

the Availability Digest

www.availabilitydigest.com

OpenStack – The Open Cloud

April 2012

As expressed by Tom O'Reilly, *"If cloud computing is the future, then understanding how to make that future open is one of the great technology challenges of our day."*

Major clouds are now forming. There is Amazon Web Services, Google Apps, Microsoft Azure, Rackspace Cloud, and Force.com, among others. Many companies are looking seriously at moving some of their applications to a cloud. Clouds provide many advantages – rapid provisioning for applications without capital expenditures, elastic resourcing depending upon an application's immediate needs, and a pay-as-you-go business model.

However, moving an application to a cloud is certain lock-in. Clouds are simply incompatible with each other. Choosing a cloud is like choosing an IBM or a Burroughs mainframe decades ago. Once you are up and running, you are owned by the vendor.

There is clearly an advantage to having common cloud standards. A company could adapt its applications to the cloud and run them on the cloud of its choice. If it becomes disenchanted with its cloud provider, it can simply move to another cloud as easily as moving a Java application to another platform. It could backup its applications in another cloud in case its cloud-of-choice goes down. And yes, there have been many disastrous failures of today's clouds.¹

OpenStack is a major initiative to achieve this goal. On its web site, <http://openstack.org>, OpenStack declares that it is "a large scale open source cloud project and community established to drive industry standards, end cloud lock-in and speed the adoption of cloud technologies by service providers and enterprises." OpenStack allows service providers, enterprises, and government agencies to build massively scalable public and private clubs using freely available Apache-licensed software.

Rackspace and NASA are OpenStack's key initial founders. They contributed the initial code to OpenStack from their cloud platforms. OpenStack is currently supported by over 150 companies including HP, Citrix, Intel, AMD, SUSE Linux, Microsoft, Cisco, Dell, and Canonical.

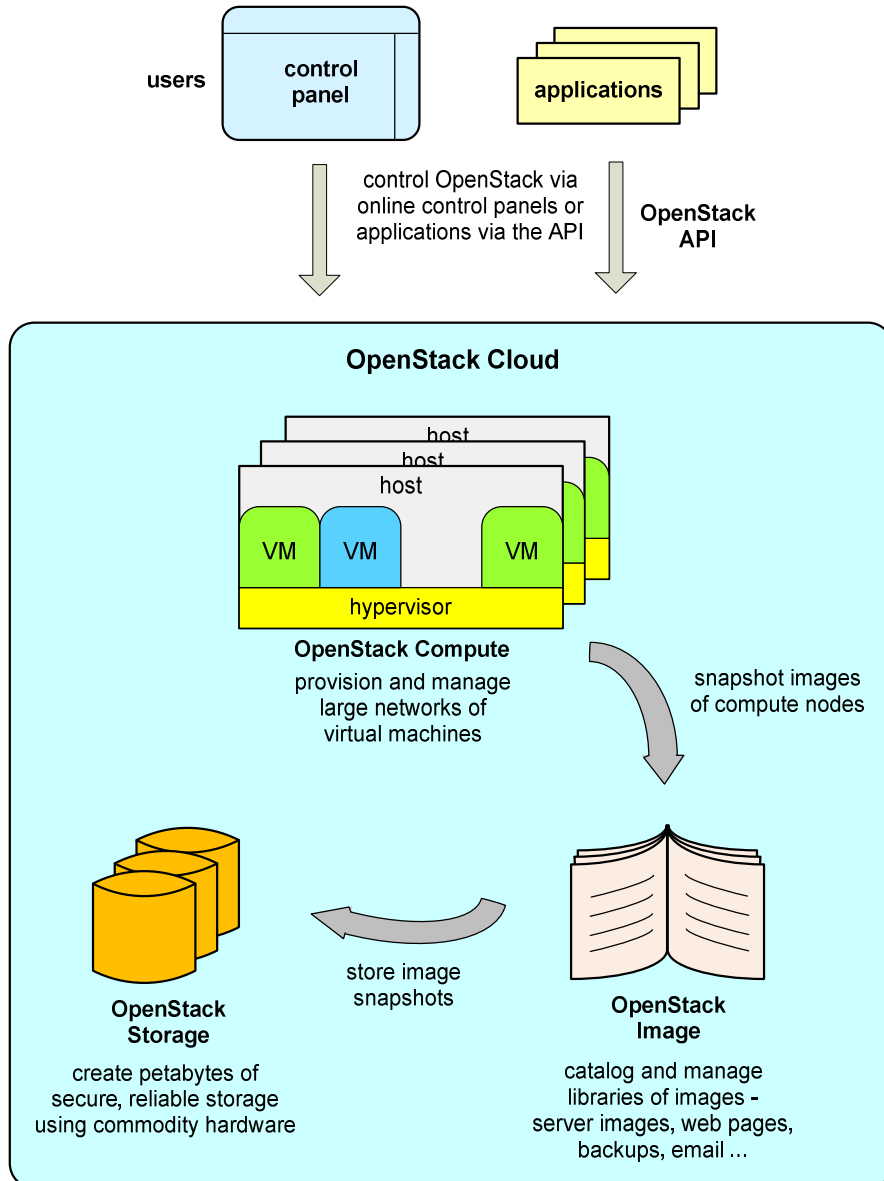
¹ [Google Troubles – A Case Study in Cloud Computing](http://www.availabilitydigest.com/public_articles/0410/google_troubles.pdf), *Availability Digest*, October 2009.
http://www.availabilitydigest.com/public_articles/0410/google_troubles.pdf
[Poor Documentation Snags Google](http://www.availabilitydigest.com/public_articles/0504/google_power_out.pdf), *Availability Digest*, April 2010.
http://www.availabilitydigest.com/public_articles/0504/google_power_out.pdf
[Amazon's Cloud Downed by Fat Finger](http://www.availabilitydigest.com/public_articles/0605/amazon_ebs.pdf), *Availability Digest*, May 2011.
http://www.availabilitydigest.com/public_articles/0605/amazon_ebs.pdf
[Windows Azure Cloud Succumbs to Leap Year](http://www.availabilitydigest.com/public_articles/0703/azure.pdf), *Availability Digest*, March 2012.
http://www.availabilitydigest.com/public_articles/0703/azure.pdf

