

COMPARISON BETWEEN CLOUD AND GRID COMPUTING: REVIEW PAPER

Hosam AlHakami, Hamza Aldabbas, and Tariq Alwada'n

Software Technology Research Laboratory (STRL), De Montfort University,
Leicester, United Kingdom
{hosam, hamza, tariq}@dmu.ac.uk

ABSTRACT

Cloud computing is the most recent announced technology that has been launched on the network world. Clouds are considered as a new generation of Grid computing. Clouds consist of data centres which are owned by the same institute. The homogeneity within each data centre in the infrastructure is the main feature for the cloud computing compared to grid computing. This paper provides a definition for the cloud, it discusses many aspects of Cloud Computing, and describes architectures for the cloud (by looking at Amazon's application (GrepTheWeb)) and how its cost definition differs from that of Grid computing. This paper focuses on comparing Cloud Computing to previous generations such as Grid Computing, by reviewing some Security and Policy Issues in Cloud and Grid Computing. At the end, this paper describes the similarities and differences between the Grid and Cloud approaches.

KEYWORDS

Grid Computing, Cloud Computing, Distributed System,

1. INTRODUCTION

The appearance of cloud computing has been observed very recently as a new promising paradigm that delivers IT services as computing utilities for companies, academic computing and enterprises. It has caused an influence in IT industries. According to IBM [6], a cloud is a pool of virtualized computer resources that hosts a variety of different workloads and allows them to be deployed and scaled-out through the rapid provisioning of virtual or physical machines; supports redundant, self recovering, highly scalable programming models and resource usage monitoring in real time to enable rebalancing of allocations when needed.

Several enterprises are deciding to use virtual data centres to facilitate infrastructure managing and trying to decrease the need of hardware maintenance. This kind of cyber infrastructure shrinks the difficulties that are brought in deployment of services. For example, many users may like to deploy applications in a distributed environment, regardless the cost of this deployment and losing flexibility. The cause of designing the Clouds to provide services to users and providers needs to be made up for sharing their capability and resources. Clouds have almost the same vision of grid computing, nevertheless still there are highly considerable differences.

A grid is a system that has the ability to manage and organize resources and services that are distributed across several control domains, utilize protocols and interfaces and supply high quality of service [27]. Grimshaw et.al. [28] define grid computing as "coordinated resource sharing and problem solving in dynamic, multi-institution virtual organizations".

Cloud Computing has been provided by a number of large organisations such as Amazon, Google, Sun and Yahoo. Individual users through other organisations are also adopted it. Amazon has been a key organisation in the development of cloud computing. Amazon modernised their own internal data centres, which resulted in significant increases in internal efficiency. In 2005 Amazon's cloud computing system called Amazon web services was born. Amazon was the one of the first organisations to provide cloud-computing facility.

This paper is organized as follows. We review current definitions of the Clouds and the Cloud Computing Deployment Models in section 2. In section 3 we describe architectures for the Clouds in detail by showing an example from Amazon organization (a recent popular application (GrepTheWeb)). Section 4 discusses cost models in the Clouds. Section 5 focuses on comparing Cloud paradigm to the Grid and the conclusions are given in last section.

2. CLOUD DEFINITION

2.1. Definition

There is no standard or agreed definition for Cloud. According to Gartner [4], Cloud computing is a style of computing where massively scalable IT-related capabilities are provided as a service across the cyber infrastructure to external users. It has been claimed for some aspects that cloud systems are narrow Grids, in the sense of exposing reduced interfaces.

However, various attempts were made for defining Cloud computing by researchers and practitioners, based on their observation of prerequisites or fundamentals of what the probability and the future of cloud computing are going to be. Many experts have proposed many definitions, which give various concepts about the cloud. The table below indicates some of those experts' definitions:

Table reference as extracted from [14] .

Author/Reference	Year	Definition/Excerpt
M. Klems	2008	you can scale your infrastructure on demand within minutes or even seconds, instead of days or weeks, thereby avoiding under-utilization (idle servers) and over-utilization (blue screen) of in-house resources...
P. Gaw	2008	using the internet to allow people to access technology-enabled services. Those services must be 'massively scalable...
R. Buyya	2008	A 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 agreements established through negotiation between the service provider and

		consumers
R. Cohen	2008	Cloud computing is one of those catch all buzz words that tries to encompass a variety of aspects ranging from deployment, load balancing, provisioning, business model and architecture (like Web2.0). It's the next logical step in software (software 10.0). For me the simplest explanation for Cloud Computing is describing it as, "internet centric software..."
J. Kaplan	2008	a broad array of web-based services aimed at allowing users to obtain a wide range of functional capabilities on a 'pay-as-you-go' basis that previously required tremendous hardware/software investments and professional skills to acquire. Cloud computing is the realization of the earlier ideals of utility computing without the technical complexities or complicated deployment worries...
D. Gourlay	2008	...the next hype-term...building off of the software models that virtualization enabled
D. Edwards	2008	...what is possible when you leverage web-scale infrastructure (application and physical) in an on-demand way...
B. de Haff	2008	...There really are only three types of services that are Cloud based: SaaS, PaaS, and Cloud Computing Platforms. I am not sure being massively scalable is a requirement to fit into any one category.
B. Kepes	2008	...Put simply Cloud Computing is the infrastructural paradigm shift that enables the ascension of SaaS. ... It is a broad array of web-based services aimed at allowing users to obtain a wide range of functional capabilities on a pay-as-you-go basis that previously required tremendous hardware/software investments and professional skills to acquire
K. Sheynkman	2008	Clouds focused on making the hardware layer consumable as on-demand compute and storage capacity. This is an important first

		step, but for companies to harness the power of the Cloud, complete application infrastructure needs to be easily configured, deployed, dynamically-scaled and managed in these virtualized hardware environments
O. Sultan	2008	...In a fully implemented Data Center 3.0 environment, you can decide if an app is run locally (cook at home), in someone else's data center (take-out) and you can change your mind on the fly in case you are short on data center resources (pantry is empty) or you having environmental/facilities issues (too hot to cook). In fact, with automation, a lot of this can be done with policy and real-time triggers...
K. Hartig	2008	..really is accessing resources and services needed to perform functions with dynamically changing needs...is a virtualization of resources that maintains and manages itself.
J. Pritzker	2008	Clouds are vast resource pools with on-demand resource allocation...virtualized ...and priced like utilities
T. Doerksen	2008	Cloud computing is ... the user-friendly version of Grid computing
T. von Eicken	2008	outsourced, pay-as-you-go, on-demand, somewhere in the Internet, etc
M. Sheedan	2008	...'Cloud Pyramid' to help differentiate the various Cloud offerings out there...Top: SaaS; Middle: PaaS; Bottom: IaaS
A. Ricadela	2008	...Cloud Computing projects are more powerful and crash-proof than Grid systems developed even in recent years
I. Wladawsky Berger	2008	...the key thing we want to virtualize or hide from the user is complexity...all that software will be virtualized or hidden from us and taken care of by systems and/or professionals that are somewhere else - out there in The Cloud
B. Martin	2008	Cloud computing encompasses any subscription-based or pay-per-use service that, in real time over the Internet, extends IT's existing capabilities

R. Bragg	2008	The key concept behind the Cloud is Web application... a more developed and reliable Cloud. Many find it's now cheaper to migrate to the Web Cloud than invest in their own server farm ... it is a desktop for people without a computer
G. Gruman and E. Knorr	2008	Cloud is all about: SaaS...utility computing...Web Services... PaaS...Internet integration...commerce platforms....
P. McFedries	2008	Cloud Computing, in which not just our data but even our software resides within the Cloud, and we access everything not only through our PCs but also Cloud-friendly devices, such as smart phones, PDAs... the megacomputer enabled by virtualization and software as a service...This is utility computing powered by massive utility data centers.

