

CloudStackProposal

Abstract

CloudStack is an IaaS ("Infrastructure as a Service") cloud orchestration platform.

Proposal

CloudStack provides control plane software that can be used to create an IaaS cloud. It includes an HTTP-based API for user and administrator functions and a web UI for user and administrator access. Administrators can provision physical infrastructure (e.g., servers, network elements, storage) into an instance of CloudStack, while end users can use the CloudStack self-service API and UI for the provisioning and management of virtual machines, virtual disks, and virtual networks.

Citrix Systems, Inc. submits this proposal to donate the CloudStack source code, documentation, websites, and trademarks to the Apache Software Foundation ("ASF").

Background

Amazon and other cloud pioneers invented IaaS clouds. Typically these clouds provide virtual machines to end users. CloudStack additionally provides baremetal OS installation to end users via a self-service interface. The management of physical resources to provide the larger goal of cloud service delivery is known as "orchestration". IaaS clouds are usually described as "elastic" – an elastic service is one that allows its user to rapidly scale up or down their need for resources.

A number of open source projects and companies have been created to implement IaaS clouds. Cloud.com started CloudStack in 2008 and released the source under GNU General Public License version 3 ("GPL v3") in 2010. Citrix acquired Cloud.com, including CloudStack, in 2011. Citrix re-licensed the CloudStack source under Apache License v2 in April, 2012.

Rationale

IaaS clouds provide the ability to implement datacenter operations in a programmable fashion. This functionality is tremendously powerful and benefits the community by providing:

- More efficient use of datacenter personnel
- More efficient use of datacenter hardware
- Better responsiveness to user requests
- Better uptime/availability through automation

While there are several open source IaaS efforts today, none are governed by an independent foundation such as ASF. Vendor influence and/or proprietary implementations may limit the community's ability to choose the hardware and software for use in the datacenter. The community at large will benefit from the ability to enhance the orchestration layer as needed for particular hardware or software support, and to implement algorithms and features that may reduce cost or increase user satisfaction for specific use cases. In this respect the independent nature of the ASF is key to the long term health and success of the project.

Initial Goals

The CloudStack project has two initial goals after the proposal is accepted and the incubation has begun.

The Cloudstack Project's first goal is to ensure that the CloudStack source includes only third party code that is licensed under the Apache License or open source licenses that are approved by the ASF for use in ASF projects. The CloudStack Project has begun the process of removing third party code that is not licensed under an ASF approved license. This is an ongoing process that will continue into the incubation period. Third party code contributed to CloudStack under the CloudStack contribution agreement was assigned to Cloud.com in exchange for distributing CloudStack under GPLv3. The CloudStack project has begun the process of amending the previous CloudStack contribution agreements to obtain consent from existing contributors to change the CloudStack project's license. In the event that an existing contributor does not consent to this change, the project is prepared to remove that contributor's code. Additionally, there are binary dependencies on redistributed libraries that are not provided with an ASF-approved license. Finally, the CloudStack has source files incorporated from third parties that were not provided with an ASF-approved license. We have begun the process of re-writing this software. This is an ongoing process that will extend into the incubation period. These issues are discussed in more detail later in the proposal.

Although CloudStack is open source, many design documents and discussions that should have been publicly available and accessible were not publicized. The Project's second goal will be to fix this lack of transparency by encouraging the initial committers to publicize technical documentation and discuss technical issues in a public forum.

Current Status

Meritocracy

[CloudStack](#) was originally developed by Sheng Liang, Alex Huang, Chiradeep Vittal, and Will Chan. Since the initial [CloudStack](#) version, approximately 30 others have made contributions to the project. Today, Sheng and Will are less involved in code development, but others have stepped in to continue the development of their seminal contributions.

Most of the current code contributors are paid contributors, employed by Citrix. Over the past six months [CloudStack](#) has received several contributions from non-Citrix employees for features and bug fixes that are important to the contributors. We have developed a process for accepting these contributions that includes validating the execution of a CLA and incorporating the contribution in the [CloudStack](#) in a manner that reflects the contributor's identity. This process has not followed the Apache model.

The [CloudStack](#) Project has had an open bug database for two years. While this database includes ideas for enhancements to [CloudStack](#), the committers have historically not asked the greater community for pointed assistance. Going forward the Project will encourage all community members to become committers and will make clear suggestions for features and bug fixes that would most benefit the community and Project.

