Cloud Services in Different Cloud Deployment Models: An Overview

Tarun Karnwal
Dept.of Computer Science
Pondicherry University
Puducherry, India

T. Sivakumar
Dept.of Computer Science
Pondicherry University
Puducherry, India

G. Aghila
Dept.of Computer Science
Pondicherry University
Puducherry, India

ABSTRACT
Cloud service provider provides resource as a service to consumers. These resources may be storage resource or processing power etc. Cloud provides services as infrastructure as a service (IAAS), platform as a service (PAAS) and software as a service (SAAS). This paper brings an insight into the various services provided by cloud environment. It elaborates the different cloud deployment model, overview of how cloud consumers take services from different cloud models. Moreover different models have different techniques to provide service. The objective of this paper gives the detail view of different roles or phases required while consumer takes services in different cloud deployment models. In which each phase have specific purpose to describe itself. Each model explains the issues in its each phase and different methods and technique used by these phases.

Keywords
SLA, federated cloud computing, Intercloud computing, local cloud (LC), global cloud (GC);

1. INTRODUCTION
Cloud computing is a combination of distributed system, utility computing and grid computing. In cloud computing we use combination of all these three in virtualized manner. Cloud computing gives advanced facility like on demand, pay per use, dynamically scalable and efficient provisioning of resources.

Cloud computing provide three types of services: IAAS, PAAS, and SAAS.

In present cloud computing, clouds are deployed in four scenarios:

1) Private cloud: In this a local organization has its internal virtualized data center for using only within the organization. They make their data center virtualized by using open software like eucalyptus or open nimbus or xen.

2) Public cloud: This cloud can be used by anyone publically on payment basis. In this cloud services provide by companies like Amazon , Microsoft azure, Google appengine.

3) Community cloud: This is private federated cloud , in this some private cloud share their data but this data will be private in these organization only , nobody other than these organization cannot access this data. It means data or resource will not available publically for customers.

4) Hybrid cloud: This is public federated cloud in this some public clouds connect with each other in federated manner and share their resources.

Now days when a customer request for services from cloud then cloud provides its resources by using any one of these three scenario

1) Monolithic cloud: These are the homogeneous cloud. These are big cloud service providers which have large amount of resources and fulfill customer requirement by delivering these resources.

2) Automatic federated resource enhancement (intercloud): this has the combination of small and big cloud providers. In this scenario if the cloud unable to fulfill all the requirement of customer then cloud ask resources from other cloud. We will discuss about this scenario in level 5 of our architecture.

3) Virtual federated catalog system (intercloud): in virtual federated catalog system a catalog is used which is attached with multiple cloud providers. This catalog has information of all the resources in each cloud so broker in our architecture layer2 will directly contact with this catalog. Each Cloud provides different type of resources as processing power, storage power, different applications etc. we will use cloud services and cloud resources interchangeably.

So further we divided our whole cloud service in four layers. These layers are consumer, broker, federated catalog system (internal intercloud), federated cloud system (external intercloud).

2. CLOUD SERVICE ARCHITECTURE
In our architecture consumer take service either from private cloud provider or public cloud provider and then further these Cloud providers also distinguished on the basis of services providing by these different service providers

Private cloud divides in tow deployment model
(1) Single private cloud
(2) Community cloud

Public cloud divides in three deployment models
(1) Single public cloud
(2) External public intercloud
(3) Internal public intercloud

