CloudSigma AG
Badenerstrasse 549
8048 Zürich, CH
www.cloudsigma.com

CloudSigma Platform Capabilities

CloudSigma is one of the most customizable cloud providers on the market with a focus on open design and flexibility with regards to computing deployments. The cloud platform is designed to provide an environment with the same degrees of freedom that private in-house environments end-users might have are able to offer. All functionality is available via an API or the end-user WebApp.

Customers are able to provision processing, storage, networks and other fundamental computing resources in an unbundled way meaning CPU, RAM, storage and bandwidth can be purchased and combined independently to allow the best combination of cloud resources without the limitation of fixed sizes. So any combination of CPU and RAM can be achieved along with multiple drives mounted to a server and multiple networking interfaces. Each resource is billed separately and transparently as either a subscription or as pay-as-you-go in 5-minute billing segments, enabling customers to track exactly how much their cloud servers are costing over time and be billed for it accurately.

Uniquely, any x86 based operating system and software can be used with complete administrator/root control including all variants of BSD, Linux and Windows. End-users can upload raw ISO image, attach CPU and RAM to it and boot it up. This allows full backwards compatibility from the platform.

All our cloud servers and drives are persistent and modelled around the same methodology as physical dedicated server equivalents - i.e. drives, NICs etc. VLANs and IP addresses are also controlled using standard behaviour and support all types of traffic including multicast and broadcast traffic, which is critical for high availability infrastructure in failover.

By creating a tighter link between the application and infrastructure layer through the virtual abstraction of the hypervisor, the cloud platform exposes a number of power tools allowing end-users to achieve greater performance levels from their cloud servers. For performance sensitive workloads this is especially important. These power tools include the ability to define the virtual core size, to expose NUMA topology and to tweak hypervisor timer settings for maximum performance. Additional power features for example include exposing the full CPU instruction set to the virtual machine, allowing much faster processing of certain calculations. This is ensured by enabling the ‘host CPU’ setting as can be observed below.

For example, an end-user may have a virtual machine of size 20GHz and 32GB of RAM. For an application that benefits from parallel processing, the number of virtual cores can be set to 20 thus giving twenty CPU threads of 1GHz each. On the other hand for an application that’s bound by core speed, the virtual core number would be set to eight, giving eight threads of 2.5GHz each. In this way, small changes can have a profound impact on performance without costing the end-user anything.

A screenshot from the provisioning portal shows an example of some of these power tweaking tools:

Power Tweaking Tools

At an account level, account administrators are able to assign specific access and control rights over certain account related operations using access control policies. This allows one account to grant certain rights to another account (or user). This is done through a ’labeling’ feature. The account administrator may for example label a set of servers ‘project x’ and through an access control policy grant full access to ‘project x’ labelled infrastructure to one or more other users. On the other hand, the account administrator might wish to grant read-only access to billing information for the accounting department. The flexibility and powerful possibilities achieved by pairing labeling with ACLs is clearly evident.

Virtual machines (VMs) can be provisioned in a matter of seconds with a high degree of flexibility and control. The average provisioning time for a new cloud server/VM is 15 seconds. This is critical for running end-user facing web services that need to flexibly provision new resources in line with fluctuating end-user demand.

Users have the option of full API access with all account actions available, i.e. 100% API coverage, allowing complete automation and remote infrastructure monitoring, or the option of a feature-rich, yet intuitive web browser based GUI. The CloudSigma WebApp has been designed to allow easy resource management via any web browser. CPU and RAM can be specified to the nearest MHz and MB.

