ABSTRACT

The new developments in the field of information technology offered the people enjoy, comfort and convenience. Cloud computing is one of the latest developments in the IT industry also known as on-demand computing. It provides the full scalability, reliability, high performance and relatively low cost feasible solution as compared to dedicated infrastructures. It is the application provided in the form of service over the internet and system hardware in the data centers that gives these services. This technology has the capacity to admit a common collection of resources on request. It is proving extremely striking to cash-strapped IT departments that are wanted to deliver better services under pressure. When this cloud is made available for the general customer on pay per use basis, then it is called public cloud. When customer develops their own applications and run their own internal infrastructure therefore is called private cloud. Integration and consolidation of public and private cloud is called hybrid cloud. But having many advantages for IT organizations cloud has some issues that must be consider during its deployment. The main concern is security privacy and trust. These issues are steps during the deployment of mostly public cloud because in public cloud infrastructure customer is not aware where the data store & how over the internet.

Cloud Computing has all the attributes and potential to support a global BPO environment. These attribute are: virtualization, service oriented architecture (SOA), utility based pricing and grid computing. Cloud Computing involves the movement of IT services – application, infrastructure and platform – onto the Internet and deployment models. Because of the high availability, high bandwidth and the increased use of the Internet it has become easier to access a variety of services, traditionally originating from within a company’s data center.

Keywords

Cloud computing; Data center; Virtualization; SOA; Service models; Virtualization; Grid Computing.

1. INTRODUCTION

A recent phenomenon in the domain of outsourcing is called Cloud Computing. “Clouds are a large pool of easily usable and accessible virtualized resources (such as hardware, development platforms and/or services). These resources can be dynamically re-configured to adjust to a variable load (scale), allowing also for an optimum resource utilization. This pool of resources is typically exploited by a pay-per-use model in which guarantees are offered by the Infrastructure Provider by means of customized SLAs” (Vaqüero, 2009). This means actually that more and more IT services – applications and technology – are outsourced to external vendors over the Web, which eventually will lead to a change in the traditional business model – where IT is in-house organized to a virtual enterprise. This virtual enterprise, based on mainly Cloud services, could be the future perspective. Meanwhile organizations are looking into business process outsourcing (BPO), which involves the delegation of an entire business process to a third party provider, including its supporting services There are generally three types of IT services which an organization can send into a cloud environment, namely: PaaS (platform as a service), SaaS (software as a service) and IaaS (Infrastructure as a Service)This research will address the most significant differences between them and provide the basic architecture and layers of the Cloud Computing modelA Cloud is a type of parallel and distributed system consisting of a collection of interconnected and virtualized computers that are dynamically provisioned and presented as one or more Unified Computing resources based on service-level agreement established through negotiation between the service provider and consumers. Cloud Computing refers to both the applications delivered as services over the Internet and the hardware and systems software in the datacenters that provide those services (Software as a Service) The datacenter hardware and software is what we will call a Cloud. When a Cloud is made available in a pay-as-you-go manner to the public, we call it a Public Cloud; the service being sold is Utility Computing Researchers propose a three way model for provision and usage of “cloud Services”, that could also be seen as recursive in case of mash-up provider that is a cloud user of another platform at the same time:

Cloud Provider → SaaS Provider / Cloud User → SaaS User

2. TECHNOLOGIES PRINCIPLES

The concept of Cloud Computing is based on a collection of many old and few new concepts in several research fields like Service-Oriented Architectures (SOA), distributed and grid computing as well as virtualization. Cloud computing is often compared to the following technologies, each of which shares certain aspects with cloud computing.

2.1 Service oriented Architecture

2.2 Distributed and Grid Computing

2.3 Virtualization

2.4 Utility Computing

2.5 Autonomic Computing

2.1 Service-oriented Architecture

Service-oriented Architecture refers to a modular design principle in Software architecture. Service-orientation aims at separating individual functions into distinct units or “services”, that could be accessed, e.g. via a network, by developers to integrate them in a reusable manner in their applications.SOA is an architecture model, not a technology itself. SOA is an architectural concept that aims to achieve loose coupling between service requester and service provider, by means of well-defined technology-agnostic contracts supporting interoperability of services independent of operating systems, programming languages or any other technology specifics that underlie applications, supporting the implementation and delivery of services

2.2 Distributed and Grid Computing

Distributed computing is a technique where computing tasks are processed by a collection of networked computers and therefore can be seen a cluster of computers Grid computing creates a fast virtual computer from a network of computers by using their idle cycles.Even though grid computing is an important successor to distributed computing, the computing environments are essentially different. For distributed computing, resources are homogeneous and are reserved, leading to guaranteed processing capacity. On the other hand, grid environments are highly unpredictable. The computers are heterogeneous, their capacities are typically unknown and changing over time, and they may connect and disconnect from the grid at any time.