Table 1: Cloud Definitions

2.2 Cloud Computing Deployment Model

This is a new model concept that can be divided into the following four famous models (but there might be other models that can be drawn from them) [20]:

- **Public:** Services and resources are reachable to the public by using the internet. This environment emphasises the advantages of rationalization (as a user has the ability to utilize only the needed services and pay only for their use), operational simplicity (as the system is organized and hosted by a third party) and scalability. The main concern in this type of cloud environment is the security; since this environment is accessible to the public and user data in one stage is hosted by a third party.
- **Private:** Services and resources are reachable within a private institute. This environment emphasises the advantages of integration, optimization of hardware deals and scalability. The main concern is the complexity, as this environment is organized and hosted by internal resources. Security is not a main issue compared to the public cloud as the services are reachable only through private and internal networks.
- **Community:** Services and resources of this type are shared by various institutes with a common aim. It may be organized by one of the institutes or a third party [21].
- **Hybrid:** This type combines the methods from the private and public clouds, where resources can be used either in a public or a private cloud environment [22]. The advantages and the concerns are a mixture of the earlier type. Another cloud technology which has become very popular recently is called Green Cloud Computing. Its aim is to reduce resource consumption and yet fulfil quality of service needed and hold the resources switched off as long as possible. "The advantages of such technology are lower heat production and power saving by employing server consolidation and virtualization technologies; since active resources (servers, network elements, and A/C units) that are idle lead to energy waste" [23].

3. CLOUD ARCHITECTURE

3.1 Overview of Cloud

Clouds computing could be comprised of several heterogeneous components/systems such as grids computing, cluster, super computers etc. (Figure 1). This aggregation is used by millions of users. Consider, for example the case of Microsoft's Live, where the system has around 300 million users. Added to this, there are almost 330,000 application developers of Amazon EC2. Cloud Architecture consists of software applications, which use Internet-accessible on-demand services. Therefore, these applications are considered as an essential computing infrastructure that is used when it is required (such as processing a user request) and to perform a specific job by giving up unwanted resources. Also drawing the needed resources on-demand (like compute servers or storage).

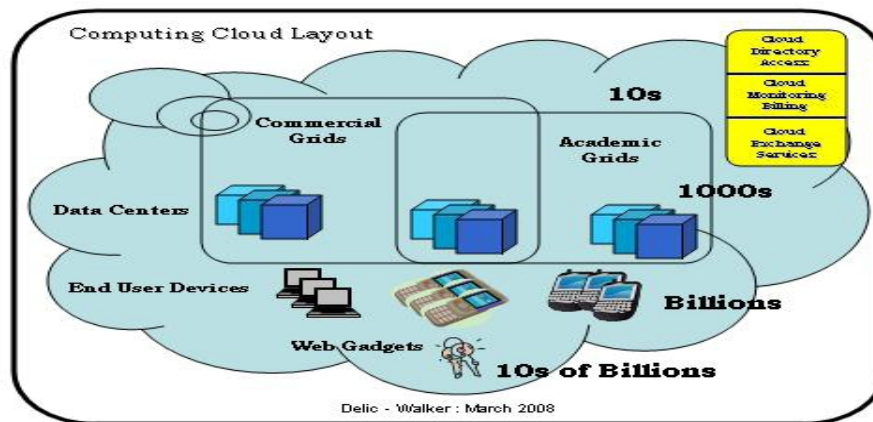


Figure 1. Computing Cloud

Source: [12]

There are some difficulties of large-scale data processing that are addressed by cloud architecture.

- First of all, it is difficult to increase the number of machines to reach an application requirement.
- It is not easy to achieve the machines when needed by any application.
- It is complex to distribute and organize a large-scale job on different machines for example running processes on them and provision another machine to recover if one of them stops working.
- It is hard to dispose of all those machines when the job is completed.
- It is difficult to auto scale up and down based on dynamic workloads. These problems have been solved by cloud architectures.

Plenty of applications may be effective because they utilize the power of the cloud Architectures. Some of these applications are widely and frequently used, for instance Processing Pipelines (Video trans coding pipelines – trans code AVI to MPEG movies), Batch Processing Systems (Log analysis – analyze and generate daily/weekly reports) and Websites(“Seasonal Websites” - websites that only run during the tax season or the holiday season “Black Friday” or Christmas).

3.2 High-level market-oriented Cloud architecture.

The figure below illustrates a high level architecture for supporting market-oriented resource allocation in Data Centres and Clouds.

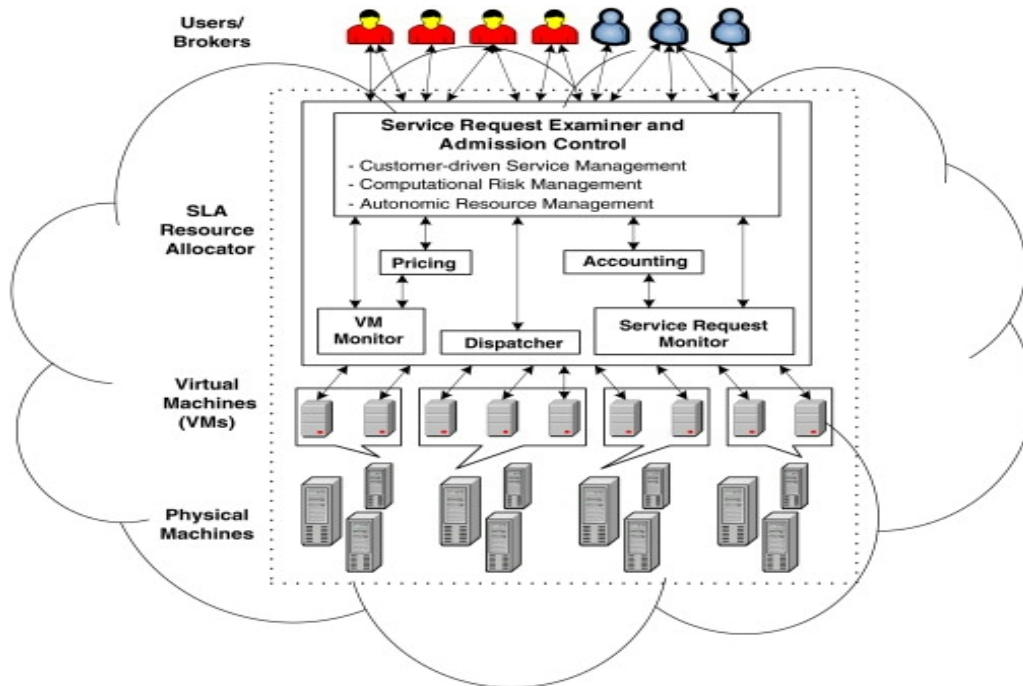


Figure 2: High-level market-oriented Cloud architecture.
Source: [17].

