

## Accounting in the CloudVeneto private cloud

*Paolo Andreetto<sup>1</sup>, Fulvia Costa<sup>1</sup>, Alberto Crescente<sup>1</sup>, Sergio Fantinel<sup>2</sup>, Federica Fanzago<sup>1</sup>, Loris Lazzaro<sup>4</sup>, Paolo Emilio Mazzon<sup>3</sup>, Matteo Menguzzato<sup>4</sup>, