ISSN : 0976-8491 (Online) | ISSN : 2229-4333 (Print)

Global Data Transfers through Cloud and Grid Computing

Rohit Seth
Dept. of Computer Applications, Dr IT Group, Banur, Punjab, India

Abstract
Networks form the global communications infrastructure that assistsmillions of phone calls each day and, even more significantly, the enormous global data transfers, chiefly resulting via Internet. These networks are operated by hundreds of companies and placed them on top of physical infrastructure(cabling and switches).

These infrastructures are usually not owned by the network operator but have to follow different regulations in each country and probably traverse several billing domains at any given point-to-point connection. This global integrated system operates with extremely low down time and transparently to the end users. This has been achieved through decades of development and standardization of interfaces, a process which has now become well established in the telecommunications industry.

The latest evolution of the global communications networks, the Next Generation Network (NGN), is designed to support converged fixed and wireless networks carrying both voice and data traffic. Furthermore, these future networks are beneficial to customers and the network providers both. Increased flexibility around network-level services has also opened the door to third party services built on top of the NGN infrastructure. With growing interest in grid computing over the past several years and more recently, cloud computing, a major question that needs to be answered is how these technologies, concepts, and capabilities can be incorporated into NGN.

Keywords
Data, Traffic, Communications, NGN, Grid, Cloud, Computing, Network, Architecture

I. Introduction to Grid and Cloud Computing
Grid and cloud computing systems would benefit from enhanced capabilities of NGN, the global reach of existing communications networks, and the stability of carrier-grade networks. Interoperability has been one of the key contributors to widespread commercial success of technologies used in the telecommunications sector,due to the interconnected nature of networks, and the plethora of network operators. Interoperability fosters diversity as well as competition in a market. Vendors can achieve interoperability of their products when they agree and implement a common set of open standards. The value of standardization has also been recognized by the grid community, and is predominantly championed by the Open Grid Forum forGrid-specific standards. Standardization, however, does not necessarily lead to interoperability. Standards have to be engineered for interoperability.

II. Standards for Grid and Cloud Domains
The monumental range of organizations involved with one or more of grid, cloud, and NGN technology each have their own priorities. However, operational systems have been designed with the range of priorities which has resulted in competing architectures and interfaces. Although NGN does not yet exist as an integrated global telecommunications platform, there is a coordinated effort to develop the suite of standards to cover a high level NGN architecture. In contrast, grid computing offers a few high level conceptual models. This wide range of high level applications connected to a wide range of heterogeneous lowlevel resources via a limited number of intermediate standard interfaces. In addition, there are a few concrete architectural models for grid infrastructures. These concrete models have Grid and Cloud Computing: Opportunities for Integration with the next generation Network a distinct disconnection: either they present an architecture which is not or only partially implemented in any operational grid, or an architecture which describes a particular grid infrastructure with limited references to standards or interfaces. In the cloud domain, there is currently a prevalence of independent services with minimal interest in interoperability or consideration of standards. While this is in the process of changing, there is currently no sufficient activity in this area to report on. The Networked European Software and Services Initiative-Grid (NESSI-Grid) review considered the impact of grid technology on business IT infrastructure, while several projects have considered the application of grid computing to e Health. The recent popularity of cloud computing also demonstrates the industry benefits of shared, distributed computing infrastructure.

Cloud computing is a network based computing that facilitates shared computer processing resources and data to computers and other devices on demand. It is a model for introducing ubiquitous, on-demand access to a shared pool of configurable computing resources (e.g., computer networks, servers, storage, applications and services) which can be rapidly provisioned and released with minimal management effort. Cloud computing and storage solutions provide users and enterprises with various capabilities to store and process their data that may be located far from the user–ranging in distance from across a city to across the world. Cloud computing is reliable on sharing of resources to achieve coherence that is similar to a utility over an electricity network. Numerous of experts claim that cloud computing allows companies to avoid infrastructure costs. Also, it enables organizations to focus on their core businesses instead of spending time and money on computer infrastructure. Proponents also claim that cloud computing allows enterprises to get their applications up and running faster, with improved manageability and less maintenance, and enables Information Technology (IT) teams to more rapidly adjust resources to meet fluctuating and unpredictable business demand. Cloud providers typically use a "pay as you go" model. This will lead to unexpectedly high charges if administrators do not adapt to the cloud pricing model.

