Optical Networking for Cloud/Grid/Data Centers

Yang Zhao
National University of Singapore
A0120575@u.nus.edu

Keywords: Cloud computing, Grid computing, Optical networks, Virtualization.

Abstract. Grid and cloud computing play the crucial role in networks in supporting today’s application. Data center nowadays has become a promising infrastructure for data storage, supporting Qos (quality of service), and even defenses against security attacks. In this paper, we address the role of network in cloud computing environment and a data center (DC) as a service architecture for future cloud computing.

Introduction

The role and scale of networks has been increasingly growing, as illustrated in Fig.1. In the late 1950s through 1960s, there were essentially standalone systems. In the 1970s, Metcalfe’s invention of Ethernet increases the value of the whole set of end points is now incontestable. In the 1990s, it is evident that local computational resources cannot keep up with the demand generated by users/applications, the increasing networking capabilities and the high capabilities massive data volumes led to the inception of computational grid. Inspired by the success of grid paradigm in scientific circles, the idea of cloud computing arose in the 2000s, building on the seminal idea of “computation provided as a public utility”. The evolution to a widespread adoption of such “utility computing” initially manifested in client-server-based architectures in local networks, and continued with remotely hosted parts in web-based solutions (Fig 1).[1]

Cloud computing begins as an essential component of the enterprise IT infrastructure and, consequently, one of the fastest growing business opportunities for Internet service providers and telecom operators. A cloud infrastructure allows realization of an end-to-end cloud service provisioning architecture that coordinates the DC (Data Center) infrastructure with the required optical network connectivity services and provides the converged services to users in the form of DC as a server (DCaaS).

In this paper, we just give some basic introduction of data center, grid and cloud. Also give some easy-going descriptions of current application of network technology on them.

The rest of this paper is organized as follow. First, we get a basic understanding of the concept of data center, grid and cloud computing. Second, I sketch optical network technology as well as its
application and virtualization. Then, I will analyze current bottleneck of realization, also give a description of challenges and future directions.

**Data Center as a Service Architecture**

**Data center.** A DC is a facility that consists of servers (physical machine), storage, and network devices, power distribution systems, cooling systems.

A data center network is the communication infrastructure used in a data center, and is described by the network topology, routing/switching equipment, and the used protocols.

**Data center virtualization.** A virtualized Data Center is a data center where some or all of the hardware are virtualized. Typically, a physical hardware is virtualized using software or firmware that divides the equipment into multiple isolated and independent virtual instances. A Virtual Data Center is a collection of virtual resources connected via virtual links, i.e. using virtualization, multiple coexisting but isolated virtual DCs plus the virtual network can be composed over the same physical infrastructure.

**Network virtualization.** A virtual network is a set of virtual networking resources: virtual nodes and virtual links; thus virtual network is a part of virtual data center. A network virtualization level is one of the layers of the network stack in which the virtualization is introduced. Both network virtualization and data center virtualization rely on virtualization techniques to partition available resources and share them among different users, however, they differ in various aspects. VNs (virtual networks) mostly consist of packet forwarding elements, virtualized data center networks involve different types of nodes including servers, routers, switches, and storage nodes. Another key difference between data center networks and virtual network is the number of nodes. What is more, data center networks are built using topologies like conventional tree, fat tree, or well defined properties, allowing to develop embedding algorithms optimized Clos topologies.

In the DCaaS (Data Center As a Service Architecture), multiple geographically distributed DCs interconnected by high-performance optical networks apply the infrastructure as a service model in order to create virtual data centers or VIs and offer them as services to large enterprise users. By using the benefits of the virtualization, multiple VDCs can be coexisting and running at the same time over the same physical infrastructure without interfacing each other. As shown in the Fig2, DCaaS comprises three layers: 1) the physical infrastructure layer 2) the infrastructure layers 3) the DCaaS layer. The elements of three layers can be clearly seen in the Fig2.