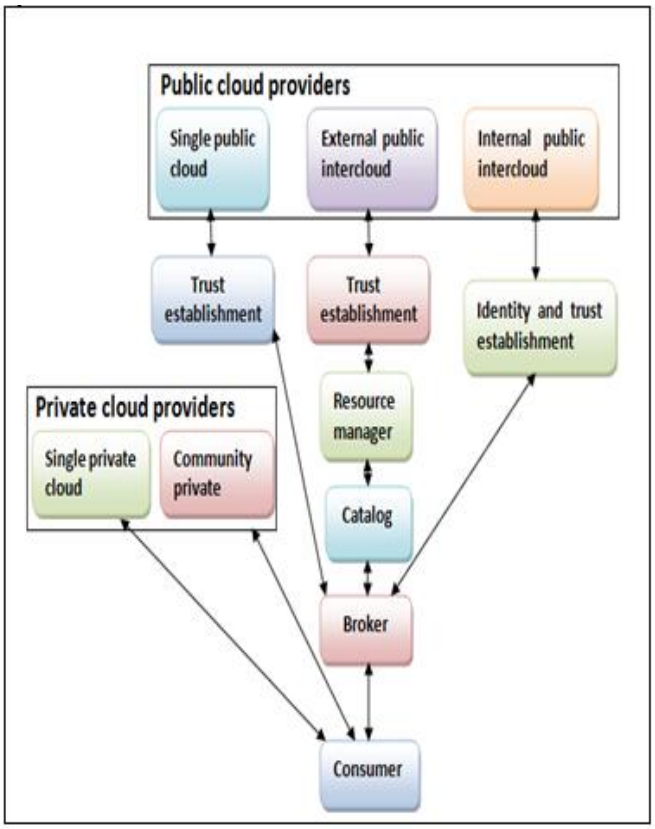

Figure 1: Cloud service architecture

3. USER OR CONSUMER

If a user will be use private cloud resources then he will be user only for that cloud provider but if that user will use public cloud resources then that user will be consumer for that public cloud because now user has to pay money to use public cloud resources. The consumer use public cloud resources on the payment basis. The consumer wants best Quality of Services (QoS) for his paying money. So there is need of a broker who will provide agreement between consumer and cloud service provide on the basis of price and QoS.

4. BROKER

Cloud computing is based on pay as use model. Cloud computing is an internet based application where resources are in distributed manner. So the consumer needs dynamic update of the resources available and also make negotiation for those resources dynamically on the basis of demand and payment.
So in this dynamic environment service level agreement (SLA) is finding a good approach for this negotiation and agreement between consumer and service provider.

The SLA first time introduced in 1989 when telecommunication industry given the services to its consumer on agreement basis. After that it gets involved in web services, network, internet and datacenters. In web services SLA take care of quality of services

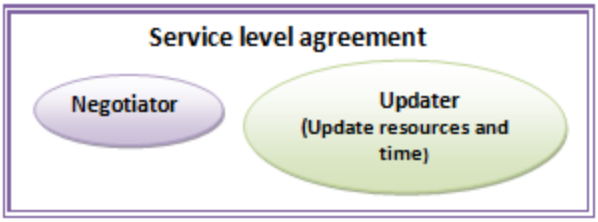

Figure 2: SLA

getting by consumer, in network it is used to take care whether it is working under decides agreement and in internet it check is there any agreement or negotiation violating.

SLA has three phases: Creation –operation- removal.

In traditional agreement problem we basically used SLA for a fully connected network, a broadcast network or a generalized network. But cloud computing is entirely different, it is internet based, dynamically scalable and virtualized. So because of all these features and to manage the cloud resources dynamically and on demand basis the SLA need some set of rules. These rules will not bind to consumer to take his decision on the basis of these rules but these rules will only suggest to the customer which service will be better for him.

SLA specification for web services mostly uses any of these two protocols: 1) Web service agreement specification [1] (WS-agreement). 2) Web service level agreement (WSLA) [5]
In 2004 [1] WS-agreement is provided by open grid form. It follows a certain name convention. In this SLA is dynamically negotiated and created by software. It checks the status of its consumer dynamically so by using web service agreement specification SLA manage any violation dynamically.
In 2007 [2] purposed a simple extension of WS-agreement protocol support for simple offer from service provider but it does not support for auction based negotiation customer and service provider.
In 2007 [3] says WS-agreement unable to solve limitation when service provider and customer have different standards, policy and languages during negotiation.
In 2008 [4] proposed a Meta negotiation architecture for SLA aware grid service and some architecture will extend in cloud computing.
In 2003 [5] WSLA framework is introduced by IBM which is using XML tags. By using WSLA framework IBM develop the SLA for web services. It is also dynamic in nature it means it can check SLA violation at run time and it can check for various (QOS) parameters dynamically.
In 2004 [6] firstly WSLA was using in grid computing, in 2009 [7] upend it in cloud computing on the basis of business model, architecture, resource management, programming model, application model and security model.
In 2005 [8] introduces a Rule Based Service Level Agreement (RBSLA) which follows knowledge based approach and use rule markup language [9] for describing SLA.