In 2009, the availability of high-capacity networks, low-cost computers and storage devices as well as the widespread adoption of hardware and utility computing led to a growth in cloud computing. Companies can scale up as computing needs increase and then scale down again as demands decrease. In 2013, it was reported that cloud computing had become a highly demanded service or utility due to the advantages of high computing power, cheap cost of services, high performance, scalability, accessibility as well as availability. Some cloud vendors are experiencing growth rates of 50% per year,but being still in a stage of infancy, it has pitfalls that need to be addressed to make cloud computing services more reliable and user friendly.

Grid computing is the combination of computer resources from different locations to approach a common goal. The grid can be thought of as a system with non-interactive workloads that involve a large number of files. Grid computing is differentiated from conventional high performance computing systems such as computing in that grid computers have each node set to perform a different purpose or application. Grid computers also tend to be more opposite and geographically dispersed (thus not physically coupled) than other computers. Although a single grid can be dedicated to a particular application, commonly a grid is used for a variety of purposes. Grids are often constructed with general-purpose grid software libraries. Grid sizes can be quite large.

Grids are a form of computing whereby a "virtual computer" is composed of many networked computers acting together to perform large tasks. For certain applications, distributed or grid computing can be seen as a special type of parallel computing that relies on complete computers connected to a network which is private or public by a conventional interface.

III. Architecture of Cloud Computing

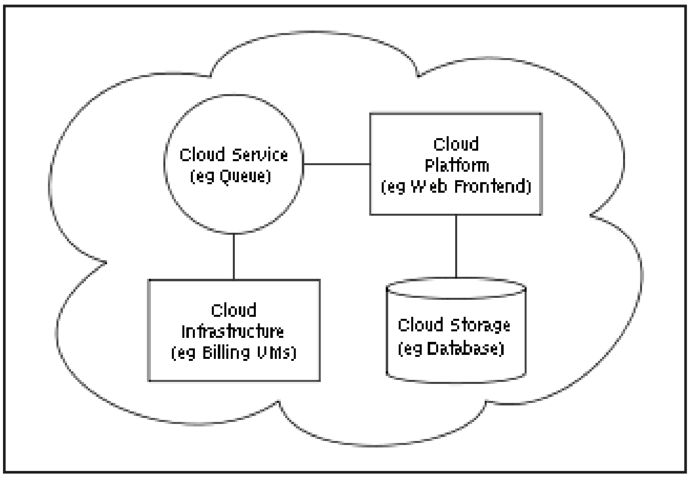

Fig. 1: Cloud Computing Sample Architecture

Cloud architecture, the architecture of software involved in the delivery of cloud computing, typically involves multiple cloud components communicating with each other over a loose coupling mechanism such as a messaging queue. Elastic provision implies intelligence in the use of tight or loose coupling as applied to mechanisms such as these and others.

Cloud engineering is the application of engineering disciplines to cloud computing. It brings a systematic approach to the high-level concerns of commercialization, standardization, and governance in conceiving, developing, operating and maintaining cloud computing systems.

IV. Security and Privacy

Cloud computing poses privacy concerns because the service provider can access the data that is in the cloud at any time. It could accidentally or deliberately alter or even delete information. Many cloud providers can share information with third parties if necessary for purposes of law and order even without a warrant. That is permitted in their privacy policies, which users must agree to before they start using cloud services. Solutions to privacy include policy and legislation as well as end users' choices for how data is stored. Users can encrypt data that is processed or stored within the cloud to prevent unauthorized access.

According to the cloud security alliance, the top three threats in the cloud are Insecure Interfaces and API's, Data Loss & Leakage, and Hardware Failure—which accounted for 29%, 25% and 10% of all cloud security outages respectively. Together, these form shared technology vulnerabilities. In a cloud provider platform being shared by different users there may be a possibility that information belonging to different customers resides on same data server. Therefore, Information leakage may arise by mistake when information for one customer is given to other. Chief technology officer at Security, said that hackers are spending substantial time and effort looking for ways to penetrate the cloud. "There are some real Achilles' heels in the cloud infrastructure that are making big holes for the bad guys to get into". Because data from hundreds or thousands of companies can be stored on large cloud servers, hackers can theoretically gain control of huge stores of information through a single attack—a process he called "hyper jacking". Some examples of this include the Drop box security breach, and iCloud 2014 leak. Dropbox had been breached in October 2014, having over 7 million of its user passwords stolen by hackers in an effort to get monetary value from it by Bitcoins (BTC). By having these passwords, they are able to read private data as well as have this data be indexed by search engines (making the information public).