Community

[CloudStack](#) has an existing community comprising approximately 8,000 forum members on cloudstack.org and 28,000 registrations for e-mail lists and newsletters relating to [CloudStack](#). All forums, developer and administrator mailing lists, and IRC channels are active. A number of commercial entities (e.g., [RightScale](#), [AppFog](#), [EnStratus](#)) and open source projects (e.g., [jClouds](#), [Chef](#)) have integrated with [CloudStack](#).

To date, the community comprises users – people that download a [CloudStack](#) binary and install it to implement an IaaS cloud. The project expects that with independent governance and the openness of the Apache development model we will significantly increase the amount of developer participation within the community.

Core Developers

[CloudStack](#) spans a wide array of technologies: user interface, virtualization, storage, networking, fault tolerance, database access and data modeling, and Java, Python, and bash programming. There is significant diversity of knowledge and experience in this regard.

Several of the initial committers have experience with other open source projects. Alex Huang contributed to SCM-bug. Anthony Xu, Edison Su, Frank Zhang, and Sheng Yang have prior experience with a combination of Xen and KVM. Chiradeep Vittal has contributed to [OpenStack](#). David Nalley has been contributing to Fedora for several years. David has also contributed to Zenoss, Cobbler, GLPI, OCS-NG, [OpenGroupware](#), Ceph, and Sheepdog.

[CloudStack](#) development to date has largely been done in the U.S. and India.

[CloudStack](#) has largely been developed by paid contributors.

Alignment

[CloudStack](#) has significant integration with existing Apache projects, and there are several exciting opportunities for future cross-project collaboration.

The [CloudStack](#) Management Server (i.e., the control plane) is deployed as a web application inside one or more Tomcat instances.

The Management Server uses Apache Web Services, Apache Commons, Apache XML RPC, Apache log4j, and Apache [HttpComponents](#) httpcore. It is built with Apache Ant.

There are strong opportunities for collaboration with other Apache Projects. Collaboration with Hadoop has at least two exciting aspects:

- [CloudStack](#) could provide an object store technology (similar to Amazon's S3 service) in conjunction with the compute service (similar to Amazon's EC2 service) that it already offers. HDFS from the Hadoop project is a promising technology for the implementation of the object store.
- It would also be possible to have [CloudStack](#) provision Hadoop compute nodes, either through virtualization or directly to baremetal. With this [CloudStack](#) could become an optional or required part of the infrastructure control plane for Hadoop.

[ZooKeeper](#) might be helpful to implement a distributed cloud control plane in the future.

Derby could be used as alternative database; [CloudStack](#) currently uses MySQL.

ActiveMQ is a good option for some of the communication that occurs in the orchestration of the cloud.

It would be natural for Apache libcloud and Apache [DeltaCloud](#) to support the [CloudStack](#) API and public clouds that expose it.

As mentioned earlier the proposers are seeking an independent foundation to provide governance for the project. ASF has clearly been successful in providing this, and we believe ASF is the best match for the future goals of the project.

Known Risks

Orphaned products

Citrix will work with the community to create the most widely deployed cloud orchestration software. Citrix's internal "plan of record" commits significant budget to developing the Project through 2014. Investment past 2014 is unspecified, but likely to continue given known and predicted revenues from derivative commercial products.

Citrix is developing a thriving business in conjunction with the prior and continued success of the community and use of [CloudStack](#). The project may be orphaned in the condition where the Project has failed to obtain either non-paid committers or paid committers from other vendors, and the committers paid by Citrix are re-assigned to another project.

Inexperience with Open Source

[CloudStack](#) has been open source since May, 2010, with the [CloudStack](#) 2.0 release by Cloud.com.

From May, 2010 to August, 2011 [CloudStack](#) was "open core", wherein approximately 95% of the code was available with a GPLv3 license and 5% of the code was proprietary. During this time the bug database was open and the source code was available. Project direction and technical discussions occurred in a closed fashion. Few technical documents were publicly available.

In August, 2011 [CloudStack](#) transitioned to 100% open source. The 5% proprietary code was released publicly with a GPLv3 license. The bug database remained open. Project direction and technical discussions occurred in a closed fashion. Some technical documents were shared publicly.

During 2012 the proposers have posted a significant fraction of technical documents pertaining to the recent [CloudStack](#) 3.0 release publicly. Some technical discussion has occurred in the open.

In April, 2012 [CloudStack](#) was re-licensed under the Apache License v2.

Several contributors have prior open source experience. This is discussed in the "Core Developers" section.

The [CloudStack](#) development process must change significantly to conform to the Apache model. These changes include: carry on all technical conversations in a public forum, develop all technical documentation publicly, follow the vote process on contribution approvals, and promote individuals beyond the initial committers to committer status, based on merit.

Homogenous Developers

The Project has committers in two locations in India, one location in the UK, and one location in the U.S. The technical knowledge of the committers is diverse, as evidenced by the wide range of technologies that converge in [CloudStack](#). The range of professional experience of the committers is diverse as well, from a few months to 20+ years.

The initial committers are all associated with the sponsoring entity. The Project will have to work with the community to diversify in this area.

Reliance on Salaried Developers

The initial committers are all salaried committers.

The initial committers have worked with great devotion to the project and have enjoyed its success. We hope this will create an emotional bond to the project that will last beyond their employment with Citrix Systems.

We expect salaried committers from a variety of companies. [CloudStack](#) is an opportunity for many vendors to enable their software and hardware to participate in the changes brought by the development of an API that can manage datacenter infrastructure. It is also an opportunity for datacenter operators to implement features they find helpful and share them with the community.

We hope to attract unpaid committers. [CloudStack](#) is interesting technology that solves many challenging problems, and cloud computing is popular in the industry media now. But, few people will run a [CloudStack](#) deployment for personal use, and this may limit our ability to attract unpaid committers. We hope that the technical domain is interesting to new committers that will join us in improving [CloudStack](#).

Relationships with Other Apache Products

Please see the Alignment section above.

Apache Brand Awareness

We expect that licensing [CloudStack](#) under the AL and associating it with the Apache brand will attract additional contributors and [CloudStack](#) users. However, we have selected the ASF as the best governance option for the project for the reasons discussed in the Rationale. Further, we expect to continue development of the [CloudStack](#) under the AL with or without the support of ASF.

Citrix currently sells a proprietary version of [CloudStack](#) released as "Citrix [CloudStack](#)". For the foreseeable future, Citrix expects to continue to sell orchestration software based on [CloudStack](#). Citrix will work with the ASF Incubator PMC and within the Podling Branding guidelines to ensure that a new branding scheme is selected for Citrix's proprietary version of [CloudStack](#) that is consistent with ASF's branding policies.

Documentation

The [CloudStack](#) project has publicly available administrator documentation, source code, forums, and technical specifications. This documentation is available at the following sites:

- <http://cloudstack.org>: forums, latest news, downloads, blogs; a good starting point.
- <http://docs.cloudstack.org>: installation guide, administration guide, API documentation, technical specifications
- <http://confluence.cloudstack.org>: past and future release plans, additional technical documentation
- <http://git.cloud.com>: current source. See the 3.0.x and master branches.

Initial Source

The genesis of [CloudStack](#)'s source is discussed in the "Inexperience with Open Source" section.

Citrix Systems currently owns the [CloudStack](#) code base. Committers use the repository at git.cloud.com to access and submit code. This repository is located in the U.S.

We propose to donate the basis for the 3.0.x series of [CloudStack](#) releases. This is the current release stream. Prior [CloudStack](#) versions have been kept as GPLv3 and currently receive limited maintenance and no feature development. The software associated with these prior versions will not be donated to ASF. Further, many branches exist and we see no benefit in recreating this historical complexity within ASF infrastructure.

Source and Intellectual Property Submission Plan

Multiple intellectual property assets are associated with the [CloudStack](#) project. First and foremost, the [CloudStack](#) source is protected by copyright. Upon acceptance into the ASF incubation program, Citrix Systems anticipates licensing the [CloudStack](#) source to the ASF. The licensed code will include all source code from the "master" branch at git.cloud.com.

In addition to the source code, Citrix systems owns a number of trademark and domain name assets that are used by the [CloudStack](#) project. Citrix anticipates donating substantially all of these trademark and domain name assets upon acceptance into the ASF incubation program. In particular, Citrix anticipates donating at least the [CloudStack](#) trademark and related domain names.

[CloudStack](#) is protected by a number of pending patent applications owned by Citrix Systems. Citrix Systems anticipates continuing to prosecute and maintain these patent applications upon entry into the ASF incubation program. Citrix Systems is dedicated to protecting the larger [CloudStack](#) community and will continue to obtain patents on [CloudStack](#) technology as a way to protect contributors and members of the [CloudStack](#) community from outside threats.

Internal Dependencies

The [CloudStack](#) Management Server has some externally developed code embedded in it. This code has come from a variety of sources and has a variety of licenses, some of which are not approved by ASF for use in Apache projects. We have already begun the process of removing and/or re-implementing code that does not have an approved license.

Component	License
GlassFish	COMMON DEVELOPMENT AND DISTRIBUTION LICENSE (CDDL) Version 1.0
Copyright (c) 2000-2005 INRIA, France Telecom	Unknown License

Contributions made to the [CloudStack](#) prior to the switch to AL were done based on a CLA that did not authorize re-licensing the contribution to AL. Citrix legal has prepared a new document that requests contributors to authorize the re-license to AL. We are asking each such contributor to sign this agreement. We will remove and/or re-implement the contributions of prior committers that do not sign this agreement. We do not expect this issue to materially impact the project.

Citrix legal has also prepared a new CLA for the project that authorizes AL licensing of contributions. This CLA will be used for contributions between the switch to AL and an eventual donation of the source to ASF.

External Dependencies

The [CloudStack](#) Management Server uses a significant number of libraries. These libraries are redistributed with [CloudStack](#) in binary form. Some of them have licenses that are not approved by ASF for use in Apache projects. We will replace them with other libraries with approved licenses or re-write the functions provided by the libraries.

We expect that it will take 3 months to remove and/or re-implement the problematic embedded source and problematic redistributed libraries.

Binary Dependencies

Component	License
GSON Closure Compiler	Apache License Version 2.0
MySQL Connector/J	GPL 2.0
paramiko	LGPL 2.1
JavaMail	Sun JavaMail 1.4 License
Java Servlet Technology	Sun Microsystems Binary Code License
JavaServer Pages Standard Tag Library	COMMON DEVELOPMENT AND DISTRIBUTION LICENSE (CDDL) Version 1.0
Apache Jakarta Commons Discovery	Apache 1.1
Apache log4j	Apache 1.1
Apache Web Services Axis	Apache 1.1
Apache Ant	Apache License Version 2.0
Apache Jakarta Commons Codec	Apache License Version 2.0
Apache Jakarta HTTP Client	Apache License Version 2.0
Apache Jakarta HttpComponents	Apache License Version 2.0
Apache Jakarta HttpComponents	Apache License Version 2.0
Apache Tomcat	Apache License Version 2.0
Apache Web Services Axis	Apache License Version 2.0
Apache Xerces Java XML Parser	Apache License Version 2.0
Apache-Jakarta Collections	Apache License Version 2.0
Apache-Jakarta DBCP	Apache License Version 2.0
Apache-Jakarta Lang	Apache License Version 2.0
Apache-Web Services Commons Util	Apache License Version 2.0
Code Generation Library	Apache License Version 2.0
ehcache	Apache License Version 2.0
excanvas	Apache License Version 2.0
google-gson	Apache License Version 2.0
iBATIS for Java	Apache License Version 2.0
Jakarta Commons-Logging	Apache License Version 2.0
Jetty - Java HTTP Servlet Server	Apache License Version 2.0
selenium	Apache License Version 2.0
jquery-easing	BSD 2.0
Orion SSH2	BSD 2.0
VMware Infrastructure Java API	BSD 2.0
XStream Library	BSD 2.0
JUnit	Common Public License
backport-util-concurrent	Creative Commons Public Domain Dedication
JSch	JSch License
Bouncy Castle Crypto APIs	MIT License V2
flot	MIT License V2
jquery-ui	MIT License V2
mockito	MIT License V2
pymysql	MIT License V2
UUID - generate UUIDs in Java	MIT License V2
iHarder.net - base64	Unspecified
iControl.jar	GPL
JnetPcap	LGPLv3

libvirt 0.4.5	LGPLv3
manageontap	NetApp EULA
NetScaler SDK	Unknown
Trilead ssh	Trilead EULA
XAPI API	GPLv2
jquery.timers	Unknown License
jquery-yafu	GPL 2.0
jquery-validate	MIT License V2
jqueryjs	MIT License V2

System Virtual Machines

The [CloudStack](#) uses multiple Debian-based virtual machines to implement features of the software. The source code that comprises the Debian-based virtual machines is GPL licensed.

The [CloudStack](#) source code includes (AL) scripts that will download and build this software. This software is downloaded from repositories external to git.cloud.com, and will presumably also be external to any Apache-owned infrastructure.

The [CloudStack](#) will download and deploy virtual machines that are built with this GPL software. Once deployed, the [CloudStack](#) will install AL-licensed software on to these virtual machines.

Since this GPL software is not present in the [CloudStack](#) repository we believe these mechanisms will be approved by ASF for use in the Project, but we have included this explanation for completeness.