In 2010 [10] telling the use of SLA in cloud computing on the base of fault tolerance. They are proposing a Dual Agreement Protocol of Cloud Computing (DAPCC). DAPCC divides the SLA in two layers, layer A and layer B. Layer A has some nodes which are directly attached with customer and layer B have set of nodes of same type which are directly attached with layer A nodes. DAPCC gets agreement as a common value among all nodes. The DAPCC try to make common value with minimum number of message exchange and which can tolerate maximum number of allowable faulty component.

In 2010 [11] core resource broker is introduces in grid computing ,it tells the traditional resource broker do not have the good policy management and the time complexity is high because three use cases scheduling is taking place in round robin. So to decrease time complexity they replace this round robin algorithm with new division based scheduling algorithm. they also introducing image repository for pre configured virtual machine(VM) image and VM metadata store in XML tag, so if new request comes for virtual machine image it firstly check in image repository which reduces time complexity ,in future we can join this core resource broker with SLA in cloud computing.

In 2010 [12] consider the computer system security. They introduce a Secure Service Level Agreement (Sec-SLA) take consideration of security related metrics where traditional approach have telecommunication metrics such as throughput, delay, packet loss etc.

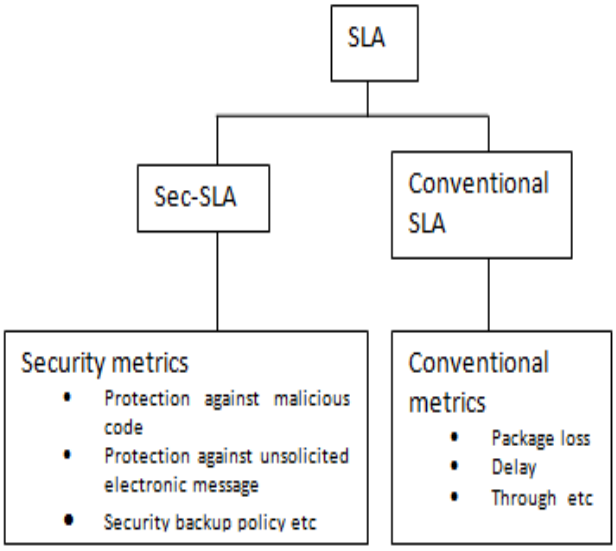

Figure 3: SLA-Classification [12]

SLA will be is more effective in web service policy(WS policy) from the World Wide Web Consortium (W3C) because by using W3C service provider and consumer both can publish their policy.

In 2011 [13] introduces aggregation property in SLA. Aggregation is needed because in cloud computing sometimes the service providers do not want to show all his services to the customer. Service provider want only to show only some of them which ever required by customer. So this paper introduces an aggregation model based on SLA views that will make the automatic hierarchal aggregation of SLA model.

