

## 0.1 Ansible

Ansible is a widely popular open-source tool used for automation of configuration management, application deployment. Ansible is popular because of its simplicity. Originally, Ansible Inc. was setup to manage the product. Later in 2015, RedHat acquired Ansible. “It uses no agents and no additional custom security infrastructure, so it’s easy to deploy - and most importantly, it uses a very simple language (YAML, in the form of Ansible Playbooks) that allow you to describe your automation jobs in a way that approaches plain English” [hid-sp18-417-doc-Ansible]. An user doesn’t have to learn a cryptic language to use it. As no agents are required to be installed in the nodes, the tool eases the network overhead. Ansible may use two kinds of server for operation. One is the controlling server that has Ansible installed. The controlling server deploys modules in the nodes through SSH channel. The basic component of Ansible architecture are:

**Modules:** This is the unit of work/task in Ansible. It can be written in any standard programming language

**Inventory:** Inventory is basically the nodes used

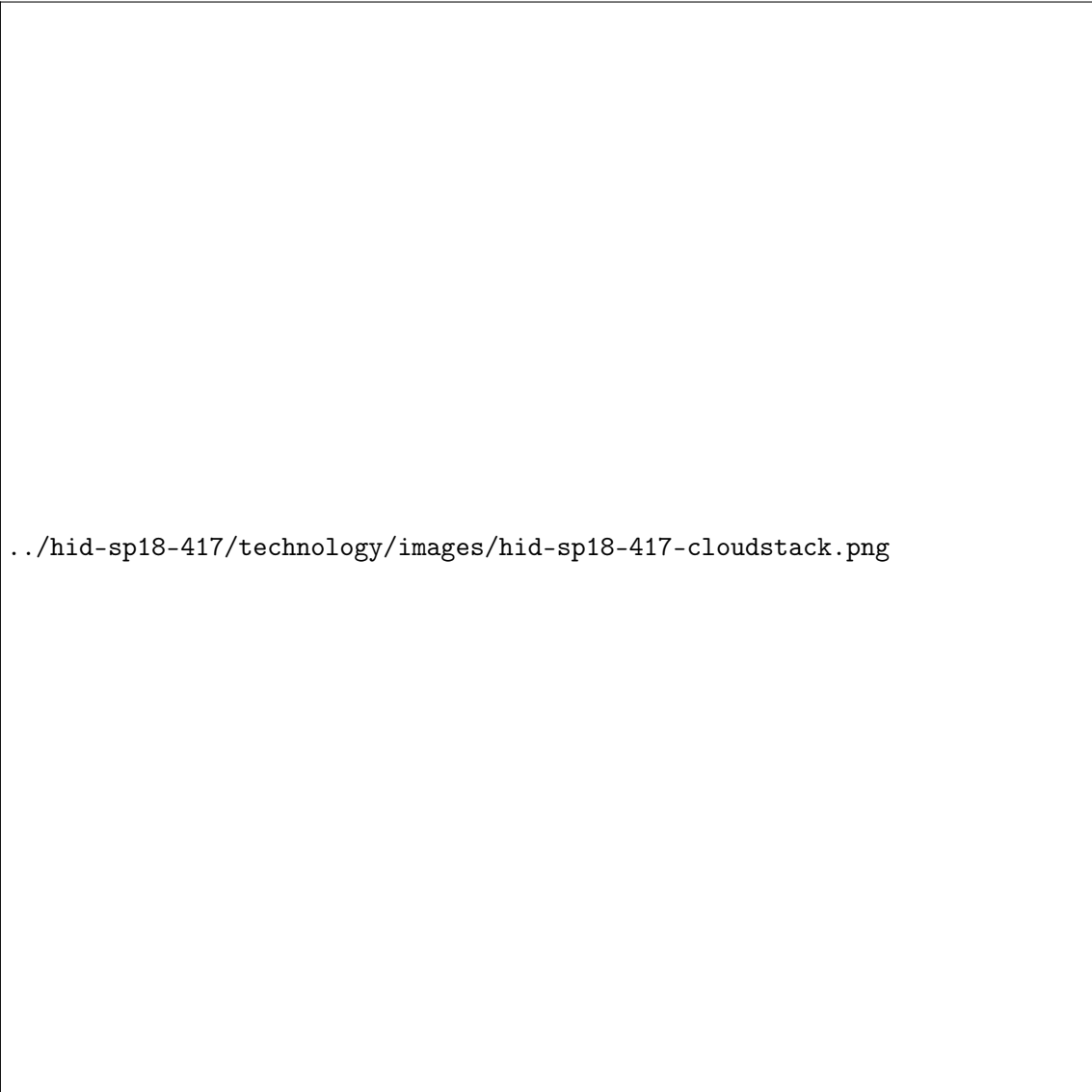
**Playbooks:** A play book in Ansible describes in simple language the infrastructure used for the deployment of the tool. This is written in YAML.