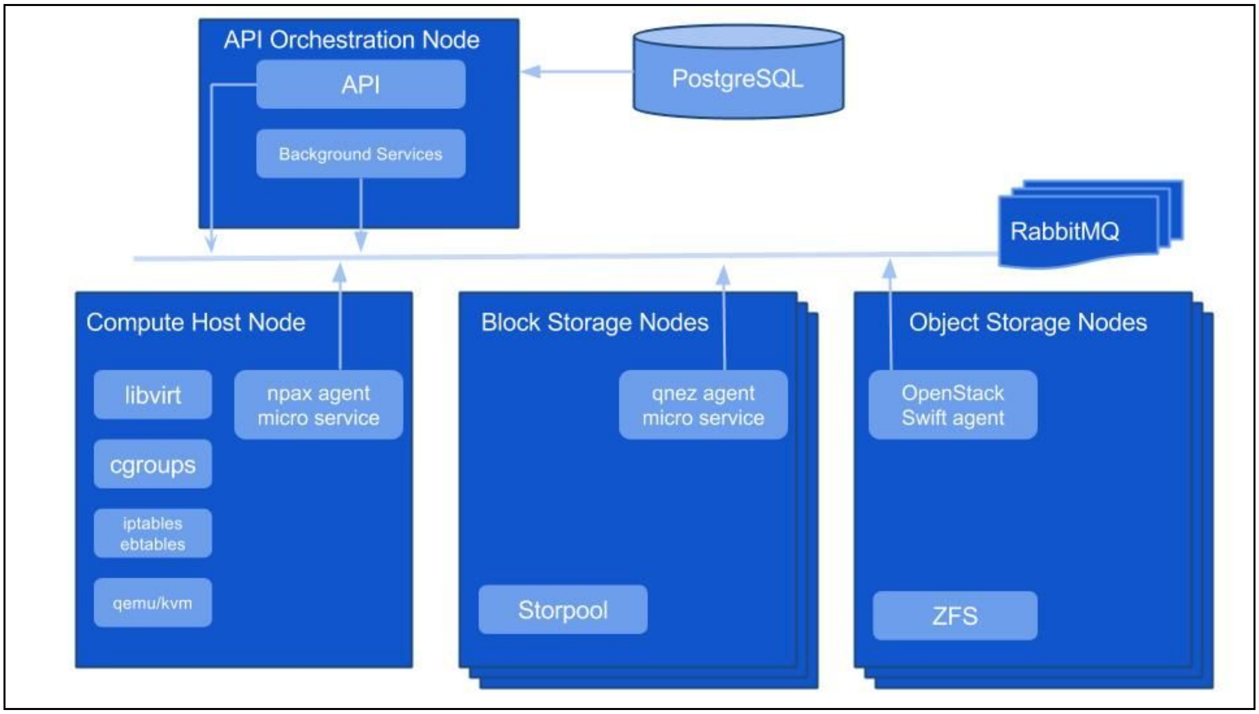
VM provisioning is achieved via a simplified wizard or custom server creation tool. The end-user can choose from a wide selection of ready system images including a number of BSD, Linux and Windows based operating systems as well as being able to quickly and easily upload their own ready ISO image. End-users can customise these marketplace images using cloud initialisation frameworks such as cloudinit (see https://help.ubuntu.com/community/CloudInit), allowing them to take a standard installation and contextualise/customise it to their specific requirements on-the-fly on first boot-up.

Cloud Architecture Overview

The cloud platform is built based upon three distinct layers - the infrastructure layer, the orchestration/management layer and the end-user facing interface layer.

The base layer is the software running on the physical servers, which provide the virtualized resources. Each physical ‘host’ server runs a Linux distribution with the KVM hypervisor exposing virtualized resources. For management libvirt is used alongside npax, a proprietary CloudSigma element plus a custom storage agent for clustered block storage. Load and VM isolation is managed via Linux cgroups.

To communicate and respond to information requests from the next layer up - the management layer, in each compute host we use Libvirt together with the proprietary ‘npax’ agent developed internally within CloudSigma which enables additional functionality not currently available from Linux KVM and/or libvirt. Block storage is delivered from a clustered SSD storage system offering high IOPS performance and low latency. This highly available system is designed to survive multiple storage node failures and to provide an always-on environment that can continue service uninterrupted through software and hardware upgrades.



Software Components in the Cloud Architecture

The clustered SSD block storage service also takes advantage of RDMA and a custom storage agent on each compute host to both lower latencies over the network (using RDMA) and speed up throughput by bypassing the standard Linux networking stack and lesser performing standards such as iSCSI by using a custom storage agent. Object storage is delivered using the OpenStack Swift object storage environment and API. This environment is a custom implementation built on top of an Exabyte scale ZFS driven system using OpenSolaris. This custom deployment still exposes the standard open source Swift API to end-users but takes advantage of ZFS' compression and deduplication abilities and it's powerful read caching to deliver both efficiency and performance. This is critical for storing large scale data archives meant for active work loads.

The management layer consists of the resource allocation logic, customer records management, and billing. This layer has been entirely coded from scratch by CloudSigma. The management layer is responsible for our ability to expose a range of unique functionalities and freedoms to users. Our utility pricing around simple charging per resource unit in short 5-minute cycles also derives from this management layer. The glue between the physical and management layers is the messaging bus which uses RabbitMQ and Celery. The ability to deliver high uptime and a low failure rate on user infrastructure requests derives in no small part from having a very robust messaging layer implementation. Records are stored using a clustered PostgreSQL database with individual services such as monitoring or billing ran as independent micro-services in Docker containers. This architecture allows easy scaling and resource prioritisation within the cloud management layer, delivering high uptime and consistent performance/responsiveness to end-users.