![Fig 2 Reference model of DCaaS architecture](From: Role of optical network virtualization in cloud computing)

In the DCaaS architecture, the virtualization of the IT resources of DCs has reached a mature and commercial stage. However, the virtualization of the optical network and the coordinated virtualization optical network is still under initial study. [2]
Clouds and Grids

Clouds and grids imply the need for efficient coordination of networking, storage, and computing resources. Grid was introduced mainly because the increasing need of e-Science applications. Grid, also called grid computing is a new network computing platform after the emergence of Internet, is a computing mode of merging scattered resources and realizing resources sharing and cooperative working. Grid—the word comes from Electric Power Grid, the original intention of grid is to make the user as convenient as using electric power infrastructure when he/she in the use of grid.

The grid has important application value in scientific research, e-government, e-business and personal entertainment and so on. Grid computing caused people’s great attention after it appeared. People connect a cluster of computers into a local grid. It is like connecting dozens of computer to become a supercomputer, computing power has greatly enhanced certainly. The local computing grid is mainly used in some scientific research. These calculations if running on a single personal computer, often it takes months, yet running in a grid, it takes only one or two days.

The characteristics of a grid are concisely summarized in three points: 1) coordinates resources that are not subject to centralize control 2) uses standard, open, general purpose protocols and interfaces 3) delivers nontrivial qualities of service.

However, cloud’s development mainly follows with the increasing commercial applications. Why call it cloud computing is that users do not know (or invisible) resources or organization beneath the cloud. Part of resources were taken out from the clouds when necessary, after using, sent them back to the clouds (this concept is the so-called "on-demand" cloud). In other words, the "cloud" is a pool of resources (i.e., a group of machine and Webpage service), it can make the shared resource provided by system is scalable, on-demand use and lease payment.

The main characteristics of cloud are user friendliness (e.g. businesses and consumers), virtualization, scalability, pay-per-use model, and service level agreement. Compared to grids, cloud virtualization allows businesses to share resources in a safe way, the response time and latency requirements of business/consumer applications can be more stringent. What is more, cloud virtualization allows migration, which has avoided failure in terms of performance. Monitoring in cloud is very challenging because users can not install or run their own monitoring infrastructures. But grids allow users to access and browse resources at various sites by identity delegation.

The application and system of cloud computing is based on large-scale inexpensive server clusters, through cooperative construction of infrastructure and upper application to achieve maximum efficiency of the utilization of hardware resources, as well as by software to tolerate more node errors in order to accomplish the scalability and reliability of the distributed computing system.

Grid computing emphasizes on a virtual organization composed of many agencies, different server of multiple organizations to form a virtual organization provides a powerful computational resources for the user. While cloud computing using virtual machines (virtual server) to polymerize and form homogeneous service, pay more attention to the share of distributed computational resources within an organization. In some aspects such as business mode, scheduling, resource allocation, whether providing service and form and so on, their difference is obvious.

Optical Technology

Optical networking plays an important role in the realization of grid and cloud computing systems. Optical transmission is generally accepted as the most cost-effective way to realize connections in the network. The technique which can transfer huge data volumes with low latency actually has made optical network connect data centers that provide computing and storage services in grid and cloud computing network. The technique named wavelength-division multiplexing (WDM) permits optical end-to-end connections over a specific frequency or wavelength. It can achieve tremendous scalability because of the high data rates offered. But some users only need low data rates causing wastes bandwidth and reduces resources utilization.
Apart from transmission technology, optical switching techniques also attempt to bridge the gap between optical bandwidth and user demands. For example, one of the optical circuit switching solutions – WDM, it reserves bandwidth exclusively for communication between source and destination. That may result in the inefficient use of bandwidth if the traffic demand does not match the full capacity of a wavelength. What is more, a two-way signaling operation to allocate bandwidth and configure switching devices takes amount of time to complete. Besides these issues, successful examples of WDM-based grids include the GineGrid and TransLight projects, which utilize optical circuits to build grid computing networks.

Optical technology is an essential component to interconnect data centers on a global scale, for example, Amazon’s EC2, Google’s cloud offerings.

Virtualization

Virtualization acts as the fundamental drive for the cloud computing, which is the process of either partitioning a single physical resource in multiple virtual resources, or aggregation of multiple physical resources in a single virtual resource. For example, a network resource can be shared among multiple users without interfering with each other when operating within their own virtual infrastructure.

