNOLOGIES

In Section III, the working mechanism and distinguish features of the proposed cloud computing platform for power system is introduced. Section IV illustrates some potential power system applications which can be developed on the cloud computing platform. Another problem is how to build such platform practically? Currently, there are many start-of-the-art products which can be used to construct cloud computing platform. Some of them are open sourced, and some of them are commercial. In this section, we will give a summary of some notable technologies.

A. Aggregating Distributed Resources

Globus Toolkit (GT) [26]: The GT is an open source grid computing software that addresses the resource sharing problems in the distributed environment. It is a middleware that let people share computing power, databases, and other tools securely online across corporate, institutional, and geographical boundaries without sacrificing local autonomy. The toolkit includes software services and libraries for resource monitoring, discovery, and management, plus security and file management. In the proposed community cloud computing platform, GT can be used to aggregate the distributed power system resources.

B. Resource Virtualization

1) Apache Virtual Computing Lab

The Apache Virtual Lab [27] is an open source software aiming to provide desktop and high performance computing environments anytime, in a flexible cost-effective way and with minimal intervention of IT staff. It has following features: self-service Web portal to reduce administrative burden; advance reservation of capacity to provide resources during classes; and deployment of customized machine images on multiple computers to provide clusters on demand.

2) Enomaly Elastic Computing Platform

The Enomaly Elastic Computing Platform (ECP) [28] is a commercial tool that offers most features a service provider needs to build an IaaS cloud. Most notably, ECP offers a Web-based customer dashboard that allows users to fully control the life cycle of virtual machines. It also allows providers and users to package and exchange applications. In summary, Enomaly ECP has following features: Linux-based controller, Web portal and web services interfaces; interface to the Amazon EC2 public cloud; virtual network; virtual clusters.

C. Parallel Programming Framework

MapReduce [29] is a framework proposed by Google for processing highly distributable problems across huge datasets using a large number of computers, clusters or a grid. Computational processing can occur on data stored either in a file system or in a database. The working mechanism of MapReduce is divided into “map” step and “reduce” step. In “map” step, the master node takes the input, partitions it up into smaller sub-problems, and distributes them to worker nodes. A worker node may do this again in turn, leading to a multi-level tree structure. The worker node processes the smaller problem, and passes the answer back to its master node. In the “reduce” step, the master node collects the answers to all the subproblems and combines them in some way to form the output.

MapReduce has been already used in cloud computing by Google. And Apache also provides an open source implementation of MapReduce, which is called Hadoop MapReduce [30].

D. Mass Data Storage and Management

1) BigTable

BigTable [31] is a compressed, high performance, and proprietary database system built on Google File System. It is currently not distributed nor is it used outside of Google, although Google offers access to it as part of their Google App Engine. BigTable is designed to scale into the petabyte range across “hundreds or thousands of machines, and to make it easy to add more machines to the system and automatically start taking advantage of those resources without any reconfiguration.
2) Apache Cassandra

Apache Cassandra [32] is an open source distributed database management system. It is designed to handle very large amounts of data spread out across many commodity servers while providing a highly available service with no single point of failure. The data model of Apache Cassandra is similar with BigTable.

VI. CONCLUSIONS

Cloud computing is a new computing model which is under rapid development. However, there are still some challenges when implementing cloud computing technologies to power system. One problem is data transferring. Currently, the data collected by sensors are transferred to each organization (such as control center) and are processed locally. In cloud computing environment, the data must be sent to cloud data center and delivered as service. As a result, ensuring the data transferring speed is a critical problem, especially for the online data processing applications. Another challenge is the security of the cloud computing environment. There are different types of users of the power cloud data center, such as generation companies, system operators, market investors, consumers, regulators, governments, etc. How to ensure the information privacy while keeping the information sharing is another problem needs to be addressed.

In this paper, the IT requirements of smart grid are analyzed firstly, followed by the explanations of distinguished features of cloud computing platform. After that, a community cloud computing platform is proposed and the detailed working mechanisms are introduced. Some potential cloud computing-based power applications are then introduced. Finally, in order to help power industry engineers build a customized cloud computing platform, several state-of-the-art products that can be applied in different areas of cloud computing are introduced.

REFERENCES