## 0.2 Apache CloudStack

Apache CloudStack is an open source that provides a highly scalable and available cloud management platform for IT Enterprises and service providers. CloudStack was originally developed by Cloud.com and was known by the name VMops. In 2011, Citrix acquired the product and donated it to Apache. “CloudStack is being developed to help managed service providers and enterprise IT departments create and operate public cloud, private cloud or hybrid clouds with capabilities equivalent to Amazon’s Elastic Compute Cloud(Amazon EC2) It uses existing hypervisors such as KVM, VMware ESXi/VMware vcenter and XenServer/XCP for virtualization. In addition to its own API, CloudStack also supports the Amazon Web Services(AWS) API[3] and the Open Cloud Computing Interface from the Open Grid Forum.” [hid-sp18-417-wiki-cloudStack]. The key feature of the product are s (1) high availability of resources s (2) network management s (3) provides GUI for ease of management s (4) compatible with most of the hypervisor/virtual monitor s (5) it provides the snapshot management. e.g. This feature is very useful is saving a state[snapshot] of a virtual machine. The vm can later be reverted to the stored state. The basic deployment of CloudStack just needs two machines: A server and a hypervisor that is a monitoring system. The process can be over simplified by configuring one machine to serve both the purpose. The same simple system can easily be scaled to a zone or a pod. Figure ?? depicts how the simplest deployment infrastructure can be scaled to provide an advanced support system.

## 0.3 Apache Delta Cloud

Apache DeltaCloud was developed in collaboration between Apache Foundation and Redhat to provide a programming application that will facilitate management of different cloud interfaces and It was supporting all the major cloud interfaces. “Each Infrastructure-as-a-Service cloud existing today[when?] provides its own API. The purpose of Deltacloud is to provide one unified REST-based API that can be used to manage services on any cloud. Each particular cloud is controlled through an adapter called a *driver*. As of June 2012, drivers exist for the following cloud platforms: Amazon EC2, Fujitsu Global Cloud Platform, GoGrid, OpenNebula, Rackspace, RHEV-M, RimuHosting, Terremark and VMware vCloud” [hid-sp18-417-wiki-deltacloud].



../hid-sp18-417/technology/images/hid-sp18-417-cloudstack.png

Figure 1: CloudStack Scalability [**hid-sp18-417-cloudstack-scaling**]

In 2009, DeltaCloud was developed for the purpose of providing one unified API for the major cloud service.

In 2011, it became a part of the Apache's top level project.

Unfortunately, in 2015 the project was closed due to inactivity. The application though inactive is chosen for the study to understand the case behind the termination of the project. It is primarily because of lack of popularity RedHat withdrew the sponsorship ultimately resulting in the termination of the project.

## 0.4 OpenNebula

OpenNebula is a useful opensource that enables seamless management and control of different cloud systems. The tools can be used for a cloud implementations to virtualize data centers and also to obtain solution for cloud infrastructure. Opennebula can be adopted on top of an existing

cloud setup. OpenNebula project started on 2005 and currently the product is available as an open-source under Apache license. “The toolkit includes features for integration, management, scalability, security and accounting. It also claims standardization, interoperability and portability, providing cloud users and administrators with a choice of several cloud interfaces(Amazon EC2 Query, OGF Open Cloud Computing Interface and vCloud) and hypervisors (Xen, KVM and VMware), and can accommodate multiple hardware and software combinations in a data center” [hid-sp18-417-opennebula-wiki]. The OpenNebula deployment needs(1) A client node(2)

../hid-sp18-417/technology/images/hid-sp18-417-opennebula.png

Figure 2: OpenNebula Deployment Model [hid-sp18-417-opennebula-deployment]

A hypervisor(3) A data storage system(4) Physical network. The deployment model is depicted in ???. The OpenNebula deployment needs (1) A client node (2) A hypervisor (3) A data storage system (4) Physical network. The deployment model is depicted in Figure ??. Due to its long steady growth, the tool is being used by customers in various industries ranging from telecom to education. The wide range of customer base is helpful in providing a solid support system to the new and existing users as well as continuous feedback becomes vital in the research and growth of the project.

## 0.5 Open Refine

OpenRefine is a useful open source that is used for data visualization and analysis. Its predominantly used for cleaning messy data and transformation of data from one format to other for ease of clarity. OpenRefine was formerly known as GoogleRefine. The tool is also used for fetching data from websites and data organization. It can import data from CSV, TSV, Excel, XML etc. It is written in Java. It works with data in tabular format like in relational data. The tool has a user interface that is available to be downloaded. “Once you get used to which commands do what, this is a powerful tool for data manipulation and analysis that strikes a good balance between functionality and ease of use” [hid-sp18-417-openrefine].

## References