International Journal of Cloud Computing and Services Science (IJ-CLOSER)
Vol.1, No.1, March 2012, pp. 36 ~ 45
ISSN: 2089-3337

An Approach to Identify the Optimal Cloud in Cloud Federation

Saumitra Baleshwar Govil*, Karthik T*, Karthikeyan S*, Vijay K. Chaurasiya*, Santanu Das*
* MBA & MS-CLIS Division, Indian Institute of Information Technology – Allahabad, India.

Article Info

Article history:

Received Jan 26th, 2012
Revised Feb 28th, 2012
Accepted March 1th, 2012

Keyword:

Federated Cloud
Optimal Cloud Service Provider
Service Level Agreement
Benchmark

ABSTRACT

Enterprises are migrating towards cloud computing for their ability to provide agility, robustness and feasibility in operations. To increase the reliability and availability of services, clouds have grown into federated clouds i.e., union of clouds. There are still major issues in federated clouds, which when solved could lead to increased satisfaction to both service providers and clients alike. One such issue is to select the optimal foreign cloud amongst the federation, which provides services according to the client requirements. In this paper, we propose a model to select the optimal cloud service provider based on the capability and performance of the available clouds in the federation. We use two matrix models to obtain the capability and performance parametric values. They are matched with the client requirements and the optimal foreign cloud service provider is selected.

Corresponding Author:

Vijay K. Chaurasiya,
MBA & MS-CLIS Division,
Indian Institute of Information Technology,
Deoghat, Jhalwa, Allahabad, Uttar Pradesh, India - 211012.
Email: vijay.chaurasiya@gmail.com, vijayk@iiita.ac.in

1. INTRODUCTION

Cloud computing is a service delivery model for providing on-demand access to resources in a shared pool, which can be requested and released, thereby evolving a pay per use strategy for services [1]. This emerging technology is attracting various organizations under its fold by virtue of its effectiveness in efficient management of resources with affordable cost. Cloud computing is characterized [2] by

- On-Demand Self Service: Provides the automated environment for obtaining services with least human interaction.
- Broad Network Access: Capable of providing services to various types of devices like laptop, PDA, mobile phones over the network.
- Resource Pooling: By enabling multi-tenant model, provides centralized pooling of resources thereby supporting their efficient use.
- Rapid Elasticity: The capabilities can be rapidly and elastically demanded.
- Measured Service: Metering capability is used for the controlled and optimized use of resources.

The cloud computing paradigm has brought major changes in traditional computing practices, where resources that were confined in a single stand-alone machine, are now outsourced to some other entity. The entity holds and manages resources, either for a single machine or multiple machines, thus acting as a resource pool and providing scalable and flexible environment to the efficient functioning of the machine(s). This entity is called as cloud service provider (SP).

Cloud service providers are akin to shopping malls which sell different commodities at different shops at the same place. These providers supply various services according to the need of the client(s). The movement towards cloud computing is justified only when the quality of service (QoS) is equivalent to that of traditional systems. This QoS is maintained by customizing the resources with the needs of the customer.

The customization of services can be established by a document that contains the definition of services, the conditions under which the services are provided and the obligations in the form of penalties if the guarantees are not met. This document is known as the Service Level Agreement (SLA). Though the SLA defines the penalty for non availability of services, clients suffer the wrath of service interruption. To avoid this, enterprises are moving from traditional cloud to federated cloud.

The paper is organized as follows. Section 2 deals with background study regarding clouds and federated clouds. Section 3 contains the objective of the proposal and Section 4 discusses the proposed model comprehensively. Section 5 talks about the future enhancements necessary in the model. Section 6 concludes the paper.

2. BACKGROUND

According to NIST [3], there are three types of cloud services. The services provided to the customer by cloud service providers are:

- Software as a Service (SaaS): Application and its associated data are provided as a service. Example: Google Docs[4].
- Platform as a Service (PaaS): Development environment is offered as a service. Example: Microsoft Azure.
- Infrastructure as a Service (IaaS): Basic storage and computing capabilities are provided as a service. Example: SalesForce.com.

