Cloud computing, has become gaining popularity because of pay-per-usage model and also by preventing the need for users to own and maintain complex computing infrastructure. However, due to their inherent complexity and large scale, production cloud computing systems are prone to various runtime problems that can be caused by hardware and software failures. Typically in a cloud environment numerous workloads deployed and running parallel to satisfy SLA of the end user requirements. In case if there is misconfigurations while deploying the new workload will impact the workloads SLA as well and identifying the impacted workloads that may disrupt because of new template stack provisioning/deprovisioning, it is a tedious task for the end user. In this paper, we present a cloud impact analysis framework with mechanisms to identify the workloads that will be impacted. We design and analyze the correlation of various workloads that will be impacted by provisioning the infrastructure through templates in a cloud environment.

1. Introduction
Adoption of Cloud Computing increasing rapidly and constantly most of the IT organizations are moving towards Cloud Computing instead of traditional data center approaches. Cloud Computing [1] has been a new paradigm because of the increasing demand of adoption by Enterprises and Small/Medium scale Business (SMB). SMB doesn’t need to have the datacentre in-house and the
infrastructure (resources) can be provisioned by Cloud Service Provider. And the number of Cloud Service Providers (CSP) increasing rapidly and this will enable the end customers to build the applications rather than bothering about infrastructure or datacentre or real estate space. This enables pay-per-usage model and facilitates ease of managing datacentres. Cloud Service Providers are continuously adding new infrastructure to their datacentres to meet the needs of the workload request of the customers. One of the major challenges for Cloud Administrator is to identify the impacted workloads that may disrupt, because of the new cloud template orchestration. For onboarding the infrastructure in a Cloud, the Cloud administrator have to understand of different hardware configurations and features and also to identify which workflows will be impacted in case of failure or success, which is a tedious and difficult task from the administrator point of view.

Templates are the configuration details of resources to be provisioned, but the template configurations will not be bound to specific hardware vendor and are generic. Templates ease orchestration of the cloud resources provisioning/deprovisioning. A template provide the configuration details from the end user perspective to the Cloud Administrator but doesn’t provide which infrastructure the work load request to be provisioned/ deprovisioned. From the Cloud Administrator identifying the existing workloads that might be impacted because of new workload provisioning through template orchestration. We are proposing cloud impact analysis framework that will be deployed as a separate service part of OpenStack Cloud computing services.

2 Open Source Cloud Computing

Cloud Computing is all about providing compute, storage and network for end customers in a distributed environment (Smit M, Pawluk P, Simmons B, Litoiu, 2012) where the actual software runs as a service on reliable and scalable servers rather than on each end users datacenter or environment. Cloud computing can refer to a lot of different things, but generally running different items "as a service" software, platforms, and infrastructure. Open source community driving the Cloud Computing through different kind of Management platforms (Buyya R, Yeo CS, Venugopal S, Broberg J, Brandic I, 2009) such as Openstack (http://www.openstack.org), CloudStack, Cloud Foundry, Open Shift, etc. We have considered Openstack as the Cloud management platform, since Openstack community contributing a large way and most of the companies; end users are deploying the Openstack services. Details of Openstack are described as below.

Openstack [4] (http://www.openstack.org) is a opensource Cloud management Service which facilitates different types of services for managing and which are easy to plugin with others service. Figure-1 Openstack deployment diagram provides overview how the cloud resources will be managed. Since Openstack provides different type of deployment methodologies devstack (which is for developers can run as a sandbox for developing the services without need of physical hardware) and Openstack management software can be deployed as highly available services on physical servers as well. Openstack consists of below services for managing the cloud:

- **Heat (HOT) [5]** (https://wiki.openstack.org/wiki/Heat) is a orchestration service of Openstack which provides a template based provisioning and the template supports heterogeneous resources configuration stored in the standard format YAML or JSON or XML format. HOT also known as Heat Orchestration Template which provisions resources invoking other services such as Nova, Neutron, Cinder, Swift and Glance. But heat (HOT) engine does not
recommended resources for the given template and the cloud administrator has to identify the resources manually for template deployment.

- **Nova** is the major computing engine of OpenStack. Nova is mainly for deploying and managing large numbers of virtual machines and other instances to handle computing tasks.
- **Swift** is a storage service for provisioning and managing object storage in Openstack.
- **Cinder** is a service for managing SAN storage typically block storage management of SAN arrays. **Neutron** is for managing different type of networks between the computes, storage and also cloud management software.
- **Keystone** is a service for authenticating and authorization of cloud users.

![Openstack Cloud Deployment Diagram](image)

**Figure 1: Openstack Cloud Deployment**

Typically in a cloud environment numerous workloads deployed and running parallel to satisfy SLA of the end user requirements. Fig-2 depicts the pictorial view of workloads running on a compute provisioned to storage.