Network virtualization, in the cloud environment, plays an important role in interconnecting geographically distributed virtual IT resources provided by remote DCs. Specifically on optical links, has been studied and implemented in the form of virtual private networks, which connect a number of known end points over a dedicated communications infrastructures. For instance, in optical layer, VPNs are provisioned to establish connectivity between data centers, such as creating a high performance public or private cloud networks. Network resources virtualization, which is also the essential technology to realize the future Internet. Because it allows private grid or cloud infrastructure can be customized according to different requirements.

IT resources consist of multiple components such as a processing units, storage units, and working memory. Virtualization of computer systems results in a virtual machine, which provides all the capabilities of the host resources. Virtualization of IT resources can be partly explained by the technology’s widespread adoption and commercial success.

In recent, due to the popularity of the grid and cloud computing concepts, combined virtualization of both networking and IT resources has gained widespread attention. Which means to introduce a logical infrastructure composition layer that manages the physical infrastructure consisting of both network and IT resources, and exposes these as virtual resources in a genetic way. The main advantage of coordinated IT and optical network infrastructure virtualization is that the virtual nodes and the virtual links in a VDC can be mapped at a single stage.

Challenges, Opportunities, Future directions

Major challenges are related to the integration of cloud and grid technologies with optical networks. Which means to share the physical resources, and keep isolation among them at the same time. Besides, network design and optimization questions are still need to be addressed, as well as challenging provisioning and scheduling problems.

Second, optical clouds play a significant role in reducing greenhouse gas emissions. Compared with traditional electronic technologies, optical networks provide some key advantages in terms of power consumptions. Because of balancing of network, storage, and processing network evaluating the cloud computing requires. Elasticity and virtualization has developed to promise the accurate matching of demand. But the more revolutionary way is that locating data centers close to renewable energy sources and moving cloud jobs from one data center to another depending on the instantaneous renewable energy production. Which is still under investigating.

Third, also in the grid/cloud paradigm itself, some important questions remain unanswered. For example, the growing interest of e-Science users in cloud solutions to serve their application needs, and references, the integration of grid and cloud paradigms poses a promising challenge to address.
Grids and cloud can continue to coexist, but grids serve to the e-Science applications whereas cloud aims to serve the business and consumers. When running distributed e-Science applications in current commercial clouds networks seem have met the bottleneck. So the network requires enhanced support for Qos, monitoring, federation of different organizations.

What is more, in the future, explorations regarding data center network virtualization may concern more about edge data centers which will complement remote data centers in providing high quality online services at low cost. That is because recent protocols put forward to building small scale data centers to get close to end users. Also there are benefits in building visualized data centers, they can provide better service in some applications. The construction cost of edge data center is lower compared to large remote data centers.

Another trend may be virtual data centers embedding, which means to accommodate a large number of virtual data centers depending on an efficient mapping of virtual resources to physical ones. But challenges under this are, for example, there are quite a lot of resources besides physical servers should be considered to be virtualized. The dynamic resource demand requires embedding of VDCs is also subject to change. Besides how to reduce the energy cost in using data center still requires us to make effort, for it is very difficult to optimize both VMs and VNs for saving energy. Another concern is fault-tolerance, which is very important in minimizing performance penalty. Finally, how to satisfy users’ demand to deploy VDCs across data centers from multiple regions becomes a very challenging problem. Because it requires embedding VDCs across multiple administrative domains.

Conclusion

The role of the optical network virtualization in the cloud computing and DC environment is discussed in this paper. What is more, the concept and unique features of optical network virtualization for the DCaaS architecture is mentioned as well. Taking into account the unique characteristics of the optical network layer. In order to adapt to cloud computing environments better, coordinated virtualization of both optical network and IT resources is proposed and investigated. During the evolution of the computing paradigms, the role of the network becomes increasingly important in grids and recent clouds. By now we have introduced some novel applications (e-Science application and business and consumer-oriented applications) that give rise to this evolution, and also identify their requirements.

References

54, no. 5, 8 April 2010, Pages 862-876.