The four deployment models according to NIST [5] are:

- Public Cloud: They provide and manage resources for multiple enterprises, i.e. all customers share the same infrastructure pool.
- Private Cloud: They exclusively provide services to a single enterprise. They may be on-premise or externally hosted.
- Community Cloud: They provide services to a specific community having shared concerns and is shared by multiple enterprises.
- Hybrid Cloud: They combine multiple clouds (public, private or community).

The cloud architecture contains certain functional interfaces for interaction between the client and the service provider. Some of the functional interfaces [6] are:

- Service Catalog: Manages the service offerings
- Security Manager: Manages security related aspects
- Service Manager: Manages instance of services

Circumstances may arise where a cloud service provider is unable to provide services agreed in the service level agreement due to the lack of resources or sudden increase in incoming workload. This results in the service provider paying the monetary penalty, but the effect of interruption of services is not mitigated. In these situations, the cloud service provider rents the services from other cloud service providers and provides them to the client. The former is called the home cloud service provider and the latter is called the foreign cloud service provider. Thus the cloud service providers form a union called a federated cloud. All the cloud service providers in this federation have Service Level Agreements between them.

Federated clouds, by providing end to end quality of services, offer many advantages [7] over traditional cloud services, which are:

- Guaranteed performance: Due to limited resources, that are available with a single cloud service provider, sudden increase in workload may lead to deterioration of performance. Cloud federation overcomes this disadvantage by hiring resources from foreign cloud service providers, thereby guaranteeing the agreed QoS. Also, high priority processing is guaranteed by delegating low priority processing tasks to foreign cloud service providers.
- Guaranteed availability: During unexpected disasters, the cloud system will be able to recover the services by federating with other cloud service providers in unaffected areas. Availability may be guaranteed according to the priority of the service, as disaster recovery may not be an instant process.
- Convenience of service cooperation: Cloud federations greatly increase the convenience by providing a one stop solution such that the consumer can see all the services together. For example, while applying for a passport, all the associated services may be integrated as one single service.
- Dynamic load distribution: Geographical distribution of clients for every cloud service provider is highly uneven. In order to provide seamless services, dynamic load distribution is facilitated by cloud federations so that they could rise above their geographical shortcomings.

3. OBJECTIVE OF PROPOSAL

For implementing an efficient cloud federation, there are certain functional requirements [7]. Some of them are:

- Matching cloud service provider with client requirements: For guaranteeing the QoS according to the client requirement, SLA's between the cloud service providers must be matched with the client requirements.
- Monitoring: The usage status and the service quality of every resource need to be monitored.
- Provisioning: The resource requirements need to be properly determined and also performance degradation needs to be detected so that resources can be reconfigured and the performance of the service is maintained.
- Resource discovery: Resources that satisfy the resource requirement of the client need to be searched and discovered by mapping them through SLA matching.
- Resource management: The configuration of the resources obtained within the same cloud or other cloud systems needs to be managed.
- Service setup: The necessary environment needs to be setup and necessary data is to be moved or copied, in order to efficiently use the delivered services.
- Authenticating Internetworking: In order to provide seamless services to the customer, such that he is not aware of using a foreign cloud service provider, identity must be federated.
- Releasing Resources: Services need to be monitored and performance has to be measured periodically so that unnecessary resources are released, once the client requirements are met.

The client and the home cloud service provider negotiate the terms of the service level agreement. In a federated cloud, the resource requirements of the client are not limited to resources and services available with the home cloud service provider. In order to satisfy the client requirements, the home cloud service provider searches and identifies the appropriate foreign cloud service provider whose SLA matches with that of the client requirement. Figure 1 illustrates the architecture of the cloud federation.