5. FEDERATED CATALOG SYSTEM (EXTERNAL INTERCLOUD)

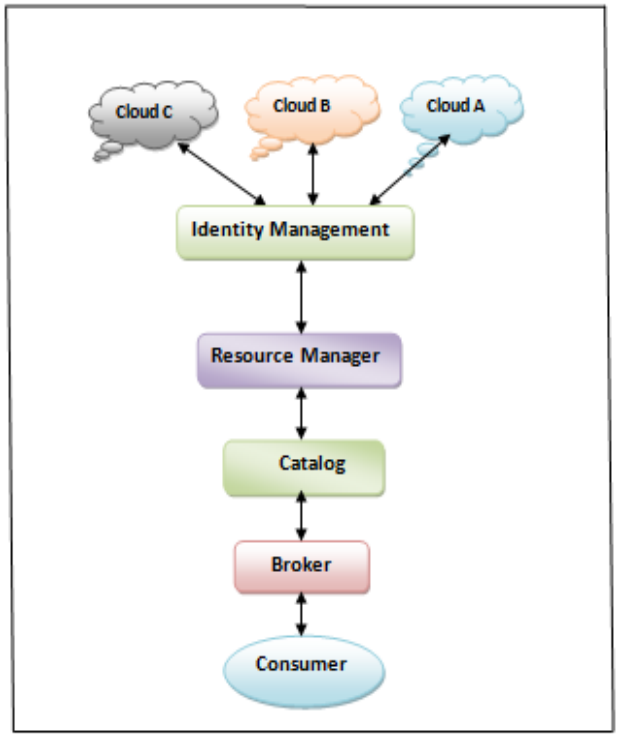

Figure 4: Federated catalog system architecture

This approach gives the most efficient, scalable and optimal provisioning of resources. To make this federated catalog we need a systematic process. We divide this process in four phases:

a) First phase: In this phase federated system need to verify the identity of cloud providers for authentication.
b) Second phase**:** Discovering different services provide by public cloud provider and updating catalog time dynamically.
c) Third phase**:** Maintain different public cloud resources information in catalog and schedule them on the basis of different parameter. These parameters called templates for different-different consumers and these templates are highly confidential and secure.
d) Fourth phase**:** It is broker which work on the basis of SLA.

5.1 First Phase

For maintaining information of resources providing by inter cloud environment firstly need to identify available clouds. Identify the cloud with different identity providers (Idp).This Idp has the key of group of trusted clouds. When a cloud or group of clouds wants to register its resources in resource

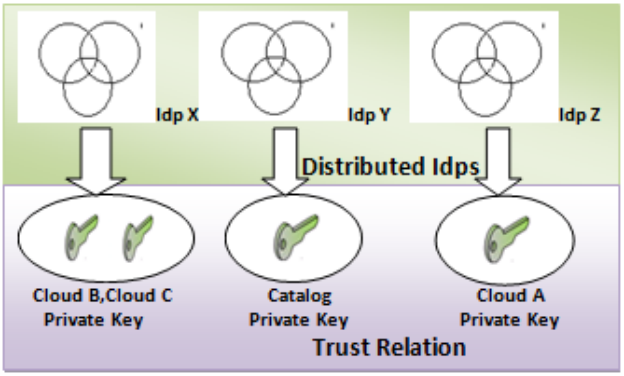

Figure 5: Identity management

or resource manager want to search available public cloud then that cloud will share its key with IDP of resource catalog. By sharing key these clouds will come under the trust of resource catalog.

5.2 Second Phase

Before maintaining the different resources providing by inter cloud environment in resource catalog firstly we need to know what are the resources providing by this intercloud environment and further need to order these resources to put same type of resources at same location in the catalog, which make customer easy to find required resources.

To order these resources need to find similar type of resources providing by different cloud providers and normalize them to maintain in catalog. To find similar type approach in inter cloud environment we can use semantic web [16]. Semantic Web provides meaning and relatedness to object so it is a good approach for our resource matching problem. Before entering these matched cloud resource information in resource catalog, need to provide normalize meaning to them. This catalog will have pricing information and trust level information also for each cloud resources. The pricing information will tell how much amount needs to pay for use any particular resource.

In 2010 [17] Raj kumar Buyya et al. used cloud coordinator to matching the resources in external federated cloud environment. The cloud coordinator used domain specific cloud .The Cloud Coordinator exports the services of a cloud to the federation by implementing basic functionalities for resource management such as scheduling, allocation,(workload and performance) models, market enabling, virtualization, dynamic sensing/monitoring, discovery, and application composition

In [18] Moreno et al. providing external intercloud by using multi cloud cluster. They implemented Sun Grid Engine (SGE) cluster. This cluster has local worker node, front end server node and several remote worker nodes. They are using three clouds Amazon EC2 (USA), Amazon EC2 (Europe) and Elastic host.