2.3 Virtualization

Virtualization is a technology that abstracts away the details of physical hardware and provides virtualized resources for high-level applications. A virtualized server is commonly
called a virtual machine (VM). Virtualization forms the foundation of cloud computing, as it provides the capability of pooling computing resources from clusters of servers and dynamically assigning or reassigning virtual resources to applications on-demand.

2.4 Autonomic Computing: Originally coined by IBM in 2001, autonomic computing aims at building computing systems capable of self-management, i.e. reacting to internal and external observations without human intervention. The goal of autonomic computing is to overcome the management complexity of today’s computer systems. Although cloud computing exhibits certain autonomic features such as automatic resource provisioning, its objective is to lower the resource cost rather than to reduce system complexity.

In summary, cloud computing leverages virtualization technology to achieve the goal of providing computing resources as a utility. It shares certain aspects with grid computing and autonomic computing but differs from them in other aspects. Therefore, it offers unique benefits and imposes distinctive challenges to meet its requirements.

2.5 Utility Computing
Utility computing proposes to allow clients to buy computing capacity as they do electricity – just by plugging in or out. For users of this utility computing principle, the cost is variable and based on the actual capacity they require, rather than a fixed cost for a capacity they only use during peak periods. Users can get the capacity they need whenever they need it, without expending resources and effort to frequently monitor and upgrade capacity.

3 CLOUD COMPUTING ARCHITECTURE
This section describes the architectural, business and various operation models of cloud computing.

3.1 A layered model of cloud computing
Generally speaking, the architecture of a cloud computing environment can be divided into 4 layers: the hardware/ datacenter layer, the infrastructure layer, the platform layer and the application layer, as shown in Fig. 1. We describe each of them in detail

The Hardware layer: This layer is responsible for managing the physical resources of the cloud, including physical servers, routers, switches, power and cooling systems. In practice, the hardware layer is typically implemented in data centers. A data center usually contains thousands of servers that are organized in racks and interconnected through switches, routers or other fabrics. Typical issues at hardware layer include hardware configuration, faulttolerance, traffic management, power and cooling resource management

The Infrastructure layer: Also known as the virtualization layer, the infrastructure layer creates a pool of storage and computing resources by partitioning the physical resources using virtualization technologies such as Xen [55], KVM [30] and VMware [52]. The infrastructure layer is an essential component of cloud computing, since many key features, such as dynamic resource assignment, are only made available through virtualization technologies.

The Platform layer: Built on top of the infrastructure layer, the platform layer consists of operating systems and application frameworks. The purpose of the platform layer is to minimize the burden of deploying applications directly into VM containers. For example, Google App Engine operates at the platform layer to provide API support for implementing storage, database and business logic of typical web applications.

The Application layer: At the highest level of the hierarchy, the application layer consists of the actual cloud applications. Different from traditional applications, cloud applications can leverage the automatic-scaling feature to achieve better performance, availability and lower operating cost. Compared to traditional service hosting environments such as dedicated server farms, the architecture of cloud computing is more modular. Each layer is loosely coupled with the layers above and below, allowing each layer to evolve separately. This is similar to the design of the OSI model for network protocols. The architectural modularity allows cloud computing to support a wide range of application requirements while reducing management and maintenance overhead.

4. SERVICE MODELS

4.1 Infrastructure as a service (IAAS)
4.2 Platform as a service (PAAS)
4.3 Software as a Service (SAAS)

4.1 Infrastructure as a Service

Infrastructure as a Service (IaaS) is the delivery of hardware (server, storage and network), and associated software (operating systems, virtualization technology, file system), as a service. Through virtualization, infrastructure providers are able to split, assign and dynamically resize these resources to build ad-hoc systems as demanded by customers. The IaaS provider does very little management other than keep the data center operational. The users must deploy and manage the software themselves. The main difference is that data is not stored on the servers which an organization owns by itself, but stored on servers in datacenters of third party service providers.

4.2 Platform as a Service (PAAS)

Platforms are an abstraction layer between the software applications (SaaS) and the virtualized infrastructure (IaaS). In a Cloud platform arrangement the provider operates the whole computing platform or operating system for the customer which is accessed over the Web. This model makes all of the facilities and services required to support the complete life cycle of building and delivering applications entirely available as an integrated suite from the internet. Therefore platform as a service means that Cloud providers supply the software platform where systems run on. Developers can write their applications according to the specifications of a particular platform without needing to worry about the underlying hardware infrastructure (IaaS). Developers upload their application code to a platform, which then typically manages the automatic up scaling when the usage of the application grows.

4.3 Software as a Service (SAAS)

Software as a Service solutions are at the top end of the Cloud Computing stack and they provide end users with an integrated service comprising hardware, development platforms, and applications. In the software as a service model, the application, or service, is deployed from a centralized data center across the web, providing access and use on a recurring fee basis. SaaS is accessed by a web-portal and is often based on the concept of utility computing, which means the organizations pays for the resources they are actually using. The next table provides an overview of the different Cloud service models as described above and states out the main current commercial vendors in the market environment.

5. DEPLOYMENTS MODELS

Clouds are often defined as public (i.e. external), private (i.e. internal) or hybrid (a combination thereof). There are three types of clouds defined in deployment models.

5.1 Public cloud

5.2 Private cloud

5.3 Hybrid cloud
6.2 Resource allocator
When an organization decides to outsource one or more IT services (infrastructure, platform or software) into a Cloud environment one should expect some coordination between the Cloud service user and provider. The services developed by the service providers are probably not one-to-one applicable into an organization.

The resource allocator forms the link between the developer and the user. Appendix A provides a schematic overview of this architecture, which states out the role of the resource allocator in relation with the Cloud service user and provider. The resource management structure has to be changed in order to regulate the supply and demand of Cloud resources. System-centric is a traditional approach to resource management that attempts to optimize system-wide measure of performance and is commonly used in managing resources in single administrative domains.

6.3 Cloud Service User
1. End consumers, who mainly use the service of the SaaS layer over a Web browser and basic offerings of the IaaS layer as for example storage for data resulting from the usage of the SaaS layer.
2. Business customers that might access all three layers: the IaaS layer in order to enhance the own infrastructure with additional resources on demand, the PaaS layer in order to be able to run own applications in a Cloud and eventually the SaaS layer in order to take advantage of available applications offered as a service.
3. Developers and Independent Software Vendors that develop applications that is supposed to be offered over the SaaS layer of a Cloud. Typically, they directly access the PaaS layer, and through the PaaS layer indirectly access the IaaS layer, and are present on the SaaS layer with their application

7. RESEARCH CHALLENGES
Although cloud computing has been widely adopted by the industry, the research on cloud computing is still at an early stage. Many existing issues have not been fully addressed, while new challenges keep emerging from industry applications. In this section, we summarize some of the challenging research issues in cloud computing.

7.1 Automated service provisioning
One of the key features of cloud computing is the capability of acquiring and releasing resources on-demand. The objective of a service provider in this case is to allocate and de-allocate resources from the cloud to satisfy its service level objectives (SLOs), while minimizing its operational cost. However, it is not obvious how a service provider can achieve this objective. In particular, it is not easy to determine how to map SLOs such as QoS requirements to low-level resource requirement such as CPU and memory requirements.

7.2 Virtual machine migration
Virtualization can provide significant benefits in cloud computing by enabling virtual machine migration to balance load across the data center. In addition, virtual machine migration enables robust and highly responsive provisioning in data centers. Virtual machine migration has evolved from process migration technique. More recently, Xen and VMware have implemented “live” migration of VMs that involves extremely short downtimes ranging from tens of milliseconds to a second. Clark et al. pointed out that migrating an entire OS and all of its applications as one unit allows to avoid many of the difficulties faced by processlevel migration approaches, and analyzed the benefits of live migration of VMs.

The major benefits of VM migration is to avoid hotspots; however, this is not straightforward. Currently, detecting workload hotspots and initiating a migration lacks the agility to respond to sudden workload changes. Moreover, the innamery state should be transferred consistently and efficiently, with integrated consideration of resources for applications and physical servers.

7.3 Server consolidation
Server consolidation is an effective approach to maximize resource utilization while minimizing energy consumption in a cloud computing environment. Live VM migration technology is often used to consolidate VMs residing on multiple under-utilized servers onto a single server, so that the remaining servers can be set to an energy-saving state. The problem of optimally consolidating servers in a data center is often formulated as a variant of the vector bin-packing problem, which is an NP-hard optimization problem.

7.4 Energy management
Improving energy efficiency is another major issue in cloud computing. It has been estimated that the cost of powering and cooling accounts for 53% of the total operational expenditure of data centers. In 2006, data centers in the US consumed more than 1.5% of the total energy generated in that year, and the percentage is projected to grow 18% annually. Hence infrastructure providers are under enormous pressure to reduce energy consumption. The goal is not only to cut down energy cost in data centers, but also to meet government regulations and environmental standards.

7.5 Traffic management and analysis
Analysis of data traffic is important for today’s data centers. However, there are several challenges for existing traffic measurement and analysis methods in Internet Service Providers (ISPs) networks and enterprise to extend to data centers. Firstly, the density of links is much higher than that in ISP or enterprise networks, which makes the worstcase scenario for existing methods. Secondly, most existing methods can compute traffic matrices between a few hundreds end hosts, but even a modular data center can have several thousand servers. Finally, existing methods usually assume some flow patterns that are reasonable in Internet and enterprises networks, but the applications deployed on data centers, such as MapReduce jobs, significantly change the traffic pattern.

7.6 Data security
Data security is another important research topic in cloud computing. Since service providers typically do not have access to the physical security system of data centers, they must rely on the infrastructure provider to achieve full data security. Even for a virtual private cloud, the service provider can only specify the security setting remotely, without knowing whether it is fully implemented.

7.7 Software frameworks
Cloud computing provides a compelling platform for hosting large-scale data-intensive applications. Typically, these applications leverage MapReduce frameworks such as Hadoop for scalable and fault-tolerant data processing. Recent work has shown that the performance and resource consumption of a MapReduce job is highly dependent on the type of the application.

The key challenges include performance modeling of Hadoop jobs (either online or offline) and adaptive scheduling in dynamic conditions. Another related approach argues for making MapReduce frameworks energy-aware. The essential idea of this approach is to turn Hadoop node into sleep mode when it has finished its job while waiting for new assignments. To do so, both Hadoop and HDFS must be made energy-aware. Furthermore, there is often a trade-off between performance and energy-awareness. Depending on the objective, finding a desirable trade-off point is still an unexplored research topic.

7.8 Storage technologies and data management
Software frameworks such as MapReduce and its various implementations such as Hadoop and Dryad are designed for distributed processing of data-intensive tasks. As mentioned previously, these frameworks typically operate on Internet-scale file systems such as GFS and HDFS. These file systems are different from traditional distributed file systems in their storage structure, access pattern and application programming interface. In particular, they do not implement the standard POSIX interface, and therefore introduce compatibility issues with legacy file systems and applications. Several research efforts have studied this problem. For instance, the work in proposed a method for supporting the MapReduce framework using cluster file systems such as IBM’s GPFS. Patil et al. proposed new API primitives for scalable and concurrent data access.

7.9 Novel cloud architectures

Currently, most of the commercial clouds are implemented in large data centers and operated in a centralized fashion. Although this design achieves economy-of-scale and high manageability, it also comes with its limitations such high energy expense and high initial investment for constructing data centers. Recent work suggests that smallsize data centers can be more advantageous than big data.

8. CONCLUSIONS

“Cloud” computing builds on decades of research in virtualization, distributed computing, utility computing, and, more recently, networking, web and software services. It implies a service-oriented architecture, reduced information technology overhead for the end-user, great flexibility, reduced total cost of ownership, on demand services and many other things. This paper discusses the concept of “cloud” computing and technologies with service and Deployment models in description Our experience with cloud computing technology is excellent and we are working on additional functionalities and features that will make it even more suitable for cloud framework construction.

9. REFERENCES