There is the problem of legal ownership of the data (If a user stores some data in the cloud, can the cloud provider profit from it?). Many Terms of Service agreements are silent on the question of ownership. Physical control of the computer equipment (private cloud) is more secure than having the equipment off site and under someone else's control (public cloud). This delivers great incentive to public cloud computing service providers to prioritize building and maintaining strong management of secure services. Some small businesses that don't have expertise in IT security could find that it's more secure for them to use a public cloud. There is the risk that end users do not understand the issues involved when signing on to a cloud service (persons sometimes don't read the many pages of the terms of service agreement, and just click "Accept" without reading). This is important now that cloud computing is becoming popular and required for some services to work. Fundamentally, private cloud is seen as more secure with higher levels of control for the owner, however public cloud is seen to be more flexible and requires less time and money investment from the user.

V. Architecture of Grid Computing

A new architecture model and technology has been developed for the establishment and management of cross-organizational resource sharing. This new architecture, called grid Architecture, identifies the basic components of a . The grid architecture defines the purpose and functions of its components, while indicating how these components interact with one another. The main focus of the architecture is on interoperability among resource providers and users in order to establish the sharing relationships. This interoperability, in turn, necessitates common protocols at each layer of the architectural model, which leads to the definition of a grid protocol architecture as shown in figure below. This protocol architecture defines common mechanisms, interfaces, schema, and protocols at each layer, by which users and resources can negotiate, establish, manage, and share resources. Figure shows the component layers of the grid architecture and the capabilities of each layer. Each layer shares the behavior of the underlying component layers. The following describes the core features of each of these component layers, starting from the bottom of the stack and moving upward.

A. Fabric Layer

The fabric layer defines the interface to local resources, which may be shared. This Includes computational resources, data storage, networks, catalogs, software modules, and other system resources.

B. Connectivity Layer

The connectivity layer defines the basic communication and authentication protocols required for grid-specific networking-service transactions.

C. Resource Layer

This layer uses the communication and security protocols (defined by the connectivity layer) to control secure negotiation, initiation, monitoring, accounting, and payment for the sharing of functions of individual resources. The resource layer calls the fabric layer functions to access and control local resources. This layer only handles individual resources, ignoring global states and atomic actions across the resource collection pool, which are the responsibility of the collective layer.

D. Collective Layer

While the resource layer manages an individual resource, the collective layer is responsible for all global resource management and interaction with collections of resources. This protocol layer implements a wide variety of sharing behaviors using a small number of resource-layer and connectivity-layer protocols.

E. Application Layer

The application layer enables the use of resources in a grid environment through various collaboration and resource access protocols. Thus far, our discussions have focused on the grid problem in the context of a virtual organization and the proposed grid computing architecture as a suggested solution to this problem. This architecture is designed for controlled resource sharing with improved interoperability among participants. In contrast, emerging architectures help the earlier-defined grid architecture quickly adapt to a wider (and strategically important) technology domain.

Table 1: Differentiating Cloud Computing and Grid Computing

	Grid computing	Cloud computing
What?	Grids enable access to shared computing power and storage capacity from your desktop	Clouds enable access to leased computing power and storage capacity from your desktop
Who provides the service?	Research institutes and universities federate their services around the world through projects such as EGI-Inspires and the European Grid Infrastructure.	Large individual companies e.g. Amazon and Microsoft and at a smaller scale, institutes and organizations deploying open source software such as Open Slate, Eucalyptus and Open Nebula.
Who uses the service?	Research collaborations, called “Virtual Organizations”, which bring together researchers around the world working in the same field.	Small to medium commercial businesses or researchers with generic IT needs
Who pays for the service?	Governments - providers and users are usually publicly funded research organizations, for example through National Grid Initiatives.	The cloud provider pays for the computing resources; the user pays to use them
Where are the computing resources?	In computing centers distributed across different sites, countries and continents.	The cloud provider’s private data centers which are often centralized in a few locations with excellent network connections and cheap electrical power.
Why use them?	- You don`t need to buy or maintain your own large computer centre - You can complete more work more quickly and tackle more difficult problems. - You can share data with your distributed team in a secure way.	- You don`t need to buy or maintain your own personal computer centre - You can quickly access extra resources during peak work periods
What are they useful for?	Grids were designed to handle large sets of limited duration jobs that produce or use large quantities of data (e.g. the LHC and life sciences)	Clouds best support long term services and longer running jobs (E.g. facebook.com)
How do they work?	Grids are an open source technology. Resource users and providers alike can understand and contribute to the management of their grid	Clouds are a proprietary technology. Only the resource provider knows exactly how their cloud manages data, job queues, security requirements and so on.