Figure2 depicts the four most valuable elements in Cloud architecture. First, Users/Brokers entity, which takes care of submitting service, requests that come to the data centre and the Cloud from any where to put them in the process.

Second, Service Level Agreement Resource Allocator entity that behaves as the interface between the Data Centre/Cloud service provider and external users/brokers. In this entity there are some required interaction between some mechanisms which appear clearly in the figure above such as Service Request Examiner and Admission Control mechanism which interprets the submitted request for quality of service requirements before deciding whether to accept the request or not, Pricing that determine the ways of charging the service requests based on several factors (for instance submission time (peak / off-peak).

In addition, pricing rates (fixed/changing) or availability of resources (supply/demand), Accounting (the calculation of the final cost that can be charged to the users will be in this mechanism), VM Monitor that checking whether Virtual Machines and their resource privileges are available or not, Dispatcher (which do a process to execute accepted service requests on allocated VMs) and Service Request Monitor (the execution progress of service requests can be Monitoring and tracking in this mechanism).

Third, VMs (Virtual Machines) entity provides a high flexibility to configure different partitions of resources on the same physical machine to various requirements of service requests. Moreover,

there is a possibility to run applications simultaneously based on different operating system environments on a single physical machine. These caused by the dynamic launch and stop of Multiple VMs on a single physical machine to meet accepted service requests. Finally, Physical Machines entity, a several computing servers that have resources to meet service requests in the data centre.

3.3 GrepTheWeb.

One of the most significant and effective applications which is currently in production at Amazon is called GrepTheWeb, which is an example of component service, integrated with cloud architecture. GrepTheWeb can be used to do more specific searches such as finding files that have different contents for instance documents that have a particular HTML tag or META tag. Also, to search for documents which have particular punctuations (“hello!”), mathematical equations (“ $f(g) = g + h$ ”), source code, e-mail addresses or other patterns.

Moreover, GrepTheWeb design aims to scales in many aspects including more powerful pattern-matching languages, more synchronized users of general database and better result characters. In addition, GrepTheWeb does not need the cost of maintaining idle machines and a heavy upfront investment, which keeps the costs of processing down. In terms of time performance, the job has been distributed into multiple tasks to execute a Distributed Grep operation that runs those tasks on multiple nodes in parallel.

Figure3 depicts a high-level illustration of GrepTheWeb architecture. GrepTheWeb has two inputs the first one is the output of the Search Results Service, which is a list of links that are zipped in a single file. The second input is a regular expression. The output is a filtered subset of document links stored and gzipped into a single file. Developers can get the status of their jobs by calling `GetStatus()` to see whether the execution is completed or not as long as the overall process is asynchronous.

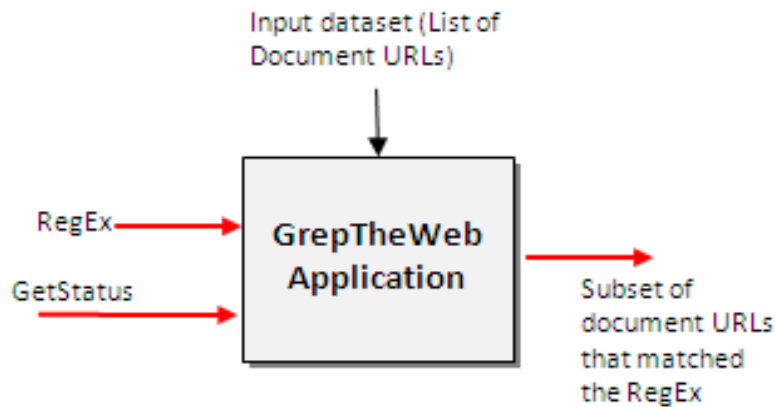


Figure 3 : GrepTheWeb Architecture - Zoom Level 1
Source: [11]

Figure 4 shows GrepTheWeb architecture consisting of Amazon Web Services components which are Amazon S3 (to retrieve input datasets and to store the output dataset), Amazon SQS (to buffer requests), Amazon SimpleDB (for storing intermediate status, log, and for user data about tasks), Amazon EC2 (to run a large distributed processing Hadoop cluster on-demand) and

Hadoop framework (for distributed processing, to schedule jobs and doing parallelization automatically).

Figure 5 illustrates the workflow of GrepTheWeb. It does its processing in four stages as appeared down. The first stage (launch phase) has the responsibility to validate and initiate the processing of a GrepTheWeb request, instantiate Amazon EC2 instances, launch the Hadoop cluster on them and start all the job processes. The second phase (monitor phase) do several significant jobs such as monitoring the EC2 cluster, maps, reduces, and checking for success and failure. The third phase (shutdown phase) to shut down Amazon EC2 instances and Hadoop processes, whilst deleting Amazon SimpleDB transient data in the Final phase (cleanup phase).

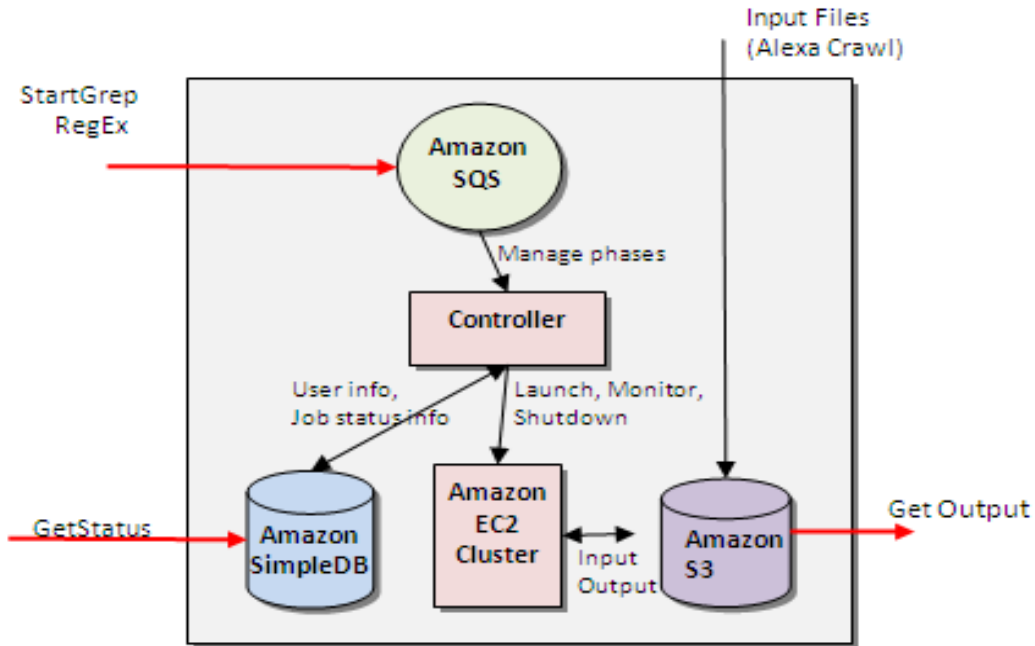


Figure 4 : GrepTheWeb Architecture - Zoom Level 2

Source: [11]