Front-end server is the master host and manages all other node. The remote worker nodes are deployed in different clouds. When a customer request for any resource then first it goes to front-end server and it assign the request to that remote worker node which gives good performance.

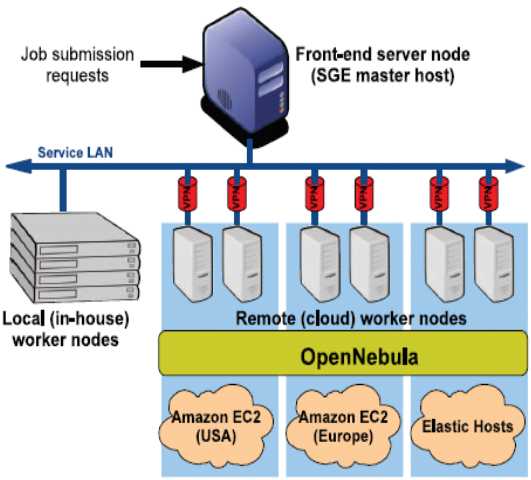

Figure 6: Experimental framework [18]

In figure: 4 federated cloud system architecture using resource manager (RM). RM will monitor and discover cloud resources and maintain them in catalog. It has six elements as given in figure: 7.

5.2.1 Discover and Update:

RM will first discover the different new public cloud available and different resources providing by those cloud or any update in already discovered public cloud resources and then add those new cloud resources or update already discovered cloud resources in catalog.

5.2.2 Cloud Analyzer:

RM will calculate each cloud each resource performance on the basis of QoS and turnaround time etc and check its prior success rate.

5.2.3 Time Management:

For each resource it give the time limit to be in catalog and calculate the trust level index for that time limit and after completion of that time limit it again update the resource status and its trust index. The trust level is necessary because A cloud provides different resources and it is not necessary that if it providing one good resource then other resources also will be good for example if one cloud providing good storage resource then it need not be necessary that its processing resource also will be good. This problem becomes big when we talk about inter cloud environment. So there is need to distinguish different type of resources in each cloud. Here we are using trust level indexing [15] to distinguish them. These trust level gives the level of trust (as 70%, 80% etc) by using fuzzy logic based trust algorithm. This trust algorithm works on some specific parameters like QoS etc

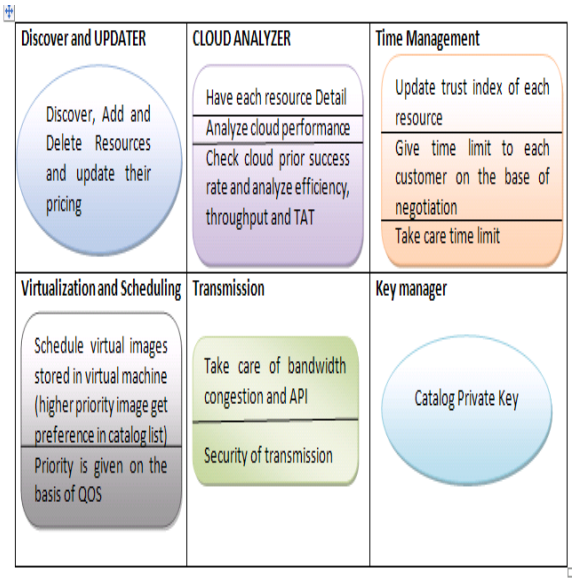

Figure 7: Resource manager (RM)

5.2.4 Virtualization and Scheduling:

Each service given to consumer in the form of virtual images. Virtual images are stored in virtual machines so on the bases of QoS we schedule each resource virtual image.

5.2.5 Transmission:

This will take care for data translation security from cloud to consumer. it will work as a third party for transmission application program interface security. it will monitor VPN and take care man in middle attack, external malicious attack like DDoS or any data leakage and provides firewall if necessary.

5.2.6 Key Manager:

RM check public cloud identity trust by providing catalog key to the identity service provider (IDP) of that public cloud

5.3 Third phase:

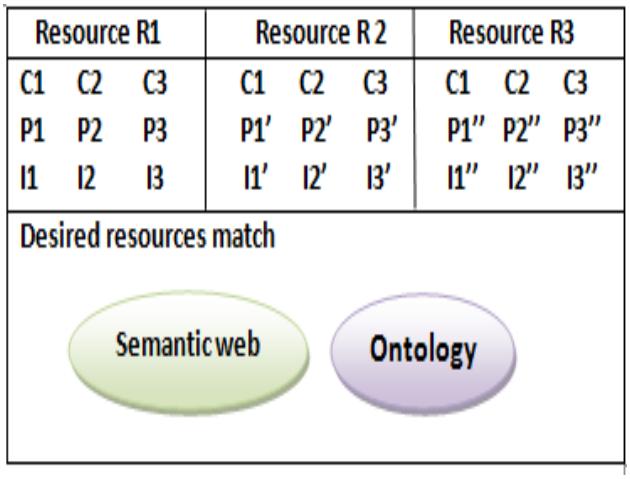

Figure 8: Catalog

For each public cloud each resource type information maintained in catalog. This catalog maintains basically two parameter of each cloud. The first parameter is each resource price and second parameter is each resource trust index.

In fig: 8 we are having three public clouds as C1, C2 and C3. Each cloud providing three type of resources R1,R2 and R3.for each type of resources cloud maintaining different prices and trust indexes as for resource R1 prices are P1,P2,P3 and trust indexes are I1,I2,I3 similarly for resource R2 prices are P1',P2',P3' and trust indexes are I1',I2',I3' and similarly for resource R3 prices are P1'',P2'',P3'' and trust indexes are I1'',I2'',I3''.

The catalog uses semantic web for choosing consumer desired resource among all available resources on the basis of trust level and price and ontology use to normalize different type of services provided by public cloud.

6. FEDERATED CLOUD SYSTEM (INTERNAL INTERCLOUD):

As in external intercloud we need to maintain an external catalog system. In internal intercloud we need not maintain any resource catalog system. In internal intercloud if a public cloud resource provider (here we called this cloud as local cloud) unable to fulfill the requirement of its customer then that public cloud provider ask for needed resources from another public cloud provider (here we called this type cloud as global cloud) and provide these resources to its customer continuesly without any interruption and without knowing by the customer about actual service provider or global cloud. In internal intercloud the local cloud make the virtual connection with global cloud and in this virtual connection the local cloud seems the global cloud resources as its own resources. In this internel federated cloud system local cloud provider will autometically ask for services from global cloud provider where customer will get unintruptted service without knowing background process. it work on Single Sign On(SSO) approach. In this approach the cloud which is directly attached by customer is called local cloud and other clouds which gives services to Local Cloud(LC) is called Global Cloud(GC).

Internal intercloud have 4 phases

a)First phase: Establishing trust between different cloud.
b)Second phase: Discovery of desired resources.
c)Third phase: Migration of resources between clouds
d)Fourth phase: Security of transmission data

6.1 Trust Establishment

In this phase local cloud identify the other global cloud to get its resources.this phase needed for authentication and autorization of our local cloud in global cloud network.in internal intercloud the identity management can done by any one of these two way:

6.1.1 Single Sign On(SSO)

In SSO local cloud needs only once to auhenticate itself with global cloud by using trust establishment and further local cloud will get directly access is any global cloud which is under trust level on the basis of trust context. For trust establishment it used intercloud identity management[14]. They provide identity to cloud by using IDP.