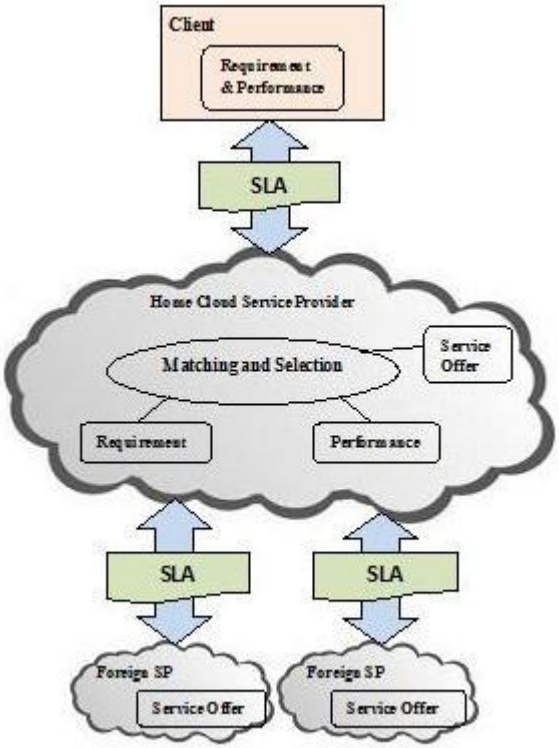

Figure 1. Architecture for cloud federation..

During searching and identification, a mechanism is needed to scrutinize multiple foreign cloud service providers and identify the appropriate foreign cloud service provider(s), which satisfies the client SLA based on their capability and performance respectively. The reason behind considering the above two step processing is that, searching and identifying the foreign cloud based on its capability alone does not

guarantee that the reliability and performance of the service mentioned in the SLA shall be achieved, even though the resources are available. To ensure reliability the cloud service providers are assessed for their performance metrics.

4. PROPOSED MODEL

As discussed above, the home cloud service provider selects the foreign cloud service provider based on the client requirement. We propose a model where the optimal foreign cloud service provider is selected. Our model is illustrated in Figure 2.

Figure 2. Phases of the proposed model.

4.1. Resolution

Availability and Reliability are the two factors which are critical to the client in a cloud service arrangement. These two factors are taken care of by the service catalog and service manager of the cloud service provider respectively. Service catalog is a functional interface which accepts the requests and offers the services. The service manager manages and modifies the deployed services. Whenever one of the critical parameters, either availability or reliability, is not assured according to the norms agreed in the SLA with the client, the cloud service provider resolves to go for a cloud federation.

4.2. Matching

The purpose of matching phase is to identify the foreign cloud service providers that are capable of satisfying the client's requirement. In order to assess the potential of the foreign cloud service providers, capability matrices are used. Every time the home cloud service provider decides to go for a cloud federation for any client requirement, it builds a corresponding capability matrix. This matrix contains certain critical requirement parameters which are decided based on the client's need. Figure 3 illustrates the matching phase.

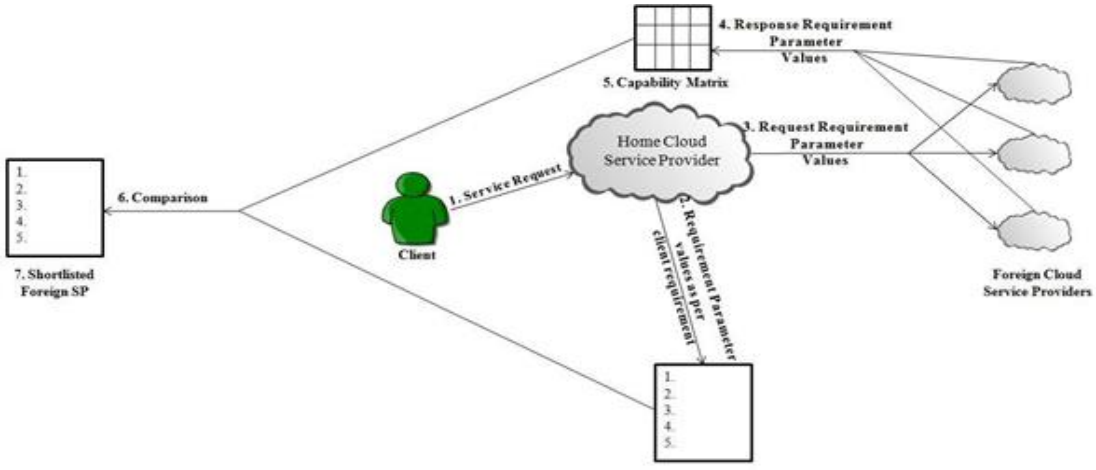

Figure 3. Shortlisting the foreign cloud service providers by matching their capabilities with the client requirements.

