

CernVM Online and Cloud Gateway: a uniform interface for CernVM contextualization and deployment

G Lestaris¹, I Charalampidis², D Berzano, J Blomer, P Buncic, G Ganis and R Meusel

CERN, CH-1211 Genve 23, Switzerland

E-mail: ¹george.lestaris@cern.ch

E-mail: ²ioannis.charalampidis@cern.ch

Abstract. In a virtualized environment, contextualization is the process of configuring a VM instance for the needs of various deployment use cases. Contextualization in CernVM can be done by passing a handwritten context to the user data field of cloud APIs, when running CernVM on the cloud, or by using CernVM web interface when running the VM locally. CernVM Online is a publicly accessible web interface that unifies these two procedures. A user is able to define, store and share CernVM contexts using CernVM Online and then apply them either in a cloud by using CernVM Cloud Gateway or on a local VM with the single-step pairing mechanism. CernVM Cloud Gateway is a distributed system that provides a single interface to use multiple and different clouds (by location or type, private or public). Cloud gateway has been so far integrated with OpenNebula, CloudStack and EC2 tools interfaces. A user, with access to a number of clouds, can run CernVM cloud agents that will communicate with these clouds using their interfaces, and then use one single interface to deploy and scale CernVM clusters. CernVM clusters are defined in CernVM Online and consist of a set of CernVM instances that are contextualized and can communicate with each other.

1. Introduction

At CernVM [1], *contextualization* is the process of applying a predefined set of parameters to a *polymorphic* virtual machine instance, converting it to a specialized processing unit. This process removes the need of distributing large OS templates for every different processing unit configuration.

By the term *polymorphic* we mean a virtual machine that provides all basic components required, but they are not configured nor specialized. This means that an instance of a *polymorphic* VM can be configured to become any kind of processing unit.

This concept is implemented in CernVM 2.x series with the image flavors. There are four different kind of CernVM images that target four different kinds of flavours: CernVM *Head Node*, CernVM *Batch Node*, CernVM *Basic* and CernVM *Desktop*. The first two are designed to be used in cluster environments. The *Head Node* contains an HTCondor¹ master and monitoring utilities while *Batch Node* contains an HTCondor client. CernVM *Basic* contains most of the

¹ <http://research.cs.wisc.edu/htcondor/>

widely-used components; CernVM *Desktop* is the same as *Basic*, with the addition of a graphical interface.

However with the introduction of CernVM 3.x series [3] there will be a single image covering all the use cases. This increases the flexibility inherited from CernVM 2.x since, with CernVM 3.x, contextualization can transform a newly booted virtual machine to either *batch head/node* or *desktop*.

There are several ways a CernVM image can be contextualized.

1.1. Automatic contextualization

1.1.1. Amiconfig The contextualization process is applied using a patched version of *amiconfig*² from *rPath*. This tool was designed to download user-data provided via an EC2 infrastructure and apply them to the system. In CernVM we have modified this tool in order to accept configuration from other sources and in order to handle more configuration cases.

In detail, *amiconfig* is a tool written in Python and has a pluggable architecture. It is composed of a downloading class and a set of plugins that apply the configuration fetched.

1.1.2. Passing user-data User-data can be provided to a CernVM instance using various techniques. Currently CernVM supports EC2 compatible method, CloudStack and HEPIX standards. These methods are used by the vast majority of cloud providers. Also with CernVM 3.x, *cloud-init*³ package is included. When AMIconfig fails to find the context, with one of the previously mentioned methods, *cloud-init* will be used and its various data sources⁴ will be applied.

1.2. Web appliance: manual contextualization

For CernVM instances that are deployed locally, earlier described contextualization is still possible with the *HEPIX compliant CD-ROM*. However, CernVM provides also a user friendlier approach aiming to help inexperienced users. This is the *web appliance* that comes with CernVM image and starts once the machine is booted. User can access the web appliance by using a URL that is printed in VM's console.

In the remaining of this paper we describe CernVM Online in section 2, a web portal to define, store and share contexts, as well as easier ways to easily apply them in local virtual machines and virtual machines with console access. Moreover we introduce a model for formulating virtual clusters in section 3, and CernVM Cloud Gateway in section 4, a distributed system for virtual cluster deployment.

2. CernVM Online

CernVM Online is a web portal that can be considered as an *online database* of contextualization information. It provides a simplified interface that allows end-users to create, archive and share contextualization information or – called for short – *contexts*. CernVM Online is also the endpoint that every CernVM instance will contact in order to fetch the appropriate contextualization information. It provides two methods to initiate this procedure: pairing and integration with WebAPI which can be seen in figure 1 and are described in the following sections.

In addition to per-machine contextualization information, CernVM Online provides management capabilities for *cluster definitions*. A cluster definition is a virtual representation of a cluster. Virtual cluster definition is described in section 3.