The third and final layer is the public layer, which provides two primary interfaces, which allow customers to control and manage their cloud infrastructure. These two public interfaces consist of the public web provisioning portal and the public API. Furthermore, we offer a wide range of 'wrappers' that allow compatibility with other mainstream IaaS APIs, such as OpenStack. We offer a 'full control' API meaning that all account functions are available via the API and can thus be fully automated. The WebApp interface uses technologies such as websockets to provide a live environment that automatically pushes infrastructure status changes to customers.

As a public multi-tenanted cloud, we provide an open computing environment for customers, whilst simultaneously protecting the cloud and existing customers from malicious behaviour. Critically, despite combining a number of open source and closed source tools, the CloudSigma cloud imposes no proprietary burden or requirements on end-user computing, allowing end-users to import and export their data and applications without vendor lock in.

The cloud platform's API is implemented using a RESTful interface, using the URL to specify an object and an action, HTTPS GET to read state and HTTPS POST to change state. The API is covered by all major cloud API abstraction libraries including jclouds, fog.io, LibCloud, Ansible, Abiquo, Flexiant and Golang. The cloud platform works with both Canonical and Flexiant to enable the on-demand deployment of software components and whole stacks using Canonical's JuJu and Flexiant's Concerto platforms. Both architectures support service monitoring, auto-scaling and load balancing.

CloudSigma has developed enabling technologies to facilitate the use of these orchestration platforms. The cloud platform offers support for OpenStack HEAT, allowing entire service stacks to be brought up using AWS and HOT templates and using any OpenStack HEAT compatible tools. With the support of these tools, full automation of CloudSigma's cloud infrastructure is possible, including automatic scaling, load balancing and more. We offer a 'full control' API meaning that all account functions are available via the API and can thus be fully automated.

Cloud Resources Availability & Capacity Management

This section describes the cloud resources capacity management.

VM sizing range and configuration

The following resource ranges are available in relation to virtual machine and drive sizing:

Resource Type	Minimum	Maximum
CPU	1 core / 1GHz	20 cores / 60GHz
RAM	1GB	128GB
Block Storage (one drive)	1GB	10TB
Object Storage (one object)	100MB	2TB

Note: A VM can have up to sixteen separate block storage devices attached to it.

The proposed solution exposes in a utility manner CPU and RAM to users. End-users can freely combine CPU and RAM to the nearest MHz and MB respectively subject to the minimum and maximum ranges outlined in the table above. It is therefore possible to size CPU independently of RAM on a VM by VM basis. The initial VM configuration can be easily scaled up and down as required by the end user, i.e. adding or removing CPU or RAM.

The CloudSigma cloud supports on demand and subscription based consumption of computing resources. In the case of on-demand resource consumption, the cloud billing system monitors usage per compute resource (i.e. CPU, RAM, storage etc.) and compares this against the end user's subscribed resource level (which may be 0). This check is conducted every five minutes. The billing system then charges the user for any resource consumption amount above the subscribed level in five minute segments. In this way a user will get charged for burst usage within five minutes of spinning up the VM and, within five minutes of shut down the system will detect this change and stop charging for those burst resources. Actual burst prices vary depending on the level of utilization of the cloud. The cloud platform uses an advanced pricing algorithm that allows our pay-as-you-go burst pricing to change over time depending on how busy our cloud is and whether the resources are being provisioned in peak or off-peak times.

Virtual Machines Provisioning Times

Virtual provisioning times for virtual machines do not vary by size. The average provisioning time is 15 seconds. The cloud provisioning layer is able to provision multiple virtual machines in parallel. The bulk aggregate provisioning times are estimated below:

Number of New VMs being Provisioned	Estimated Delivery Time
1	15 seconds
100	2.5 minutes
500	12.5 minutes

Virtual machine creation, spin-up and spin-down times are identical. The above table can be used as a guideline for typical lead times to be expected when provisioning VMs in volume. As you can see, VM provisioning times scale linearly on CloudSigma's high-performance stack. The billing system is based on five minute billing segments, so VMs once shutdown will stop being billed for within a maximum of five minutes.