Table 1 contains certain cloud services and their respective critical requirement parameters. Now the values of these critical parameters are obtained from various foreign cloud service providers and populated in the capability matrix. Once the capability matrix is completed, the values of it are compared with that of the requirements of the client to shortlist eligible foreign cloud service providers. Table 2 is an example of capability matrix.

Table 1. Critical requirement parameters for certain cloud services.

Services→ **Critical Requirements↓**	**Virtual Business Desktop**	**E-commerce Website Hosting**	**Web Videoconferencing**	**Rich Media Hosting**	**Game Hosting**
Elastic Computing	✓	✓	✓		
Persistent Storage	✓			✓	✓
Data Availability	✓	✓	✓	✓	
Data Backup	✓	✓		✓	
Data Isolation	✓		✓		
Capability to Scale	✓			✓	✓
Load Balancing	✓	✓	✓	✓	✓
Risk Management	✓	✓			
Identity and Access Management	✓	✓	✓		
Compliance	✓	✓	✓		

Table 2. Capability Matrix: Critical requirement parameters of certain cloud service providers.

Service Provider → **Critical Requirements↓**	**Amazon AWS**	**Microsoft Azure**	**Rackspace Cloud**	**Joyent Cloud**
Elastic Computing	Xen Based VM	Windows Azure VM	Xen Based VM	Joyent Smart OS
Data Availability	Cluster formation of data centre	Geo Replication	DC-DC Replication	Smart data centre availability management
Data Backup	Own Tape Backup Service Provider	Not Available	Managed Backup Service	Paid Backup Storage
Data Isolation	Simple Data Isolation	VLAN Isolation Mechanism	Simple Data Isolation	Smart data centre dynamic VLAN
Capability to Scale	Auto Scaling	Scaling on demand using CDN	Scalable on demand up to 192 TB	Scaling by automatic CPU bursting over 3GB per second
Load Balancing	Elastic Load Balancing	Load Balancing by traffic manager	Dedicated Rackspace cloud load balancers	By Paid Joyent Smart Machine Appliances
Risk Management	AWS Control Environment	Microsoft Security Development Lifecycle	Dedicated threat management system alert logic threat manager	Smart Technologies Architecture
Identity and Access Management	AWS IAM	Service Management API	Simple Identity Management	Smart data centre API Security
Compliance	SAS 70 Type II, SOX, ISO 27001, HIPAA	US Safe Harbor, ISO 27001	US Safe Harbor	US Safe Harbor

4.3. Selection

Selection phase tries to identify the best choice among the shortlisted foreign cloud providers from the matching phase. Once the foreign cloud service providers are shortlisted based on client requirement, performance matrices are used to select the best foreign cloud. These performance matrices contain the evaluation values of certain standard benchmark tests. These standard benchmark tests are selected according to the client requirement at the sole discretion of the home cloud service provider. Table 3 contains some of the standard benchmarks used to evaluate certain types of performance [8].

Table 3. Standard benchmark tests for measuring the performance of cloud service providers.

Performance Parameters → **S. No↓**	**CPU Performance**	**Memory I/O**	**Disk I/O**	**Multithreaded CPU Performance**	**Encoding**	**Compression**
1.	7-Zip Compression	Cache Bench Read Cache	BlogBench Read/Write Performance	C-Ray Tracer	FFmpeg AVI to NTSC VCD	7-Zip Compression
2.	C-Ray Tracer	GeekBench	Bonnie++	Crafty Chess Engine High	FLAC Audio Encoding	GZip Compression
3.	CPU Performance Metric	Memory to Performance Metric	DBench 128 Clients	Performance Linpack	Encoding Performance Metric	Compression Performance Metric
4.	John -The Ripper	RAM Speed	Disk I/O Performance Metric	Timed Apache Compilation	GnuPG	LZMA Compression
5.	Monkey Audio Encoding	Redis - Benchmark	Flexible I/O Tester	UNIX Bench - Parallel	LAME MP3 Encoding	Parallel BZIP2 Compression

These benchmark tests are run in all the shortlisted cloud service providers and the results are populated in the performance matrix. Now, the service catalog of the home cloud service provider compares the performance metrics agreed with the client in the service level agreement and the performance metrics of the foreign cloud service provider. The foreign cloud service provider which satisfies the performance parameters of the client is selected to provide services to the client. Figure 4 illustrates the selection phase.

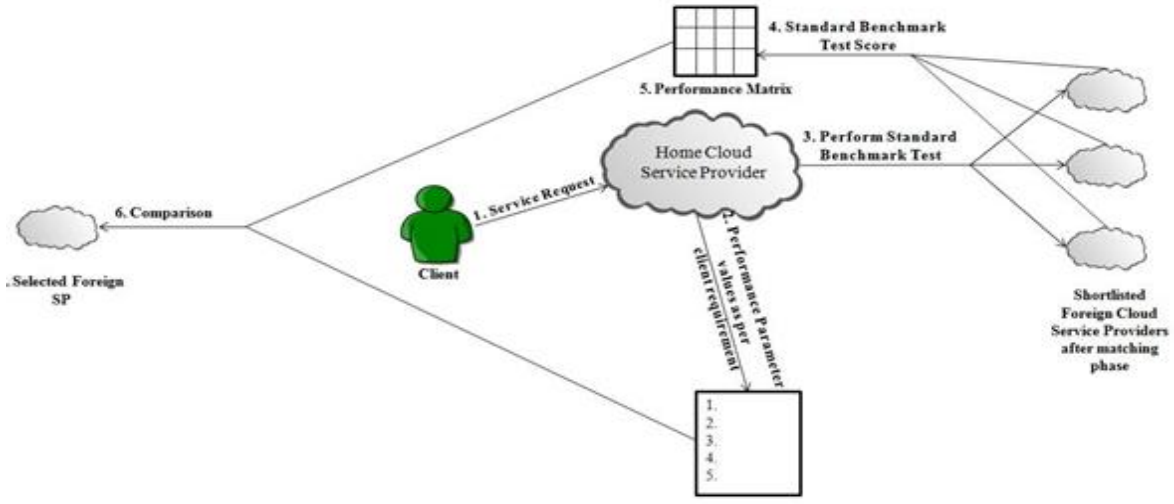

Figure 4. Selection of the optimal foreign cloud service provider by comparing their performance values.

Now let us consider the example where Online High Definition Video Encoding service is provided in the cloud. The necessary parameters for evaluating performance of this service are Multithreaded CPU Performance, Encoding, Disk I/O and Memory I/O. So appropriate benchmarks tests, as available in table 3,

are selected according to the client requirements by the home cloud service provider and their scores are compared. This is illustrated in Table 4.

Table 4. Performance Matrix: Test results for critical performance parameters of certain cloud service providers.

Cloud SP→ Performance Parameters(Benchmark)↓	Amazon AWS – EC2 (ec2-us-linux-m1.large)	Joyent Cloud (jy-us-west-16gb)	RackSpace (rs-16gb-us)	Microsoft Azure (az-16gb-us)
Multi-Threaded CPU Performance(C-Ray Tracer)	117.04	122.67	134.76	116.33
Memory I/O (GeekBench)	4185	3970	4420	4123
Disk I/O (Bonnie++)	20312	25143	23290.25	20115
Encoding(FLAC Audio Encoding)	18.43	13.64	13.26	13.98

Based on Figure 5, considering the standard benchmark test values and the relative position of each foreign cloud service providers in these test results, RackSpace may be selected as the foreign cloud service provider for the client with the Online High Definition Video Encoding service requirement.