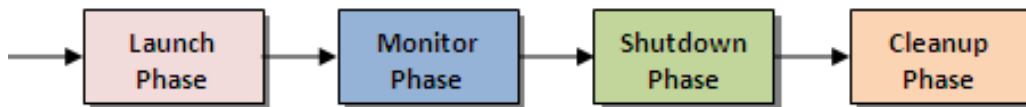


Figure 5: Phases of GrepTheWeb Architecture

Source: [11]

Detailed Workflow for GrepTheWeb architecture appeared in Figure 6 below:
The following steps are summarised from [11]

1. If queues and controller threads have not been initially created at the start of the application, they are made to do so. Their relevant controller threads for any messages they may contain poll queues.
2. A lunch message is added to the queue as soon as startGerp request is received.

3. Launch phase: The launch controller thread does a series of actions as follows:

- It picks up the launch message.
- Carries out the lunch task.
- Updates the current status and timestamps in the Amazon SimpleDB domain.
- Adds a new message in the monitor queue and finally araises the message from the lunch queue after processing.

4. Monitor phase: In this phase the monitor controller performs the following actions :

- Picks up this message,
- Validates the status or error in Amazon SimpleDB
- Carries out the monitor task
- Updates the status in the Amazon SimpleDB domain
- Adds a new message into the shutdown and billing queues
- Finally deletes the message from monitor queue after processing.

5. Shutdown phase in this phase the shutdown controller takes the following actions:

- Picks up the message from the shutdown queue.
- Carries out the shutdown task.
- Updates the status and timestamps in Amazon SimpleDB domain.
- Deletes the message from the shutdown queue after processing.

6. Cleanup phase: the purpose of this phases to archives the SimpleDB together with user information.

7. In the end to get the statues of the system, users can execute GetStatus on the service end point and then download the filtered results from Amazon S3 interface component.

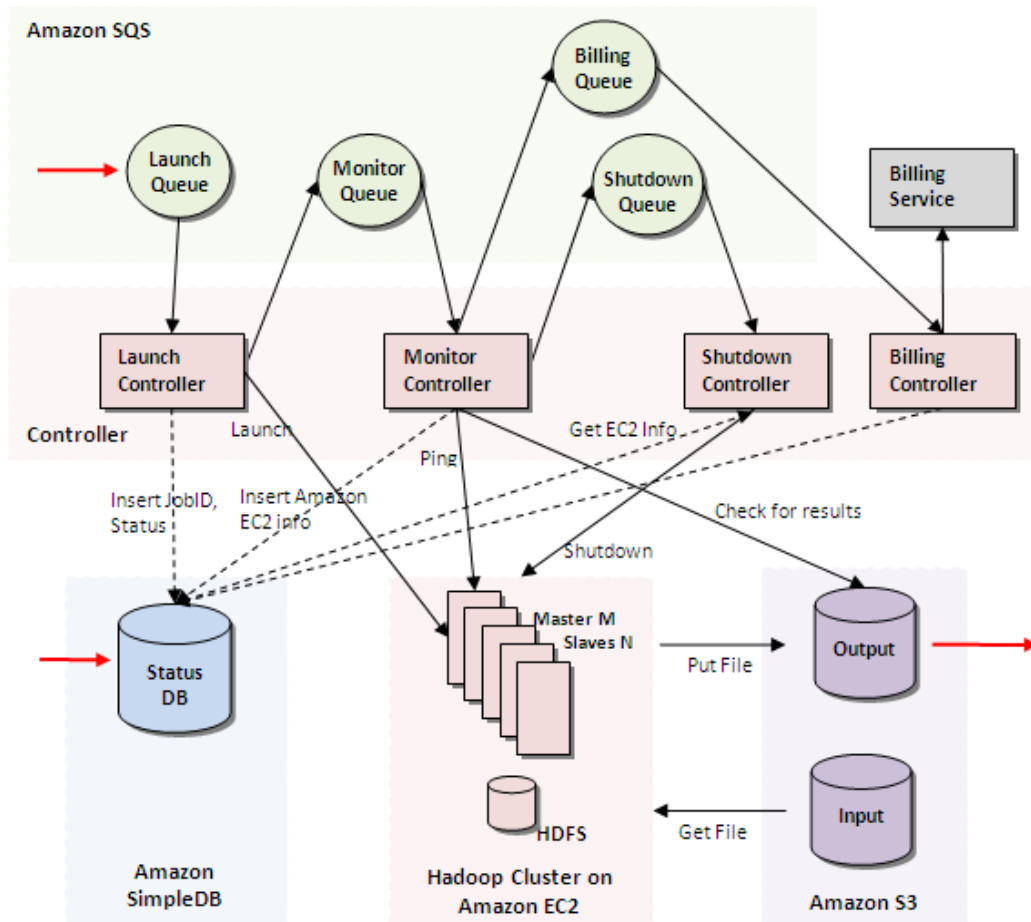


Figure 6: GrepTheWeb Architecture - Zoom Level 3
Source: [11]

HADOOP MAP REDUCE

According to [11]:

“Hadoop is an open source distributed processing framework that allows computation of large datasets by splitting the dataset into manageable chunks, spreading it across a fleet of machines and managing the overall process by launching jobs, processing the job no matter where the data is physically located and, at the end, aggregating the job output into a final result”.

It typically works in three phases, as has shown in figure7 below:

- A map phase that transforms the input into an intermediate representation of key value pairs.
- A combine phase (handled by Hadoop itself) combines and sorts by the keys.
- A reduce phase recombines the intermediate representation into the final output.

Since developers implement two interfaces, Mapper and Reducer, all the distributed processing (automatic parallelization, job scheduling, job monitoring, and result aggregation) is taken care

by the Hadoop which has a master process running on one node to oversee a pool of slave processes (also called workers) running on separate nodes. Thus, the mechanism of Hadoop can be basically described as following [11]:

Firstly, the inputs are split into chunks, which are assigned to slaves where each slave carried out the map task (logic specified by user) on each pair found in the chunk. Then, the result will be written locally and the master is informed about the completed status. Next, Hadoop combines all the results and sorts the results by the keys. Finally, these keys are assigned to the reducers which pulls the results using an iterator, runs the reduce task (logic specified by user), the final output is sent back to distributed file system.