² <https://bitbucket.org/rpathsync/amiconfig>

³ <http://cloudinit.readthedocs.org/>

⁴ <http://cloudinit.readthedocs.org/en/latest/topics/datasources.html>

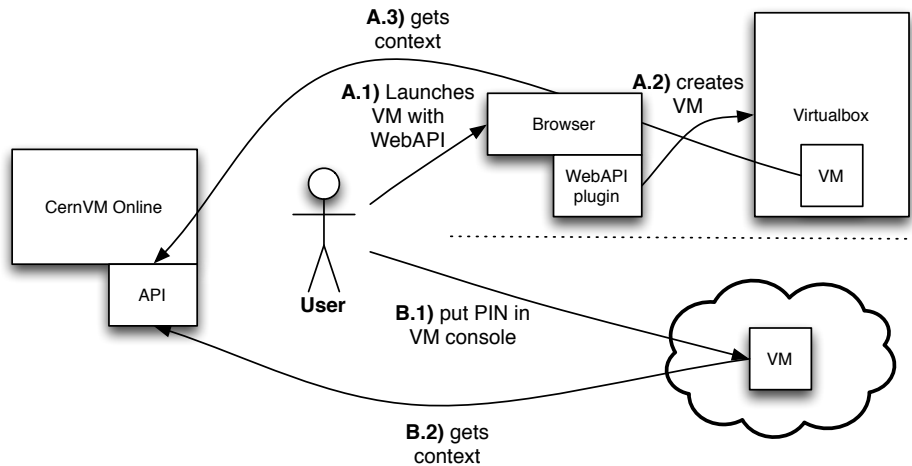


Figure 1. CernVM Online use cases

This service is currently hosted at <http://cernvm-online.cern.ch>. You can log-in with your CERN credentials, or you can create an external account. Note however that this is a public service and does not target only CERN users.

2.1. Context definition

A context is the file that contains all the information required to configure a CernVM instance. Contexts in CernVM Online are created using an interface that follows the same principle of dividing the context into sections that provide configuration parameters for different contextualization plugins. Hence the interface itself allows for selecting which plugins will be enabled and defining their settings.

2.1.1. Immutable contexts Contexts stored in CernVM Online are immutable. User is not allowed to modify them. This is done to keep consistency of stored context and context already applied on a VM instance. Nevertheless, if the user needs to modify a context, CernVM Online generates a new context (with a new id) that inherits all the settings from the “parent” context and provides user with an interface to make required modifications. This procedure is called *cloning*.

2.1.2. Security Contexts sometimes contain sensitive information such as passwords and keys. CernVM Online provides the *encrypted context* feature that enables encryption of the context contents with a user defined pass phrase. The contents are stored encrypted in the database and any access to the context (for cloning and pairing) requires the context password.

2.1.3. Marketplace CernVM Online is also a place for sharing contexts for common use cases within the user community. A user that defines a context for a common purpose can *publish* it to the CernVM Online *marketplace*. The contexts are then organised into categories, indexed by tags and users can clone and pair them.

2.2. Pairing

CernVM Online contexts can be extracted in AMI config format using the web interface or the API. This solves the problem of manually writing the contexts. Additionally the *pairing*

mechanism enables contextualization of running virtual machines by just typing a PIN number.

Pairing requires access to VM's console. There the user can specify the pairing PIN obtained from CernVM Online. The VM will contact CernVM Online to get the context and it will start the contextualization process to apply it. The VM information such as name, CernVM version and IP address will be stored in CernVM Online and the user is able to get a list of paired virtual machines.

2.3. Integration with WebAPI

CernVM WebAPI browser plugin has been developed to deploy local CernVM instances from the browser window using the VirtualBox hypervisor. It supports Internet explorer, Firefox and Chrome and it has been integrated experimentally with CernVM Online. This integration allows user to deploy a CernVM 3 instance by selecting the context and the instance hardware configuration (CPU, memory, disk). The WebAPI plugin will install VirtualBox in the system (if it is not already there), download the tiny μ CernVM image and start the machine. Once the machine boots it will be contextualized with the selected context.

3. Modeling virtual clusters

A virtual cluster is a computing cluster that consists solely by *virtual machines*. Clusters contain machines with *different roles*. Each role is defined by the services a machine runs. However machine's role also defines the configuration of the machine itself. There are machines in a cluster which are more important and that all the other machines have to connect to. These machines require special contextualization, different hardware configuration, and in strict networks different firewall settings.

The model developed to define virtual clusters separates these different machine types based on their role. This distinction is known by the name *services*. Each service in our virtual cluster can be implemented by a single VM or by many. There are two types of services based on their scalability behaviour. The *fixed services* and the *scalable services*.