6.1.2 Third Party

In this third party will take care for identiry management of local cloud. The third party will provide digital identity to global cloud each time to establish connection with global clouds.the literature says to maintain and manage of these digital identity by the third party for each connection is very complicated. So according to us single sign on using IDP is better approach.

6.2 Resources discovery

Once the identity of cloud completed now need to discover required resources in GC. In [19] provide the three phase approach.they are using secure assertion markup language which will do two work together first it provide matching of resource and second it provide them authenticity. In [15] they are using extensible acccess control markup language (XACML) by OASIS for providing matching and access control

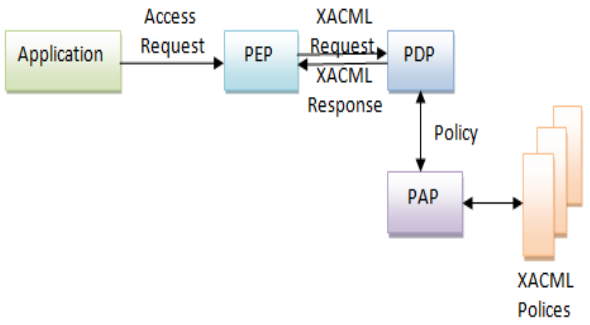

Fig 9: OASIS XACML processing environment [15]

Here policy enforcement point(PEP) will send the consumer access request to policy decision point (PDP) to decide whether access should be given or not by using XACML message. PDP decide on access permission on the basis of authorized policies.so a policy administration point (PAP) is used to get to the policies the PDP uses the PAP where policies are authored and stored in an appropriate repository.

6.3 Resource Migration

In clouds data is in the form of virtual images which are store on heterogenious virtual machines. These virtual images have all the resources as system kernal, separate operating system, processor and user data so when any consumer demand for any resource then inter cloud gives its resources to consumer on some policy basis but transformation of highly confidential data in the network need high efficiency, throughput, bandwidth and security. So [20] given a nobel approach to improve virtual machine migration in federated cloud enviornment.they are using separate migration agent for transfering the data in network.

6.4 Transmission Security

Transmissinon security needed to take care if any data leakage or any malicious attack like man in middle attack, DDoS attack etc. In [21] they are using cloud traceback to detecting the X-DDoS attack and using cloud protector to remove that attack.

7. CONCLUSION AND FUTURE WORK

This paper tells about different ways of providing resources to consumer or user from cloud. Firstly paper distinguish between private cloud and public cloud and further public cloud divides in three parts single public cloud, external public intercloud and internal public intercloud. Paper described about the different methods, technique and roles using in each phase by single public cloud, external public intercloud and internal public intercloud.

For future we are working on our external intercloud architecture and trying to provide different solution for the security of these internal intercloud and external intercloud.

8. REFERENCES

[1] Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Pruyne, J., Rofrano, J., Tuecke, S., Xu, M.: Web services agreement specification (WS-Agreement). In: Global Grid Forum. (2004).

[2] Andrieux, A., Czajkowski, K., Dan, A., Keahey, K., Ludwig, H., Nakata, T., Pruyne, J., Rofrano, J., Tuecke, S., & Xu, M. (2007). Web Services Agreement Specification (WS-Agreement). OGF proposed recommendation (GFD.107).

[3] Antoine Pichot, Philipp Wieder, Oliver W¨aldrich, Wolfgang Ziegler(2007). Dynamic SLA-negotiation based on WS-Agreement, CoreGRID TR-0082,June 24, 2007.

[4] Brandic, I, Venugopa S., Mattess, M., & Buyya, R. (2008). Towards a Meta-negotiation Architecture for SLA-Aware Grid Services. *International Workshop on Service-Oriented Engineering and Optimization*, (pp. 17). Bangalore, India.

[5] Ludwig, H., Keller, A., Dan, A., King, R., Franck, R.: Web service level agreement (WSLA) language speci_cation. IBM Corporation (2003).