The number of Hadoop slave instances can be decided by reservation logic for GrepTheWeb, based on the complexity of the regex and the input dataset [11].

4. THE COST OF A CLOUD

In this section the discussion focuses on cost models in some Components in Cloud architecture such as those in table2. The important question about the cost of cloud is Where does the cost go in today's cloud service data centres?. The table below provides a rough guide to associated costs [2].

Amortized Cost	Component	Sub-Components
~45%	Servers	CPU, memory, storage systems
~25%	Infrastructure	Power distribution and cooling
~15%	Power draw	Electrical utility costs
~15%	Network	Links, transit, equipment

Table 2: Guide to where costs go in the data centre.

Source: [2]

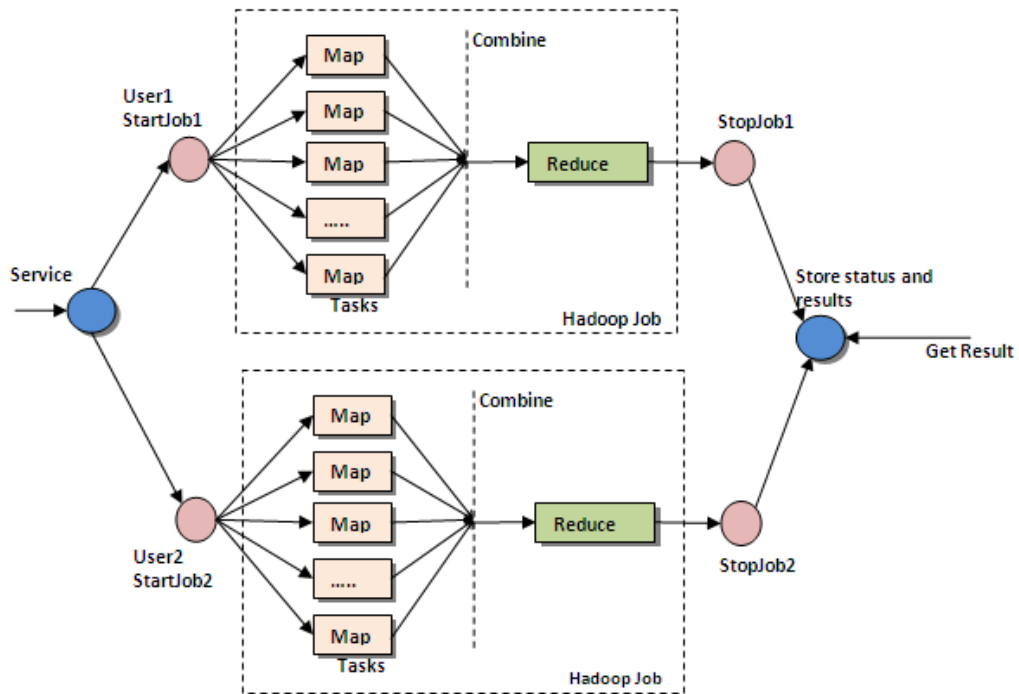


Figure 7: Map Reduce Operation (in GrepTheWeb)
Source: [11]

4.1 Server Cost

As shown in table2 above, the greatest cost goes to servers. There are some structural factors that prevent accomplishing high utilization of data centre. First of all, Uneven Application Fit, incorporating unnecessary components such as memory, network and storage, which might be unutilized by applications, fitted on the server. Secondly, Uncertainty in Demand Forecasts, service demands in Cloud particularly new service need to be high state of certainty by going beyond what forecasts would predict, otherwise it might spike quickly. Thirdly, Long provisioning time scales, purchases have a tendency to be large, with components bought in bulk. Infrastructure is intended to extend over long time periods. Servers are intended to continue as many years as possible, with annual growth ordered [2]. In addition, there are some other valuable features (e.g. Risk Management, Hoarding and Virtualization short-falls), which reflect badly utilization of data centre [2].

Thus, the main effective factor to raise utilization and solve the above problems is agility, the ability to increase and decrease the number of resources to match demand and to extract those resources from optimal location [2].

4.2 Infrastructure Cost.

As table 1 represents the aggregate cost is substantial. In addition, as shown in Figure 8 the drawing of electric power from the utility leads to capital investments in large-scale generators, transformers, and Uninterruptible Power Supply (UPS) systems. The cost of the difference in time between ordering and delivering is 8 months. However, relaxing the requirement for individual server can be used to reduce the infrastructure cost [2].

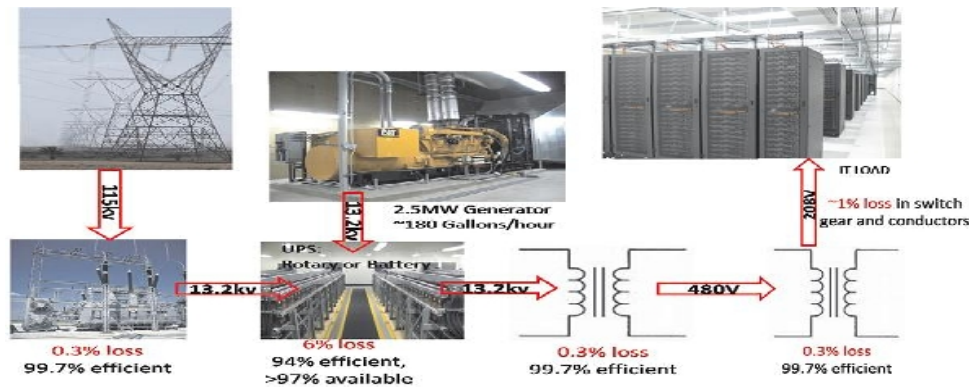


Figure 8: Infrastructure components.
Source: [2]

4.3 Power cost.

The metric provided by the green Grid [19] is used to describe the data centre Power usage efficiency (PUE) as $PUE = \frac{\text{Total Facility Power}}{\text{IT Load}}$. A PUE of the art is 1.7, which is below the average of the normal state. To estimate power draw for a mega data centre, a PUE of 1.7 is assumed, a reasonable utility price of \$.07 per KWH, 50,000 servers with each drawing on average 180W (servers draw as much as 65% of peak when idle), the total cost comes to $50,000 \cdot 180/1000 \cdot 1.7 \cdot \$0.07 \cdot 24 \cdot 365 = \9.3 million a year.

However, the largest impact on the power cost of data centres is decreasing the power draw of each server. Hardware innovation, including use of high efficiency power supplies and voltage regulation modules are an important improvement to reduce the power cost. In addition, another approach, which has been proposed by Barroso and Hlzle [13], is based on energy proportionality to create servers that are closer to implementing energy proportionality to improve efficiency.

4.4 Network cost.

The networking costs are based in wide area networking [2]:

1. Peering, in delivering packets to end-users.
2. The inter-data centre links carrying traffic between geographically distributed data centres.
3. Regional facilities (backhaul, metro-area connectivity, co-location space) needed to reach wide area network interconnection sites.