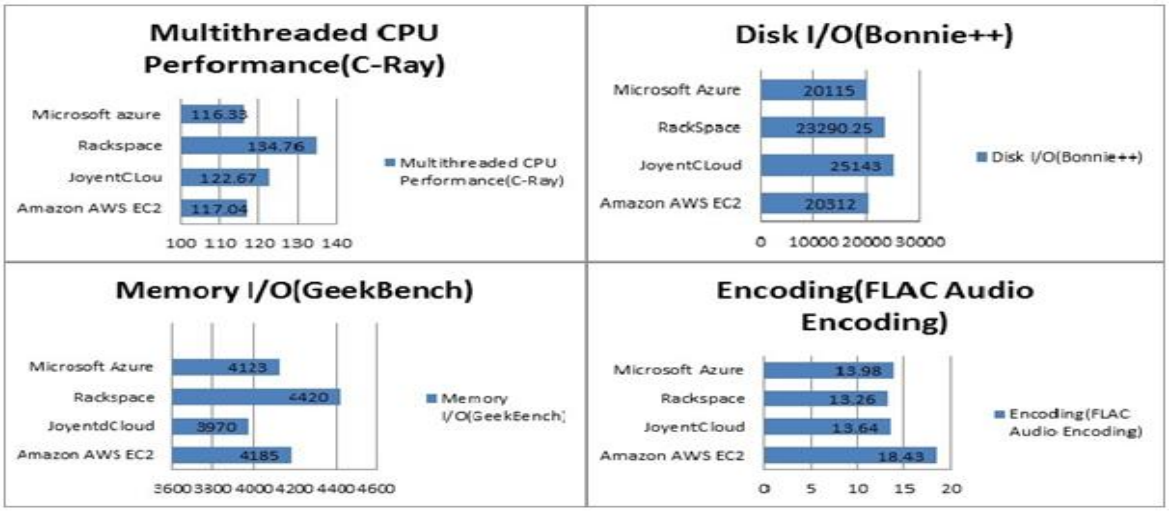

Figure 5. Performance Graphs depicting the results of standard benchmark tests.

In case of multiple foreign cloud service providers, eligible to provide services to the same client, a quantity called 'Utility Value' may be used. Utility value is maintained by the monitoring agent of each cloud service provider to assess the reliability of other cloud service providers, based on its past performance. Utility value is a counter incremented whenever a cloud service provider successfully uses the service of another cloud service provider. Higher the utility value, probability of the particular cloud service provider being selected is high. Even if the utility value is unable to determine the optimal cloud service provider, the home cloud service provider selects a provider at random from the set of eligible clouds.

Ideally, the client needs to be provided with uninterrupted services. However practically it is not possible. The term downtime is used to refer to periods when the service agreed in the service level agreement, is unavailable to the client. Maximum permissible downtime is the downtime which is feasible to the client without any serious negative business impact. The summation of the time taken for selection of foreign cloud service provider from the instant the home cloud service provider decides to go for cloud federation, and the time taken for the selected cloud service provider to provide the services together should be less than the maximum permissible downtime.

$$\Delta s+\Delta d<=MPD$$

Δs=Time taken for selection of foreign cloud service provider, Δd=Time taken for delay in providing service by the selected provider, MPD=Maximum permissible downtime. When Δs becomes higher than MPD, the client is alerted to impose the agreed penalty on home cloud service provider.

4.4. Usage

Once the foreign cloud service provider is selected, the services are rented by the home cloud service provider and delivered to the client. Now the home cloud service provider acts as a cloud broker between the client and the foreign cloud service provider. The services provided by the foreign cloud service provider are continuously monitored for its performance parameters by the home cloud service provider. If the performance parameters is not satisfying the service level agreements, a penalty chain is triggered which is paid by the foreign cloud service provider to the home cloud service provider and the home cloud service provider to the client.

4.5. Suspension

When the service manager and the service catalog of the home cloud service provider determine that there is no further requirement of services to a particular client, the connection path between the home cloud service provider and the foreign cloud service provider is suspended. All further client requests are routed to the home cloud service provider.

5. DISCUSSION AND FUTURE WORK

Selection of the optimal cloud in a cloud federation is a complex task requiring perusal of various parameters, in which our work is still in its early stages. Though we have considered capability and performance as our important consideration for matching and selection, cost remains another important parameter that needs to be integrated in this model to make it more comprehensive. Also automating the process of selection of foreign cloud also needs urgent attention.