Virtual Machine Templates

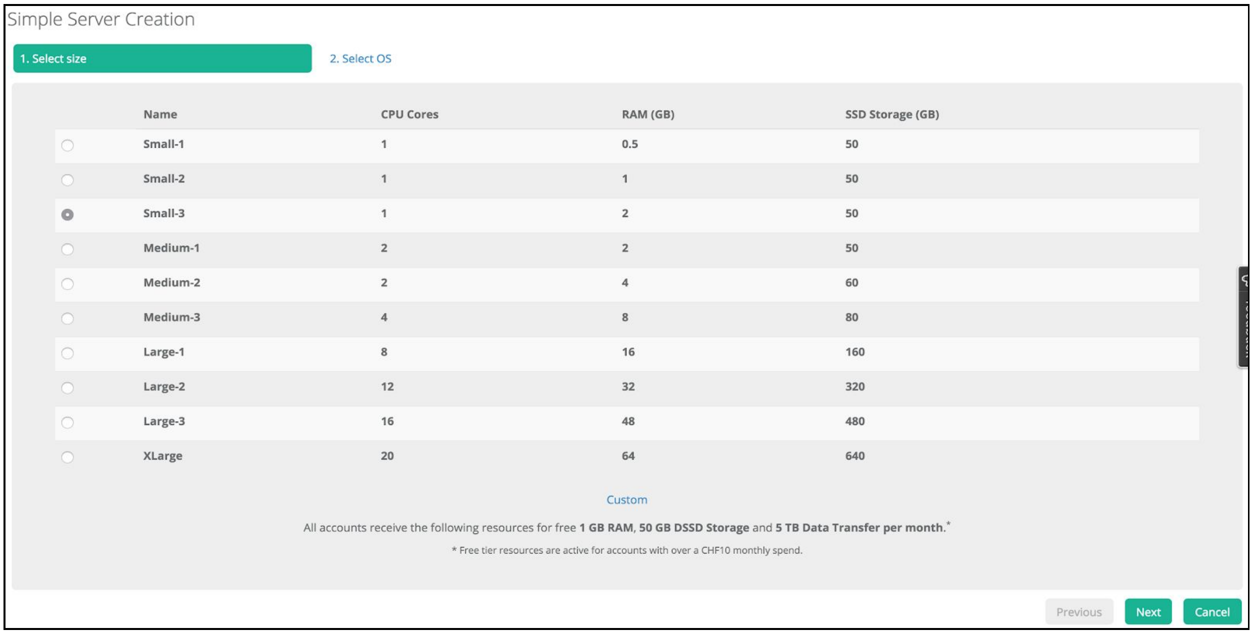
CloudSigma offers two approaches to VM creation - a quick and simple creation process and a detailed, flexible and extremely granular custom one. Both are outlined below.

- *Simple Server/VM Creation Process*

Using this process, an end-user can quickly create a VM from a predefined list of sizes and available operating systems and applications. VM settings are optimised automatically by the cloud based on the user's choice of operating system.

Next, the user can choose from the operating system they wish to deploy. After this final step, the virtual machine will automatically be created with associated storage and networking and will be booted up. Details such as the username to be used when logging in to the server, and the VM's public IP address is succinctly communicated with the end-user. Where appropriate, a VNC connection in the browser will be opened to allow the user immediate access.

The quick server/VM creation process takes less than 30 seconds with a provisioning time of 15 seconds. A new VM can be configured, provisioned and accessible in under one minute using this method.



Simple Server Creation Sizing Menu

- *Custom Server/VM Creation*

For users wishing to specify their VMs in more detail, the custom server creation tool is available, allowing users to specify every aspect of their VM exactly how they would like it. Firstly, they are able to choose the exact CPU / RAM sizing and give the VM a name:

Customer Server Creation Sizing and Options

Next, the user can add one or more SSH keys in order to enable secure access from from the get-go (in the case of Linux and BSD based systems). Once this step is completed, it is possible to create and/or attach and mount one or more drives of the user's choosing to the new VM. After this, the networking set-up can be defined. Multiple public and private networking interfaces can be assigned to a single VM. Finally, the end-user can tweak the advanced settings of the virtual machine to maximise its performance in relation to the specific application requirement that they have.

Cloud Storage Capacity Management

We provide more details on both storage systems in the sub-sections below.

Block Storage Environment

The cloud platform solution provides the ability to users to create and (re)size multiple persistent virtual drives. It is an SSD-based storage system perfectly suited to the high-performance requirements associated with big data mining and scientific computing requiring high performance storage. Below is a screenshot of the end-user provisioning portal showing the current drives in the user's account:

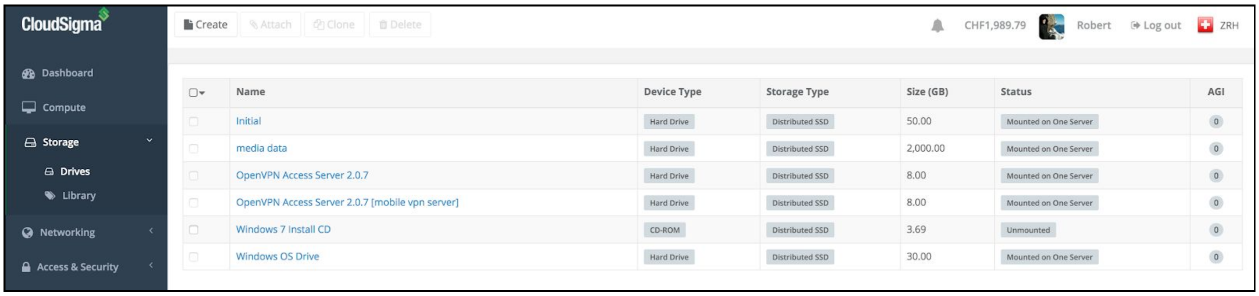
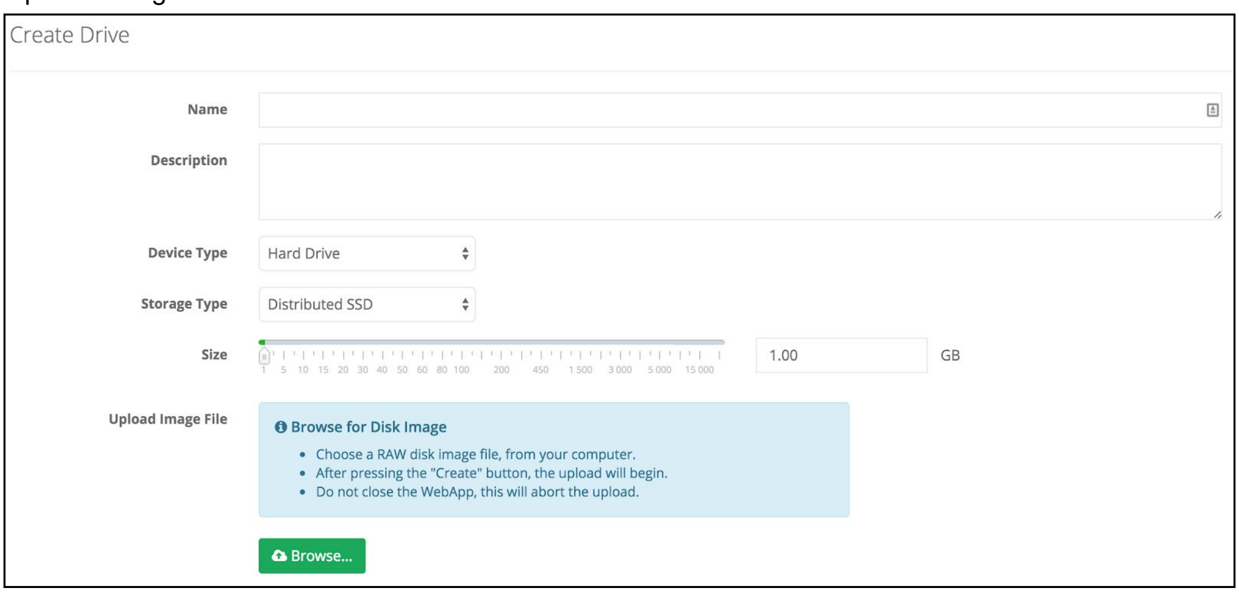


Figure WebApp Provisioning Portal

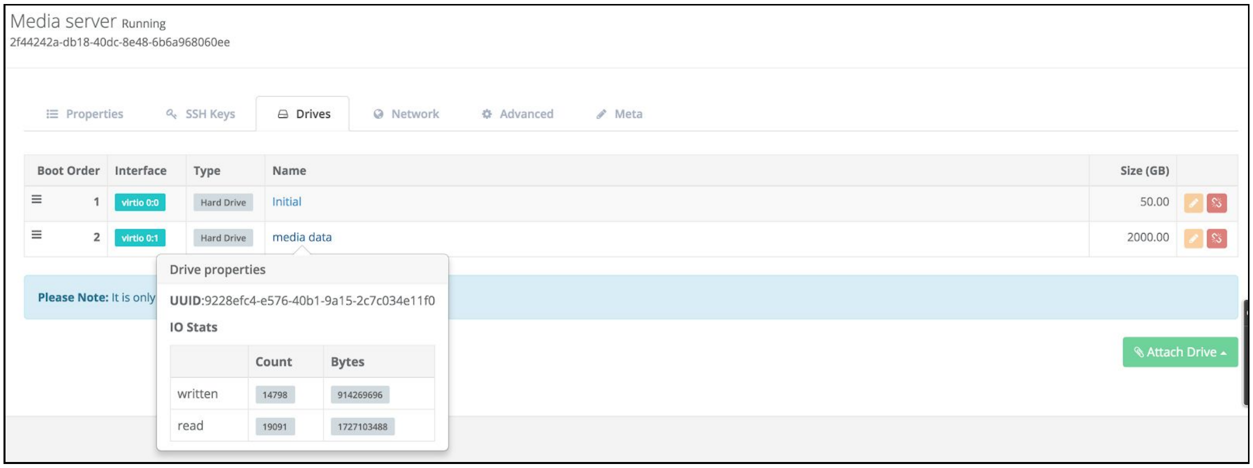
Drives can be mounted or unmounted. The system supports the creation of regular drives as well as CDs / DVDs. Users can directly upload RAW images of CDs and hard drives, complementing the extensive selection of install CDs and pre-installed systems available from the marketplace.