Reducing the cost of wide area networking can be achieved using a clever design of peering and transit strategies, combined with optimal placement of micro and mega data centres. Moreover, a better design of the services themselves partitioning their functionality and their state between data centres can be also used to reduce the cost.

5. COMPARISON BETWEEN CLOUD AND GRID COMPUTING

5.1 Grid Definition

While still there are several different conceptions upon the definition of the grids Ian Foster has indicate [8] a definition of the Grid as “a system that coordinates resources which are not subject

to centralized control, using standard, open, general-purpose protocols and interfaces to deliver nontrivial qualities of service”.

Nowadays, it can be clearly observed that Clouds are the latest paradigm to emerge that promises reliable services delivered through next-generation data centres which are built on virtualized compute and storage technologies [1].

The popularity of different paradigms varies with time. The web search popularity, as measured by the Google search trends during the last 12 months, for terms “cluster computing”, “Grid computing”, and “Cloud computing” is shown in Figure 2. From the Google trends, it can be observed that cluster computing was a popular term during 1990s, from early 2000 Grid computing become popular, and recently Cloud computing started gaining popularity [17].

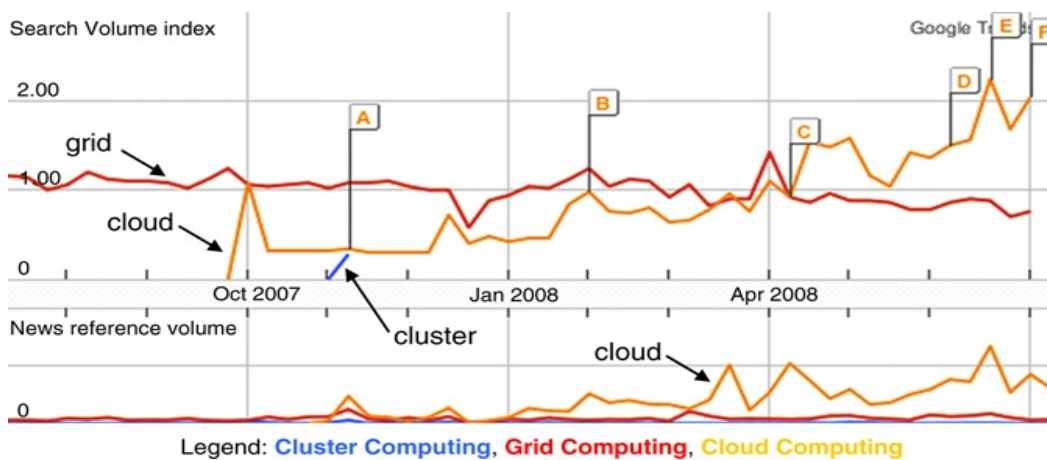


Figure 9: Google search trends for the last 12 months.
Source [17]

The following points are adopted from [17]:

Spot points in this figure indicate the release of news related to Cloud computing as follows:

- A. IBM Introduces 'Blue Cloud' Computing, CIO Today - Nov 15 2007.
- B. IBM, EU Launch RESERVOIR Research Initiative for Cloud Computing, IT News Online Feb 7 2008.
- C. Google and Salesforce.com in Cloud computing deal, Siliconrepublic.com - Apr 14 2008.
- D. Demystifying Cloud Computing, Intelligent Enterprise - Jun 11 2008.
- E. Yahoo realigns to support Cloud computing, 'core strategies', San Antonio Business Journal Jun 27 2008.
- F. Merrill Lynch Estimates "Cloud Computing" To Be \$100 Billion Market, SYS-CON Media Jul 8 2008.

Other more recent news includes the following:

- Yahoo, Intel and HP form Cloud computing labs, Reseller News - Jul 29 2008.

- How Cloud Computing Is Changing The World, Pittsburgh Channel.com - Aug 4 2008.
- SIM tone Corporation Takes Cloud Computing to the Next Level with Launch of First Wireless, "Zero Touch" Universal Cloud Computing Terminal, TMC net - Sep 8 2008.

5.2 Feature Comparison

There are many important features that are common to both Cloud and Grid. Table 2 below presents various features of Grids and Clouds, highlighting the similarities and differences between both paradigms. The first one of these features is Resource Sharing; Grids appear to be fairly sharing resources across organizations, whereas Clouds provide the resources that the Service Provider requires on demand. Another feature is Heterogeneity, both Cloud and Grid Computing support the aggregation of heterogeneous resources. Add to these, Virtualization feature that covers both, data and computing resources [15]. Mutually, Clouds and Grids add the virtualization of hardware resources.

As far as the Security feature is concerned, it has been argued by some authors that security has not been seriously explored [18]. Grids have not dealt with end user security, however in Clouds each user has an access to its individual virtualized environment. Clouds system is facing a serious problem caused by lack of high level services; this may be as a result of the low level of maturity associated to Clouds paradigm. In contrast, Grids have a number of these high level services for instance data transfer, metadata search [7].

Moreover, Still there are some valuable features (e.g. Scalability and Self Management, Usability, Standardization and Payment Model Quality of Service), which are summarized in the table below.

Table reference as extracted from [14]

Feature	Grid	Cloud
Resource Sharing	Collaboration (VOs, fair share).	Assigned resources are not shared.
Resource Heterogeneity	Aggregation of heterogeneous resources.	Aggregation of heterogeneous resources.
Virtualization	Virtualization of data and computing resources.	Virtualization of hardware and software platforms.
Security	Security through credential delegations.	Security through isolation.
High Level Services	Plenty of high level services.	No high level services defined yet.
Architecture	Service orientated.	User chosen architecture.
Software Dependencies	Application domain-dependent software.	Application domain independent software.
Platform Awareness	The client software must be Grid-	The SP software works on a customized

	enabled.	environment.
Software Workflow	Applications require a predefined workflow of services.	Workflow is not essential for most applications.
Scalability	Nodes and sites scalability.	Nodes, sites, and hardware scalability.
Self-management	Reconfigurability.	Reconfigurability, self-healing.
Centralization Degree	Decentralized control.	Centralized control (until now).
Usability	Hard to manage.	User friendliness.
Standardization	Standardization and interoperability.	Lack of standards for Clouds interoperability.
User Access	Access transparency for the end user.	Access transparency for the end user.
Payment Model	Rigid.	Flexible.
QoS Guarantees	Limited support, often best-effort only.	Limited support, focused on availability and uptime.

Table 3: Grid vs. Cloud Characteristics

There are many similarities between Grid and Cloud computing, both systems share the same basic goal [9]: “to reduce the cost of computing, increase reliability, and increase flexibility by transforming computers from something that we buy and operate our-selves to something that is operated by a third party”.

However, there are some differences Grids come out from existing heterogeneous resources to add new abstraction layers, allowing the definition of new and complex services by aggregation. In contrast, Clouds are not effectively composed of homogenous components, and it is possible to be designed on top of an existing grid.