OpenStack Architecture

OpenStack has three major components:

- OpenStack Compute provisions and manages large networks of virtual machines.
- OpenStack Object Store creates massive, secure, and reliable storage using standard hardware.
- OpenStack Image catalogs and manages libraries of server images stored on OpenStack Storage.

Furthermore, there are dozens of ancillary projects adding more value to OpenStack.



In addition to its own REST-based (Representational State Transfer) API, OpenStack supports Amazon's EC2 (Elastic Compute Cloud) and S3 (Simple Storage Services) APIs. This allows companies to continue to use their legacy Amazon Web Services APIs to manage their applications until they can take advantage of the advanced functionality of the native OpenStack API.

OpenStack Compute

OpenStack Compute (code-named Nova) provides all of the facilities necessary to support the life cycle of instances in the OpenStack cloud. It creates a redundant and scalable computing platform comprising large networks of virtual machines. It provides the software, control panels, and APIs necessary for orchestrating a cloud, including running instances, managing networks, and controlling access to the cloud.

OpenStack Compute is hypervisor agnostic. It supports many of today's leading hypervisors including VMWare ESX, Microsoft Hyper-V, and Citrix Xen, among many others.

Typical uses include implementing an IaaS platform, scaling up and down to meet demand for web resources and applications, processing massive amounts of data with tools like Hadoop, and provisioning compute resources for teams and projects.

OpenStack Compute scales to thousands of compute nodes with a shared-nothing architecture. It is inherently multi-tenant, including billing, logging, auditing, and end-user control panels.

A pluggable authentication system makes it possible to easily integrate an existing authentication system including those that use an LDAP backend.

A web control panel suitable for rebranding is available for customer-facing control. OpenStack Compute also includes reference Android and iOS clients suitable for rebranding to provide provider-specific support for its customers.

OpenStack Storage

OpenStack Storage (code-named Swift) is modeled after Amazon's EBS (Elastic Block Store) mass store.² It provides redundant, scalable data storage using clusters of inexpensive commodity servers and hard drives to store massive amounts of data. It is not a file system or a database system. Rather, it is intended for long-term storage of large amounts of data (blobs). Its use of a distributed architecture with no central point of control provides great scalability, redundancy, and permanence.

Objects are written to multiple storage devices to create a highly redundant storage cluster. Storage clusters scale horizontally by adding new nodes. OpenStack is responsible for ensuring replication and integrity across the cluster. Should a node fail, OpenStack replicates its contents from a surviving active node to another node to ensure that the level of redundancy is maintained.

Storage attached to an OpenStack virtual machine is nonpersistent. It disappears when the virtual machine is terminated. Consequently, OpenStack Storage is integrated with OpenStack Compute to provide persistent storage for OpenStack virtual-machine images and applications. Other uses for OpenStack Storage include web images, backup storage, archiving, and email storage.

OpenStack Storage may be shared publicly, or it can be kept private. It can scale to multiple petabytes and billions of objects.

OpenStack Image Service

OpenStack Image Service (code-named Glance) is a retrieval system for virtual-machine images. It provides registration, discovery, and delivery services for these images. It can use OpenStack Storage or Amazon S3 (Simple Storage System) for storage of virtual-machine images and their associated

² Amazon's Availability Zones, *Availability Digest*, November 2011.
http://www.availabilitydigest.com/public_articles/0611/amazon_availability_zones.pdf.

metadata. It provides a standard web RESTful interface for querying information about stored virtual images.