[6] Chuan He; Lei Gu; Bin Du; Zhenchun Huang Sanli Li; , "A WSLA-based monitoring system for grid service - GSMon," Services Computing, 2004. (SCC 2004). Proceedings. 2004 IEEE International Conference on , vol., no., pp. 596- 599, 15-18 Sept. 2004.

[7] Pankesh Patel, Ajith Ranabahu, Amit Sheth(2009). Service Level Agreement in Cloud , Computing Cloud Workshops at OOPSLA09.

[8] Paschke, A.; , "RBSLA A declarative Rule-based Service Level Agreement Language based on RuleML," Computational Intelligence for Modelling, Control and

Automation, 2005 and International Conference on Intelligent Agents, Web Technologies and Internet Commerce, International Conference on , vol.2, no., pp.308-314, 28-30 Nov. 2005.

[9] Boley, H., Tabet, S., Wagner, G.: Design rationale of ruleml: A markup languagefor semantic web rules. (2001) 381-401.

[10] Shun-Sheng Wang, Kuo-Qin Yan, Shu-Ching Wan "Achieving efficient agreement within a dual-failure cloud-computing environment",Expert Systems with Applications, 38 (2011) Elsevier journal on, 906–915,2011.

[11] P. Balakrishnan , Thamarai Selvi Somasundaram " SLA enabled CARE resource broker" , Future Generation Computer Systems 27 (2011) Elsevier journal on ,265–279,2011.

[12] De Chaves, S.A.; Westphall, C.B.; Lamin, F.R.; , "SLA Perspective in Security Management for Cloud Computing," Networking and Services (ICNS), 2010 Sixth International Conference on , vol., no., pp.212-217, 7-13 March 2010.

[13] Irfan Ul Haq, Altaf Ahmad Huqqani, Erich Schikuta(2011). Hierarchical aggregation of Service Level Agreements, Elsevier journal on Data & Knowledge Engineering 70 (2011) 435–447.

[14] Celesti, A.; Tusa, F.; Villari, M.; Puliafito, A.; , "Security and Cloud Computing: InterCloud Identity Management Infrastructure," Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE), 2010 19th IEEE International Workshop on , vol., no., pp.263-265, 28-30 June 2010.

[15] Bernstein, D.; Vij, D.; "Intercloud Security Considerations," Cloud Computing Technology and Science (CloudCom), 2010 IEEE Second International Conference on , vol., no., pp.537-544, Nov. 30 2010-Dec. 3 2010.

[16] W3C Semantic Web Activity, http://www.w3.org/2001/sw/.

[17] Rajkumar Buyya, Rajiv Ranjan, and Rodrigo N. Calheiros(2010). InterCloud: Utility-Oriented Federation of Cloud Computing Environments for Scaling of Application Services, Springer Verlag Berlin Heidelberg , C.-H. Hsu et al. (Eds.): ICA3PP 2010, Part I, LNCS 6081, pp., Melbourne, Australia , 13–31, 2010.

[18] Moreno-Vozmediano, R.; Montero, R.S.; Llorente, I.M.; , "Multicloud Deployment of Computing Clusters for Loosely Coupled MTC Applications," Parallel and Distributed Systems, IEEE Transactions on , vol.22, no.6, pp.924-930, June 2011.

[19] Celesti, A.; Tusa, F.; Villari, M.; Puliafito, A.; , "Three-Phase Cross-Cloud Federation Model: The Cloud SSO Authentication," Advances in Future Internet (AFIN), 2010 Second International Conference on , vol., no., pp.94-101, 18-25 July 2010.

[20] Celesti, A.; Tusa, F.; Villari, M.; Puliafito, A.; , "Improving Virtual Machine Migration in Federated Cloud Environments," Evolving Internet (INTERNET), 2010 Second International Conference on , vol., no., pp.61-67, 20-25 Sept. 2010.

[21] Ashley Chonka,YangXiang n, WanleiZhou, AlessioBonti(2011), "Cloud security defense to protect cloud computing against HTTP-DoS and XML-DoS attacks" Network and Computer Applications 34 (2011) 1097–1107.