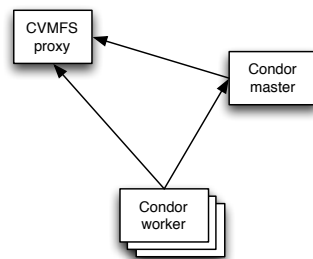


Figure 2. A typical virtual cluster

Fixed services They are services that are deployed once for each virtual cluster instance. A fixed service has a fixed number of instances implementing it. In batch systems the fixed service would be the head node. Other examples of fixed services are proxies and VO boxes.

Scalable service These are services that are been implemented by a varying number of VMs. As the name suggests, these services can be scaled up or down. Typical example is the worker nodes of a batch system.

3.1. Service dependancies

Usually the scalable service instances communicate with the fixed service instances. This implies that the fixed service instances must be available when a scalable service instance is created. To formulate the order of VM deployment we use service dependancies. Each service can depend to a number of services that must be deployed for it to start. The dependancies also used to destroy VMs. Once a fixed service instance is not used (required) by any other VM, it can be destroyed.

In the current implementation of CernVM Online interface for defining virtual clusters, dependancies are set automatically by the order of services. Also all the scalable services depend on all the fixed services. This interface will be extended in the future to allow to define more flexible dependancies between services.

3.2. Service definition

A service requires, and is identified by, a *name*. Also it describes the software and hardware configuration of its instances. *Software configuration* is selected by defining the *CernVM Online context* and *CernVM image/version* to use for a service. *Hardware configuration* is defined by the concept of offerings. There are three types of offerings: *compute*, *disk* and *network* offerings.

4. CernVM Cloud Gateway

CernVM Cloud Gateway is a distributed system used to spawn virtual clusters in multiple and different clouds. Its front end is a web server, the *cloud gateway server*, that provides the interface to the users as well as a REST API. The cloud gateway server generates requests that are put into the shared database that *gateway agents* poll. The the gateway agents communicate with the remote *cloud agents* (one for each connected cloud) through XMPP.

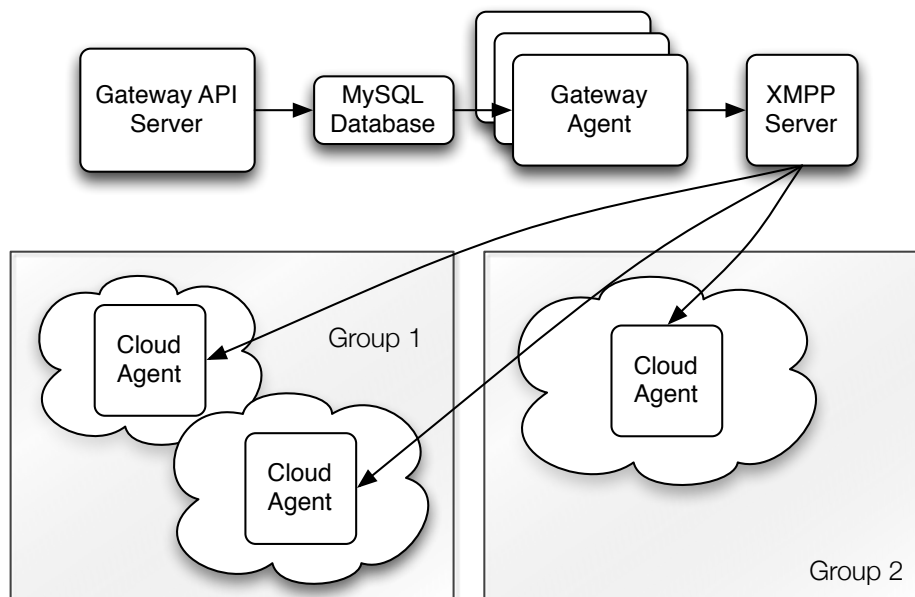


Figure 3. The architecture of the CernVM Cloud Gateway

More information on the CernVM Cloud Gateway API documentation, implementation details and administrator reference can be found in the technical report [4] of the project.

4.1. Cloud gateway server

The cloud gateway server provides the user interface and the API of the system. Access requires authentication which is based on CernVM Cloud Gateway accounts for the user interface and API credentials for the API. Note that a user can create and use more than one set of API credentials and revoke them independently. The provided interface allows for:

- Creation of new clusters, according to a cluster definition stored in CernVM Online
- Scale up cluster services
- Pause, resume or destroy individual instances and
- Pause, resume or destroy the entire cluster

The gateway server is using a MySQL database to keep the state of the system, the deployed clusters, their virtual machines, the connected clouds, the users and groups, etc. Though all the above mentioned actions required communication with the actual cloud APIs. This communication is asynchronous and it is handled by the gateway agent.