![Typical workloads running in a cloud](image)

**Figure 2 – Typical workloads running in a cloud**
2.1 Template

Templates contain the data related to the configurations gathered from existing system or snapshots of existing templates or predefined templates. Typically templates are represented in json or xml format for preserving the data. But the data present in templates are generic doesn’t interrelated to any resources in the cloud. Below is the sample data of the templates for virtual machine in yaml format:

A simple Heat template that spins instances and a private network.

```
{  
heat_template_version: v1  
description: >  
"TemplateFormatVersion" : "2014-09-09",  
"Description" : "CloudFormation Sample  
Template for instance creation"  
"Parameters" : {  
  "KeyName" : {  
    "Description" : "Name of an existing EC2 KeyPair to enable SSH access to the instances",  
    "Type" : "String"  
  },  
  instance1:
    type: OS::Nova::Server  
    properties:  
      name: heat-instance-02  
      image: 01b0eb5d-14ae-4c9e  
      flavor: m1.xsmall  
    networks:  
      - port: {  
          name:"template"  
          type_resource: "VM",  
          description: {  
            "VM provisioning"  
          }  
        }  
    features: {  
      details: {  
        os-type: ‘debian,  
        version: ’6.5,  
      }  
    proc:{  
      type: ‘x86’,  
      proc: ‘distributed,  
      required_units:2.0,  
      max:4.0,  
    }
```
mem: {
    mem: ‘distributed
    required_units:2.0
    maxunits:4.0
}

storage: {
    volume_type: ‘external’
    sizeoflun:’20GB’,
    mapped: ‘enabled’
}

adapters: {
    networkadapters{
        number 1
        speed: 5 mbps
    }
}

3 Cloud Impact Analyser Service workflow

From Fig-2, if we observe multiple workload are parallel running in a cloud environment, eg:

1) SAP Workflow running on VM-SAP on Server-1 on Storage LUN-1
2) HR Workflow running on VM-HR on Cluster-1 on Storage LUN-2
3) Apache Webserver running on VM-WebApp on Server-1 on Storage LUN-3
4) MailServer running on VM-MailServer on Cluster-1 on Storage LUN-2.

For example, if the user want to deploy a new template which has configuration to deploy workload for “websphere” application on Server-1, then deployment of new template should not disrupt the existing workloads and the administrator should be able to identify the workloads that will be impacted.

Below is the workflow [6] for cloud impact analyser framework:

1) User chooses the templates and chooses the resources to provisioned
2) Cloud Impact analyser gathers the details of the resources to be provisioned and template details.
3) Computing the topology map for compute resources, network and cinder services.
   a. Get the compute details from Nova service using the resource uuid.
   b. Find the associated networks from the neutron services for compute based on uuid.
   c. Find the associated of cinder volumes for the given compute from cinder service.
   d. Build the topology map
4) Cloud Impact analyser (CIA) builds the Map from the step-3 such as the key is resource-id and the value will be a lists of object that comprises the Path-Tree with metadata.

<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>25c4ab5c-646d-</td>
<td>VM-SAP ➔ Server-1 ➔</td>
<td>Metadata contains multiple attributes along</td>
</tr>
<tr>
<td>11e5-9d70-</td>
<td>Zone1 ➔ Storage1 ➔</td>
<td>with application type</td>
</tr>
<tr>
<td>feff819c9c9f</td>
<td>LUN1</td>
<td></td>
</tr>
<tr>
<td>93a805ac-646d-</td>
<td>VM-HR ➔ Cluster-1 ➔</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Zone 3</td>
<td></td>
</tr>
</tbody>
</table>

Table -1: lists of object that comprises the Path-Tree with metadata
Cloud Impact Analyser Framework For Identifying Workloads Impacted Due To Stack Orchestration

5) CIA service reads the templates meta data and reads the application meta data.
6) CIA iterates through the Map built in Step 3 to identify the list of the applications.
7) CIA builds the topology tree of applications workloads that will be impacted because of the new template workload provision request.

![Diagram of impacted workloads]

Figure 3 – Impacted workloads are highlight.

Figure-3 depicts the workloads that will be impacted running on Server-1 in case if the new template provision request fail or disrupt because of any misconfigurations.

3.1 Proposed Architecture

Below is the proposed architecture to find the impacted workloads, we are proposing a new service CIA [7] service which interacts with Nova, Cinder and Neutron service [8] to find the topology map and identify the impacted workloads [9].
4. Results
We have tried proposed approach by devstack deployment and experimented on the commodity based servers mostly KVM based hypervisors. We tried to find the impacted resources for the new Workload deployment on the infrastructure server-1 and from the Cloud Impact Analyser service provides the detailed outcome of the list of the workloads that might be impacted.

Figure 4 – CIA is the service interacts with Openstack services

Figure 5 – Outcome of the impacted resource
5. References


Biographies

Swaroop Jayanthi is currently a research scholar in the Department of Information Technology in GITAM University, Visakhapatnam. His area of research activities are Cloud and Data Mining.

Padma Poosapati is a Associate Professor in Department of Information Technology at GITAM University, Visakhapatnam, Andhra Pradesh, India. She received her Master degree in Computer Science and Engineering from Andhra University in 1999 and PhD degree in Computer Science and Engineering from Andhra University 2010. Her current research interests include Cloud Computing, Clustering and Classification in data mining and graph mining.