Benefits?	- Collaboration: grid offers a federated platform for distributed and collective work. - Ownership : resource providers maintain ownership of the resources they contribute to the grid - Transparency: the technologies used are open source, encouraging trust and transparency. - Resilience: grids are located at multiple sites, reducing the risk in case of a failure at one site that removes significant resources from the infrastructure.	- Flexibility: users can quickly outsource peaks of activity without long term commitment - Reliability: provider has financial incentive to guarantee service availability (Amazon, for example, can provide user rebates if availability drops below 99.9%) - Ease of use: relatively quick and easy for non-expert users to get started but setting up sophisticated virtual machines to support complex applications is more difficult.
Drawbacks?	- Reliability: grids rely on distributed services maintained by distributed staff, often resulting in inconsistency in reliability across individual sites, although the service itself is always available. - Complexity: grids are complicated to build and use, and currently users require some level expertise. - Commercial: grids are generally only available for not-for-profit work, and for proof of concept in the commercial sphere	- Generality: clouds do not offer many of the specific high-level services currently provided by grid technology. - Security: users with sensitive data may be reluctant to entrust it to external providers or to providers outside their borders. - Opacity: the technologies used to guarantee reliability and safety of cloud operations are not made public. - Rigidity: the cloud is generally located at a single site, which increases risk of complete cloud failure. - Provider lock-in: there's a risk of being locked in to services provided by a very small group of suppliers.
When?	The concept of grids was proposed in 1995. The Open science grid (OSG) started in 1995 The EDG (European Data Grid) project began in 2001.	In the late 1990`s Oracle and EMC offered early private cloud solutions. However the term cloud computing didn't gain prominence.

References

[1] Raicu, Y. Zhao, C. Dumitrescu, I. Foster, M. Wilde."Falkon: A Fast and Light weight tasK executiON framework",IEEE/ACM Super Computing 2007.

[2] N. Karonis, B. Toonen, and I. Foster. MPICH G2: A GridEnabled Implementation of the Message Passing In terface.Journal of Parallel and Distributed Computing, 2003.

[3] Introduction to Grid Computing by Bart Jacob Michael Brown

[4] Kentaro FukuiNihar Trived Buyya, Rajkumar; Kris Bubendorfer (2009). Market Oriented Grid and Utility Computing. Wiley. ISBN 978-0-470-28768-2.

[5] Smith, Roger (2005). "Grid Computing: A Brief Technology Analysis" (PDF).

[6] Berstis, Viktors. "Fundamentals of Grid Computing". IBM.

[7] [Online] Available: https://en.wikipedia.org/wiki/Grid_computing

[8] Jacob, Bart; et al. "Enabling Applications for Grid Computing". IBM.

[9] [Online] Available: https://en.wikipedia.org/wiki/Cloud_computing

[10] Rajkumar Buyya Market-Oriented Cloud Computing: Vision, Hype and Reality for Delivering ITServices as Computing Utilities 2008.

[11] Jean-Daniel Cryans, Criteria to Compare Cloud Computing with Current Database Technology 2008.

[12] Christopher Moretti, All-Pairs: An Abstraction for Data-Intensive Cloud Computing IEEE 2008.

[13] http://cloudcomputing.ieee.org/publications/articles

Mr. Rohit Seth earned his Bachelor Degree in Computer Application from Panjab University, Chandigarh, in 2008 the Master of Computer Application from Department of Computer Science and Application, Panjab University, Chandigarh, in 2011. He is Assistant Professor in Department of Computer Science and Applications, Dr It Group of institutes, Banur, Dist-Patiala, India, since August 2011. He is having 5 plus years of experience. His research expertise include cloud computing Wireless technological updated areas, digital signal processing, networking, Smart phones with latest technological enhancements and VoLTE concepts.