There is no fixed drive sizing, users can create drives subject to a minimum and maximum size in a granular way. All storage is persistent and the provisioning time in creating a new drive is 10-20 seconds with an average provisioning time of 15 seconds, similar to starting or creating a new virtual machine. Once created, a drive can be resized subject to the maximum drive size as outlined in the figure below. Drive resizing is executed instantly with a provisioning time of 10-20 seconds for the increased drive size.



WebApp Storage Creation Section

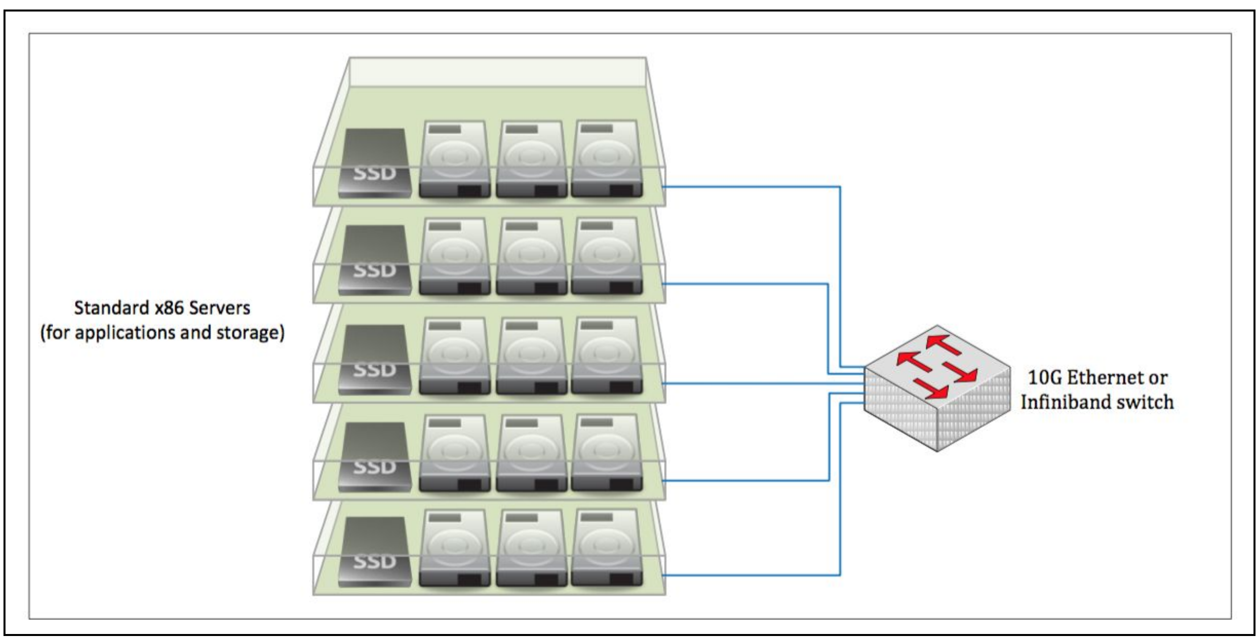
Unlike less flexible cloud systems that bundle storage and compute resources together in pre-defined sizes, the cloud platform allows multiple drives to be mounted to a single VM. In the example VM below, two drives are attached to the same cloud server. One smaller one (50GB) for the operating system and a second drive for media data (2TB). Users can also track read and write activity as shown in the pop-up from the screenshot.



Drives Section

The block storage system is driven by StorPool - an advanced proprietary high-performance distributed storage technology chosen by CloudSigma after extensive research into best-in-breed block storage solutions. It pools the attached storage (hard disks or SSDs) of standard servers in order to create a single pool of shared block storage. The software is installed on each server in the cluster and combines the performance and capacity of all drives attached to the servers into one global namespace.

The system provides standard block devices which can then be attached to virtual machines/cloud servers. In this implementation, the system has been configured with no file system in order to use volumes to store VM images directly.