4.2. Gateway agent

All the asynchronous requests are stored in the database. One or more (for scaling) gateway agents are polling this database for new requests. Once a new request is made, a gateway agent will process it accordingly.

The gateway agent will need to communicate with connected clouds to process the requests. In the case of a cluster creation or scaling up, the gateway agent will speak to more than one clouds to select those that will be used. For virtual machine specific operation (pause, resume, destroy) the gateway agent will need to speak directly with the cloud that runs this virtual machine.

The communication with the clouds is done through cloud agents which are remote processes. Gateway agents and cloud agents are both implemented with the iAgent framework and they communicate through XMPP PubSub channels and directed messages.

4.3. Cloud agent

A cloud agent is the process that will listen to requests from gateway agents and that will eventually make the API calls to the connected cloud. In contrast to gateway agents, cloud agents can speak with a single cloud.

Since cloud agent is a remote process it can be deployed and hosted by the cloud administrator himself. Thus, in order to contribute cloud resources to CernVM Cloud Gateway no cloud API credentials need to be stored centrally. The administrator of the cloud agent is able to define soft quotas for limiting the used resources as well as ACL for defining which users and groups of CernVM Cloud Gateway will be able to use a certain cloud.

Finally the administrator has to define the mapping between the globally defined images, and offerings to cloud specific images and VM flavors. The cloud agent is able to communicate with different cloud APIs as this communication is handled by a module called *cloud driver* which provides a uniform interface to different cloud APIs.

4.4. iAgent framework

iAgent framework [2] is a Perl framework that is used to develop distributed, agent-based systems that communicate through XMPP. The concept is similar to the actors model. Each agent is a process that has its own configuration. It requires XMPP credentials and defines programmatically accessible endpoints that expose its functionality.

iAgent is based on POE Perl library. That allows us to define hooks that are used for message passing between the modules of the agent. Again, each module exports its functionality and iAgent kernel calls for message passing are invoking its callbacks.

4.5. Overflowing clouds

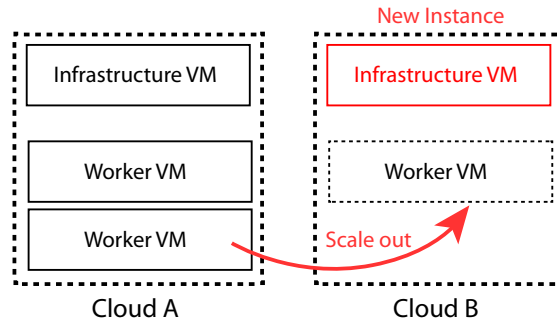


Figure 4. Scaling a cluster beyond the capabilities of the current host cloud

CernVM Cloud Gateway is able to use multiple clouds. Also the definition of a virtual cluster is based on services and dependencies. These two facts are related and combined into CernVM Cloud Gateway feature of scaling up a cluster even if it does not fit to the current used cloud, by expanding it to different cloud(s). The communication between cluster services though has to be ensured and since inter cloud networking might be slow or even not possible, CernVM Cloud Gateway provides all the required services to the scaled service, in the new cloud.

Service replication ensures that the service that needs scaling will be able to locally connect to other required services while it will not be limited to the cloud that was selected in the cluster deployment. Also this enables the usage of multiple cloud at the same time and for the same cluster.

5. Conclusion

We presented the developments in CernVM contextualization processes and the introduction of CernVM Online web portal for managing and sharing contexts. Additionally we described the methods provided to apply a context to local virtual machines, deployed on user's PC, or to CernVM instances that are already running and where user has console access.

Then we proposed a model for defining virtual clusters by separating types of machines into services, having fixed and scalable services as well as setting service dependencies. We showed the CernVM Cloud Gateway which uses this model to deploy and contextualise virtual clusters based on CernVM instances over multiple and different clouds.

Finally we briefly mentioned the CernVM Cloud Gateway feature of overflowing clouds by making use of our flexible model for defining virtual cluster services dependencies and by replicating required services for a cluster instance to new clouds. The later is certainly something worth exploring further to enable the simultaneous use of different, heterogenous and geographically distributed cloud resources.

References

- [1] P Buncic et al 2010 CernVM - a virtual software appliance for LHC applications *Journal of Physics: Conference Series* **219** 042003
- [2] I Charalampidis et al 2012 Managing the virtual machine lifecycle of the CernVM project *Journal of Physics: Conference Series* **396** 032022
- [3] J Blomer et al 2013 Micro-CernVM: slashing the cost of building and deploying virtual machines *Journal of Physics: Conference Series*
- [4] 2012 CernVM Cloud technical report URL <http://cern.ch/cernvm-cloud/static/docs/report.pdf>