Clients can register new virtual-machine images with OpenStack Image Service, obtain information on publicly available images, and download images. The service is compatible with all common image formats.

Networking

OpenStack ensures that virtual-machine instances are configured so that they can communicate with themselves and with the outside world. Just as in Amazon Web Services, an instance can have two IP addresses associated with it. One is a private address used for communicating between instances. The other is a public IP address used for communicating with the outside world.

Security

OpenStack security is addressed at all layers throughout the stack. It provides ingress filtering for a virtual-machine instance based on security groups. A security group is a named set of rules that get applied to incoming packets for the instance. Each security group can have multiple rules associated with it such as the source IP address and network, the protocol type, the destination ports, and so forth. Any packet matching the rules of the security group is accepted; all others are rejected. A null security group blocks all traffic.

OpenStack provides no egress filtering. All outbound traffic is allowed.

The OpenStack development team recognizes that security is a complex, evolving issue that needs to be aggressively pursued and exposures eliminated. It encourages the reporting of security vulnerabilities to its OpenStack Vulnerability Management Team.

OpenStack Commands

OpenStack provides commands for administrative tasks such as user management, instance management, and network management. The command classes include the following:³

Euca2ools Commands

Eucalyptus (an acronym for “Elastic Utility Computing Architecture for Linking Your Programs To Useful Systems”) is a software platform for the implementation of cloud computing clusters. It implements the Amazon Web Services API. OpenStack provides the Eucalyptus command line tools called Euca2ools, which can be used internally to interact with an OpenStack cloud or externally to interact with Amazon’s EC2 cloud.

User/Role Management

With the commands to manage users and their roles, the system administrator can add, list and delete users, associate users with existing projects, view and modify access keys of users, assign user roles, and check to see if a user has a specific role.

³ OpenStack Compute Starter Guide, <http://docs.openstack.org>.

Instance Management

OpenStack defines instance types by specifying the amount of memory, the number of virtual CPUs, and the disk storage that can be allocated to an instance type. With the instance management commands, an administrator can define and remove instance types and can list the current instance types.

Service Management

The administrator can check the state of all available services, enable or disable a service, and view the resource utilization (number of CPUs, memory, and disk) of various OpenStack components.

Project Management

The administrator can create, manage, and delete projects. He can list the registered projects, download the files for a specific project, and check the resources allocated to a project such as storage and virtual-machine instances.

Database Management

OpenStack stores the data related to a project, its users, and its resources in a relational database. The administrator can obtain the current database version and can sync the database schema to the current version.

Announced Implementations

Trial implementations of OpenStack clouds are now becoming available.

Ubuntu Cloud Live

Ubuntu is an open-source operating system sponsored by Canonical Ltd. It is based on the Debian Linux distribution. Canonical is one of the sponsors of OpenStack, and its operating system, Ubuntu, has been ported to OpenStack.

Ubuntu is the reference operating system for the OpenStack cloud. Ubuntu Cloud Live delivers a fully functional OpenStack cloud that can be run on a PC for demonstration and development purposes.⁴

HP Cloud Services

HP has announced the availability of HP Cloud Services, which provides an Infrastructure as a Service (IaaS) public cloud offering.⁵ HP Cloud Services uses the open-source OpenStack cloud and is now available for private Beta use. On May 10th, 2012, HP Cloud Services will become available in public Beta form.

The HP Cloud Services components, HP Cloud Compute and HP Cloud Storage, have a web-based interface that is mobile-accessible. HP Cloud Services is replicated across multiple availability zones and includes proactive HP 24/7 monitoring.

HP has chosen Linux-based Ubuntu as one of its host and guest operating systems. Several languages, including Java, Python, Ruby, and PHP are supported.

Companies can apply for participation in HP's private Beta at www.hpcloud.com.

⁴ OpenStack Projects, <http://openstack.org/projects/>.

⁵ HP Announces Support for OpenStack, *HP Blogs*; July 27, 2011.

[Ubuntu Powers HP Public Cloud](#), *Canonical Announcement*; October 6, 2011.

[HP Cloud Services Now in Private Beta](#), www.hpcloud.com; March 6, 2012.

Summary

In the early days of Edison, companies depended upon their own electrical generators to power their factories. These generators were the equivalent of today's data centers. Then local standardization allowed communities to share a common power system for lighting, manufacturing, and other uses. This is where we are presently in our cloud technology. It wasn't until national and international standards were developed that nationwide power grids could serve all communities. Thus was created the true electric utility.

Similarly, a true compute utility is awaiting the development of accepted standards. Only then can a company plug into the public cloud of its choice for its computing needs. OpenStack is a heavily supported industry effort to create such standards via a free and open mechanism. Thousands of developers are now contributing their time and talent to this important project; and as time goes on, it is likely that thousands more will join the effort. Hopefully, OpenStack will be a major step forward in achieving a true compute utility.