The essential objective for creating the Cloud system is to provide a determined set of capabilities to the user; therefore the principle of the design is a specific goal interface. In contrast, Grids are considered to be general purpose, so they display a complete set of available system capabilities and the resulting interface available for users and applications remains low level [10].

5.3 Convergence of Grids and Clouds

Various approaches exist, combining Clouds and Grids together can be seen as a combination of advanced networking with sophisticated virtualization. However, Grids need to accelerate the incorporation of virtualization technologies to be nearby Clouds, by gaining some advantages that Clouds show (migrability, hardware level scalability). Moreover, Grids need to provide easier

entry points so as to enable a wider adoption by end users, i.e., Grids are meant to be virtualized and automatically scalable utilities, which clearly shows a convergence with current Clouds [14].

The problem of delivering the user's needs leads to reducing and escalating in service's computing resource. Therefore, implementing the cloud on third-party infrastructure can optimize the payment cost. In addition, to widen the spectrum of offers and stimulate competition the same Grid services can be deployed with different computing problems [5].

There are two solutions to deal with increased processing power, additional low-level services can be deployed automatically, and to deal with increased number of users, additional high-level services can be instantiated to serve the interface [5]. Firstly, solution is based on incorporating an index service to register grid services. Secondly, load balancing can be used to achieve this approach at interface nodes. For example, a master-worker is implemented as multiple workers who need only to register themselves with the index service to become available to the master. The figure below shows a scenario compatible with the master-worker Paradigm.

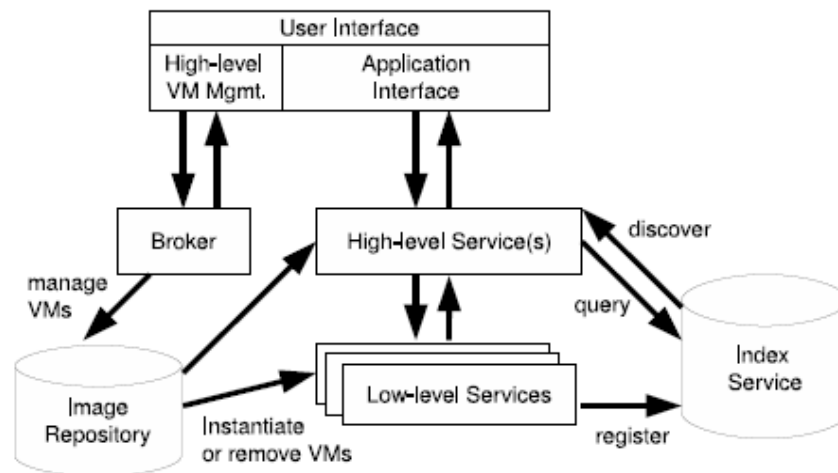


Figure 15 : Grid application in a cloud environment
Source:[5]

5.4 Security and Policy Issues in Cloud and Grid Computing

Cloud and grid computing technologies are used as inexpensive systems to gather and utilize computational capability together. These technologies try to improve jobs and application services by arranging machines and distributed resources in a single huge computational entity [24]. Clouds mainly consist of data centres which are owned by the same institute. The homogeneity within each data centre in the infrastructure is the main feature for the cloud computing compared to grid computing. In this case, any conflict between a heterogeneous data centre and/or different administration domains can become a serious issue for cloud interoperability. It can be seen that the stages of anonymity and privacy provided by cloud to the external users will be less than the user of desktop in numerous situations [25]. On the other side, grids was originally established on the idea that resources infrastructure are dynamic and heterogeneous in their nature. This means different organization with different administrative domains. This also means that security was taken into account from the beginning when the grid system was originally built. Presently, the security paradigm for clouds appears to be fairly less secure than the model in grids environment. The infrastructure of the cloud normally depends on

web models (over SSL) to establish and access account information for the external users, and allows them to change or reset their keys (passwords) with a new ones by email in what can be considered as unencrypted communication [26].

In general, the data in the cloud environment is distributed over multiple servers in the cluster and hosted by third parties (at least for some time) who are hidden from the data owner or the external users. This means the privacy of the users' data and ability to give the permissions to handle this data to a specific domains or users by the original data owner is less than expected. On the other hand, the situation in the grid environment is completely different, where the grid administrators or grid users can control the handling of their data or the permission and policies over their resources.

6. CRITICAL DISCUSSION

In section 2 we reviewed several definitions of Clouds, trying to clarify the Cloud paradigm. However, the commonality of these definitions focuses on software and infrastructures as services, with an explicit attention to scalability of software and application provision and the benefits of economics of scale and sharing resources under uncertainty.

One simple way is to look at the Cloud as a large-scale (or even ultra-large scale) distributed environment, where a group of computers operates in the sky providing services to anyone from anywhere regardless of location. Clouds are virtual, as the locations of the servers are unknown and irrelevant to the end user. Clouds set the promises for scalability, henceforth, providing promises for continuous dependability on the infrastructure to address the growing business needs, expansion of the service provision, and utilisation of the infrastructure through resource sharing.

Despite the fact that [11] has reported on examples of "cloud architectures", this is meant to be examples of architectures of components constituting the cloud. The description of the architecture of GrepTheWeb provides an example of how application architectures could be integrated with the overall cloud infrastructure and architectures.

Amazon has created a very challenging application (GrepTheWeb) that facilitates trend aspects that were addressed as a complicated problem in early days of the cloud. Therefore, large organizations should contribute and treat such this application to find out and improve using of many features in the cloud.

As the promised unlimited scalability, data sharing and resource sharing are characteristics of the Cloud; it is worth mentioning that security becomes a concern, which future research should address. This is because of insufficient control over user's information and data, whose location is unknown due to the nature of the Clouds

By moving services to the cloud, architects will be faced with the challenges in making the cloud infrastructure dependable. This is because the services will be centralised and be part of the infrastructure. This may require more effort to ensure that infrastructure is reliable, available and secure.

In terms of costs related to servers, using the cloud systems is a cost-effective alternative to using company's own servers and related hardware. Since, servers are pooled and shared by a large number of companies as per service charge, it is a good option.

The infrastructure required for a successful cloud computing is very expensive specially the cost of power which is required by the dedicated data centres. Although, there are steps to reduce these costs, however they are at their early stages.

7. CONCLUSION

In conclusion, cloud computing as mentioned above is a new technology of computer network, providing the web services in high quality and lower cost comparing to normal technique. Using cloud computing might contribute to improvement of services in other related technologies specially the previous generations such as Grid computing. Cloud computing is almost certainly set to be developed and become an attractive option for many corporations. The future of computing is set to be successful. It is the next generation and can provide tremendous value to companies. It can help companies achieve more efficient use of their IT hardware and software investments, which in turn can lead the acceleration and adoption of innovations.