Implementing federated cloud introduces various security issues. Federating the identity management is an integral part of providing convenience to the user and simplifying the interface. But federated identity management brings with it serious interoperability and trust issues.

In spite of the advantages, there are some issues which act as a bottleneck for large scale use of federated clouds in enterprises. These include non availability of technology neutral interfaces for communication between different cloud service providers. Also compliance with various legislation, regulation and policies proves to be a major challenge for cloud federations. Reduction of time delay or transferring the services from one service provider to another is also an area that requires improvement.

6. CONCLUSION

Cloud federation brings together services of different cloud service providers that can be tailored to match with the client's requirements. In this paper, we propose an approach for getting the optimal cloud service in the federated cloud computing environment. This is achieved by the selection of the optimal cloud service provider, by matching the capability and performance of the foreign cloud service providers with the requirements of the client. The purpose is to facilitate rapid and uninterrupted service, and efficient resource management for a federated cloud environment.

REFERENCES

[1] M.C. Mganga and N. R. Putri, Introduction, Enhancing Information Security in Cloud Computing Services using SLA Based Metrics, p.10, 2011.
[2] T. Grance and P. Mill, Definition of Cloud Computing, The NIST Definition of Cloud Computing, p.1, 2009.
[3] T. Grance and P. Mill, The NIST Definition of Cloud Computing, The NIST Definition of Cloud Computing, p.3, 2011.
[4] V. K. Chaurasiya, R. Gupta, A. A. Pinto, S. Singh, P. Srivastava, S. Verma, Problem of Security, An architecture based on proactive model for security in cloud computing, p.2, 2011.
[5] T. Harris, Understanding public and private clouds, Cloud Computing – An overview, p.4, 2009.
[6] Distributed Management Task Force Inc. [homepage on the Internet], Interoperable Clouds. Available from: http://www.dmtf.org/sites/default/files/standards/documents/DSP-IS0101_1.0.0.pdf, December 23, 2011.

[7] Global Inter-Cloud Technology Forum [homepage on the Internet], Use Cases and functional requirements for Intercloud computing. Available from: http://www.gictf.jp/doc/GICTF_Whitepaper_20100809.pdf, December 29, 2011.
[8] Cloud Harmony [homepage on the Internet], Benchmarks, Available from: http://www.cloudharmony.com/benchmarks, January 02, 2012.

BIBLIOGRAPHY OF AUTHORS

Saumitra Baleshwar Govil completed his graduation in 2010 at Gautam Budha Technical University and is pursuing his masters in cyber law and information security at Indian Institute of Information Technology, Allahabad. His research interest includes network security and database management.
Email: sbgovil@gmail.com

Karthik Thyagarajan completed his graduation in 2010 at Anna University and is pursuing his masters in cyber law and information security at Indian Institute of Information Technology, Allahabad. His research interest includes cloud computing security and Application security.
Email: karthik29nov@gmail.com

Karthikeyan Srinivasan completed his graduation in 2010 at Anna University and right now pursuing his masters in cyber law and information security at Indian Institute of Information Technology, Allahabad. His research interest includes cloud computing security and Identity Management.
Email: karthikeya777@gmail.com

Dr.Vijay K. Chaurasiya is working as an Assistant Professor of Information Technology in MBA (IT) & MS (CLIS) Division of Indian Institute of Information Technology, Allahabad. He received his Master of Technology Degree in Wireless Communication and Computing from Indian Institute of Information Technology, Allahabad in 2004 and Ph.D. in Information Technology from the same Institute in 2010. His research interest includes Wireless Sensor Networks, Computer Networks, Routing Protocols and Database Management System.
Email: vijay.chaurasiya@gmail.com

Dr.Santanu Das is working as an Lecturer of Information Technology in MBA (IT) & MS (Cyber Law and Information Security) Division of Indian Institute of Information Technology, Allahabad. He holds a MBA degree in Finance and has over eight years of teaching experience at BIT MESRA, Ranchi. His research interest include Accural, Financial and Merger Accounting.
Email: sdas@iiita.ac.in