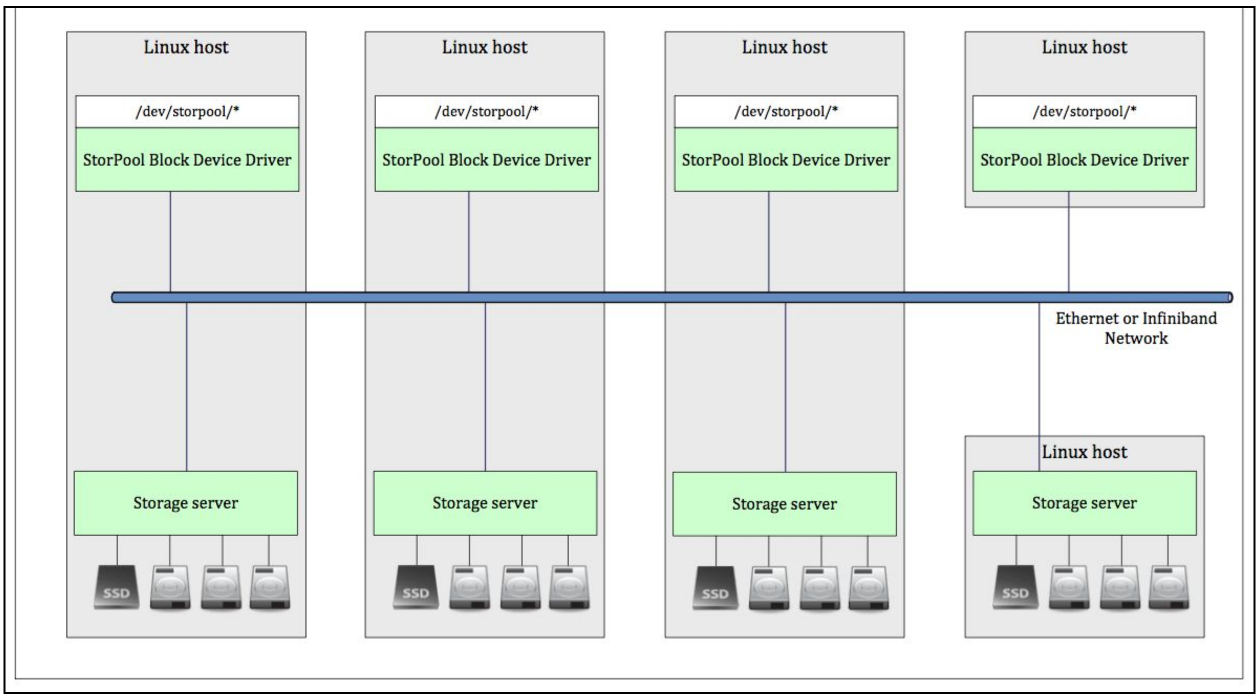


StorPool Shared Block Storage Solution

Redundancy is provided by multiple copies (replicas) of the data written synchronously across the cluster. Users can set the number of replication copies, with the CloudSigma cloud configured to store three copies of all data. This technology is superior to RAID in both reliability and performance.

Unlike RAID, our system replication distributes copies across different servers. As such, in the case of a server or component failure, data that is stored on this affected server is not lost.

Rebuilding a failed drive/node is done from multiple drives (and not from one particular drive as is the case with RAID). What this means is that rebuild times are significantly shorter and they do not impact overall system performance. The system also protects data and guarantees data integrity via a 64-bit checksum and version for each sector maintained by the system. The system provides a very high degree of flexibility in volume management. Unlike other storage technologies, such as RAID, the system does not rely on device mirroring (pairing drives for redundancy). So every disk that is added to the cluster adds capacity to the cluster, not just for new data but also for existing data. Provided that there are sufficient copies of blocks, drives can be added or taken away with no impact to the storage service. Unlike rigid systems like RAID, the system does not impose any strict hierarchical storage structure that links and reflects to the underlying disks. The system simply creates a single pool of storage that utilises the full capacity and performance of a set of commodity drives. This allows this consortium to easily scale out storage requirements as the cloud data ecosystem grows over time.



Data Replication and Redundancy

In summary, the proposed block storage system provides a flexible yet simple interface to users whilst combining advanced technology on the back-end to delivery high performance and extreme storage resilience.

Block Storage Backup Solution

The block storage system used for the proposed cloud solution by CloudSigma is used for all the cloud storage utilised by user VMs. This block storage solution is able to provide implicit non-disruptive backups at the storage block level for all user data. This includes any data contained within virtual drives including application data, databases, all operating system information etc. It provides full drive level backup of customer data. The back-up system backs up all end-user computing data each night and retains seven days of rolling snapshots. End-users can roll back one or more of their virtual drives to any of the seven off-site data snapshots at any time. Additionally, they can manually request a backup of a virtual drive at any time via the API or WebApp.