Clouds consist of data centres which are owned by the same institute. The homogeneity within each data centre in the infrastructure is the main feature for the cloud computing compared to grid computing. In this case, any conflict between a heterogeneous data centre and/or different administration domains can become a serious issue for cloud interoperability. It can be seen that the stages of anonymity and privacy provided by cloud to the external users will be less than the user of desktop in numerous situations. On the other side, grids were originally established with the knowledge that resources infrastructure are dynamic and heterogeneous in their nature. This implies different organization with different administrative domains. This also means that security was taken into account from the beginning when the grid system was originally built. Presently, the security paradigm for clouds appears to be less secure than the model in grids environment.

It has been arguably said that the cloud computing might reduce the cost of the infrastructure by shrinking the number of resources such as servers. Nevertheless, the services that provided in such environments should be working as if in Grid computing for example. Cloud computing enables innovation by minimising need for innovators to find resources to develop, test, and make their innovations available to the user's community. Innovators become free to focus on innovating rather than the logistics and management of systems. The reduced cost of infrastructure will make it a way forward for many organisations. Cloud Architectures shows how flexible is building applications on on-demand infrastructures. GrepTheWeb is a good example, demonstrating this flexibility. The Cloud seems to be the brightness future of many IT infrastructures and systems.

References

- [1] A. Weiss. (2007). Computing in the Clouds, netWorker 11 (4) 16_25
- [2] A. Greenberg, J. Hamilton, D. A. Maltz, P. Patel. (2009). <http://ccr.sigcomm.org/online/files/p68-v39n1o-greenberg.pdf>
- [3] Amazon simple storage service. Web Page <http://www.amazon.com/gp/browse.html?node=16427261>
- [4] R. Desisto, D. Plummer, Smith(2007). "Tutorial for Understanding the Relationship Between Cloud Computing and SaaS". http://www.gartner.com/DisplayDocument?ref=g_search&id=640707
- [5] G. V. Mc Evoy, B. Schulze(2008). "Using Clouds to address Grid Limitations". at <http://portal.acm.org/citation.cfm?id=1462704.1462715>
- [6] G. Boss, P. Malladi, D. Quan, L. Legregni, H. Hall. (2007). "Cloud Computing" www.ibm.com/developerworks/websphere/zones/hipods/
- [7] H. Stockinger. (2007). "Defining the grid: a snapshot on the current view". The Journal of Supercomputing, (1):3-17

- [8] I. Foster(2002). What is the Grid? A Three Point Checklist. *Grid Today*, vol. 1, no. 6, pp. 22—25.
- [9] I. Foster's. (2008). weblog. <http://ianfoster.typepad.com/blog/2008/01/theres-grid-in.html>
- [10] S . Jha, A. Merzky, G. Fox. Using Clouds to Provide Grids Higher-Levels of Abstraction and Explicit Support for Usage Modes. [Online] Available: http://www.ogf.org/OGF_Special_Issue/cloudgrid-saga.pdf
- [11] J. Varia. (2008).“Cloud Architectures”. jineshvaria.s3.amazonaws.com/public/cloudarchitectures-varia.pdf
- [12] K. A. Delic and M. A. Walker. (2008) .” Emergence of The Academic Computing Clouds”. http://www.acm.org/ubiquity/volume_9/v9i31_delic.html
- [13] L. A. Barroso and U. Hölzle. (2007). "The case for energy-proportional computing,". *Computer*, vol. 40, no. 12, pp. 33-37
- [14] L. M. Vaquero, L. R. Merino , J. Caceres, M Lindner. (2009). “A Break in the Clouds: Towards a Cloud Definition”. <http://portal.acm.org/citation.cfm?id=1496091.1496100>
- [15] Members of EGEE-II. An egee comparative study: Grids and clouds - evolution or revolution. Technical report, Enabling Grids for E-science Project, June 2008. Electronic version available at <https://edms.cern.ch/document/925013/>
- [16] Next Generation Grids (NGG) Experts Group (2008). Requirements and options for European grids research 2005-2010 and beyond (ngg expert group report). Technical report, European Commission, Available at http://www.semanticgrid.org/docs/ngg2_eg_final.pdf
- [17] R. Buyya, C. S. Yeoa, S. Venugopala, J. Broberg, I. Brandic. (2008). “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility”. journal homepage: www.elsevier.com/locate/fgcs
- [18] S. Jha, A. Merzky, G. Fox. (2008) . “Using clouds to provide grids higher-levels of abstraction and explicit support for usage modes”. Technical report, Open Grid Forum. Available at <http://grids.ucs.indiana.edu/ptliupages/publications/cloud-grid-saga.pdf>.
- [19] The Green Grid. URL <http://www.thegreengrid.org>.
- [20] De Chaves, S.A., Uriarte, R.B., Westphall, C.B.: Toward an architecture for monitoring private clouds. *IEEE Communications Magazine* 49(12), 130–137 (2011), <http://dblp.uni-trier.de/db/journals/cm/cm49.html#ChavesUW11>.
- [21] Mell, P., Grance, T.: The nist definition of cloud computing (draft) recommendations of the national institute of standards and technology. *Nist Special Publication* 145(6), 7 (2011), http://csrc.nist.gov/publications/drafts/800-145/Draft-SP-800-145_cloud-definition.pdf.
- [22] Sotomayor, B., Montero, R.S., Llorente, I.M., Foster, I.: Virtual infrastructure management in private and hybrid clouds. *IEEE Internet Computing* 13, 14–22 (September 2009), <http://dl.acm.org/citation.cfm?id=1638588.1638692>
- [23] Werner, J., Geronimo, G., Westphall, C., Koch, F., Freitas, R.: Simulator improvements to validate the green cloud computing approach. In: *Network Operations and Management Symposium (LANOMS)*, 2011 7th Latin American. pp. 1–8 (oct 2011).
- [24] Douglas O. Balen, C.B.W., Westphall, C.M.: Experimental Assessment Of Routing For Grid And Cloud. *The Tenth International Conference On Networks(Icn 2011)* Pp. 341–346 (2011).
- [25] Delaney, K.J., Vara, V.: Google Plans Service To Store Users’ Data (November 27, 2007), [Http://Online.Wsj.Com/Article/Sb119612660573504716.Html](http://Online.Wsj.Com/Article/Sb119612660573504716.Html)
- [26] Jaeger, P., Lin, J., Grimes, J.: Cloud Computing And Information Policy: Computing In A Policy Cloud? *Journal Of Information Technology Politics* 5(3), 269– 283 (2008), [Http://Www.Informaworld.Com/Openurl?Genre=Article&Doi=10.1080/19331680802425479&Magic=Crossref](http://Www.Informaworld.Com/Openurl?Genre=Article&Doi=10.1080/19331680802425479&Magic=Crossref)
- [27] Foster, I., Kesselman, K.: *The Grid: Blueprint For A Future Computing Infrastructure*. In: Morgan Kaufmann In Computer Architecture And Design (1999).
- [28] Grimshaw, A.S., Humphrey, M.A., Natrajan, A.: A Philosophical And Technical Comparison Of Legion And Globus. *Ibm J. Res. Dev.* 48, 233–254 (March 2004), [Http://Dx.Doi.Org/10.1147/Rd.482.0233](http://Dx.Doi.Org/10.1147/Rd.482.0233)