Cryptography

The [CloudStack](#) makes use of encryption functions available via Java and the underlying OS. We expect that the [CloudStack](#) will have to follow the export control procedures described at <http://www.apache.org/dev/crypto.html>. When the [CloudStack](#) was previously registered with BIS the open source version qualified for the TSU exception.

The [CloudStack](#) uses https to communicate to [XenServer](#) and vCenter. ssh and scp are used between the Management Server and hypervisor hosts as well.

The [CloudStack](#) stores an MD5 hash of user password data. The [CloudStack](#) uses MySQL encryption to store some data in an encrypted fashion.

The [CloudStack](#) stores a pair of API public/secret keypairs for users. This is done using javax.crypto.KeyGenerator with HMAC-SHA-1.

The [CloudStack](#) does not specify key lengths explicitly. It uses SSH, SCP and lets them negotiate encryption.

The [CloudStack](#) provides a public HTTP-based API to provision and deprovision VPN users. The [CloudStack](#) has internal Java-based abstractions for managing VPN users. This Java software makes private API calls to another system, which will then provision the VPN user in the VPN software on that other system. The actual set up of the VPN session is done using L2TP/IPSec.

As mentioned earlier the [CloudStack](#) includes software to build and later deploy Debian-based virtual machines. These VMs are stripped down versions of Debian that include encryption sufficient for ssh/scp, https, and IPSec VPN to work. The [CloudStack](#) does not include the source for these VMs. The maximum encrypted throughput of the VPN has not been determined.

Required Resources

Mailing Lists

We request mailing lists to match the mailing lists currently in use, plus the recommended private list. These are:

- cloudstack-private: for confidential PPMC discussion
- cloudstack-dev: for development discussions
- cloudstack-commits: for source code changes
- cloudstack-users: for administrator and discussions

Subversion Directory

The [CloudStack](#) has used git for approximately two years. We understand that there is a "prototype" git server available. We request an allocation on this git server. We believe this will be less disruptive to the committers than a change to SVN.

We request `"/repos/asf/incubator/cloudstack"`.

Issue Tracking

We would like an allocation for Jira. [CloudStack](#) uses bugzilla today, but we have been planning a move to Jira for some time. We request that the project name be “CloudStack”.

Other Resources

The [CloudStack](#) Project includes several websites. Donation of these websites was discussed in the IP submission plan. We would like to engage in discussion on the logistics of this.

Initial Committers

In the past few months several new developers have joined the Citrix [CloudStack](#) team. We are recommending that only the developers with several months of experience with [CloudStack](#) join as initial committers. The Project will then follow the meritocratic process to enable the newer team members to become committers. We believe this will be a good exercise for us as we transition to an Apache development model in the Project.

The list of initial committers follows. At this time none of the initial committers has a CLA on file with ASF.

- Abhinandan Prateek, abhinandan.prateek@citrix.com
- Alena Prokharchyk, alena.prokharchyk@citrix.com
- Alex Huang, alex.huang@citrix.com
- Anthony Xu, xuefei.xu@citrix.com
- Brian Federle, brian.federle@citrix.com
- Chiradeep Vittal, chiradeep.vittal@citrix.com
- David Nalley, david.nalley@citrix.com
- Edison Su, edison.su@citrix.com
- Frank Zhang, frank.zhang@citrix.com
- Janardhana Reddy, janardhana.reddy@citrix.com
- Jessica Tomechak, jessica.tomechak@citrix.com
- Jessica Wang, jessica.wang@citrix.com
- Kelven Yang, kelven.yang@citrix.com
- Kevin Kluge, kevin.kluge@citrix.com
- Kishan Kavala, kisha.kavala@citrix.com
- Murali Reddy, murali.reddy@citrix.com
- Nitin Mehta, nitin.mehta@citrix.com
- Prachi Damle, prachi.damle@citrix.com
- Sam Robertson, sam.robertson@citrix.com
- Sheng Yang, sheng.yang@citrix.com
- Sonny Chhen, sonny.chhen@citrix.com
- Will Chan, will.chan@citrix.com

Affiliations

The initial committers are all affiliated with Citrix Systems.

Sponsors

Champion

Jim Jagielski

Nominated Mentors

Jim Jagielski, Daniel Kulp, Alex Karasulu, Olivier Lamy, Brett Porter, Mohammad Nour, Matt Hogstrom

Sponsoring Entity

We request that the Incubator sponsor this effort.