Due to the clustered three copy block storage system used, the back-up process does not impact service availability or performance, but instead utilises free capacity to sync data for the backups in the background. End-users can therefore have a continuously running computing environment and at the same time have rolling backups of all their data.

In addition to the automatic backup system, end-users are able to create point-in-time snapshots of their drives, which can later be cloned and upgraded to create stand-alone drives. A snapshot can be created on-demand while the server is running, thus in no way affecting the performance or availability of the systems. By using snapshots, customers can protect themselves from data corruption or use them for auditing purposes.

Furthermore, there is an advanced snapshot management feature, allowing customers to create snapshot management policies and apply them to one or more drives. In this manner, customers are able to automate the snapshot process.

Supported Software & Tools

CloudSigma's stack was chosen as the cloud stack to facilitate hosted processing as it offers an open API and a very high degree of compatibility with end users who have legacy requirements. There are absolutely no restrictions on the type of operating system or application that the customer can deploy, as long as it is x86 compatible.

The following section presents the options available in CloudSigma cloud platform, which can be deployed and configured instantly.

Operating Systems Support

The cloud platform provides an extensive library of ready-to-use public drive images that include pre-installed systems that can be deployed instantly, alongside install CDs for users preferring to install their own systems from scratch. The following table provides a sample of the available drive images in the marketplace:

Linux distributions	CentOS 5.X, 6.X, 7.X Debian 5.X, 6.X, 7.X, 8.X Fedora 19, 20, 21, 22 Linux Suse 11, 12 RedHat Enterprise 5.x, 6.X, 7.x Ubuntu 12.X, 13.X, 14.X CoreOS, CloudLeap, Elastix, Finnix, Knoppix, OpenVPN, pfSense, Vyatta
Windows distributions	MS SQL Server Standard MS SQL Server Enterprise MS SQL Server Web VirtIO Drivers for Windows Windows Server Datacenter Windows Server Standard 2008, 2012
Solaris distributions	Oracle Solaris 11.X
Other drive images	FreeBSD 8.X, 9.X, 10.X GParted NetBSD 6.X OpenBSD 5.X

In addition, instead of restricting users to a curated list of predetermined images, we have provided customers with a simple and easy way to import and export data and ready operating system images into and out of the system. End-users can use their own images and upload them (or download them) directly via the end-user provisioning portal or over the API. This method also allows for pausing and resuming uploads and downloads.

End-users always have sole root/administrative access over their VMs. This approach not only eliminates vendor lock-in but allows the chosen cloud system to be fully backwards compatible with end-users able to run any version of an operating system they so wish rather than choosing from a fixed list and be forced to upgrade periodically as that list changes.

Middleware Components / Available Tools

CloudSigma offers a range of ready middleware components for a variety of end user markets. These middleware tools are exposed via a middleware services marketplace that's integrated on top of the cloud platform's infrastructure cloud layer.

- In the case of applications, they deploy into an auto-scaling platform-as-a-service environment already patched for the latest updates making them ideal for web services. All the application templates are instantly deployable into a CloudSigma Infrastructure account, offering auto-scaling and load balancing as the demand on computing resources varies through their use.The applications include Apache, Docker, Hadoop, Joomla, MongoDB, Ruby and others.
- In the case of libraries, drivers and other integrations, they are used in conjunction with other services to build wider deployments. These include OpenStack Heat, CloudInit, CoreOS, Apache jclouds Driver, Apache Libcloud Driver and others.

API Support

The proposed solution exposes 100% of all functionality available via APIs allowing full integration into wider software deployments. The core infrastructure API is provided by the CloudSigma cloud stack, however this is complimented by interoperability with a number of open source and other industry standard APIs for wider user engagement and accessibility. The consortium wishes to bring to bear the value of innovation from a commercial provider such as CloudSigma with the benefits of federation and interoperability.

The object storage environment is exposed using the OpenStack Swift interface and the compute orchestration layer itself covered by CloudSigma's own API is also available using OpenStack Heat. The object storage environment also provides additional accessibility via the S3 AWS API.

As outlined earlier, a number of commercial and open source cross cloud drivers are available that allow the management of CloudSigma in conjunction with other leading public and private cloud stacks, both commercial and open source. Overall the consortium feels that the vast majority of end users will find a tool, integration or direct API interface that fits within their programming language and legacy computing requirements in order to minimise friction to onboarding new users to the platform.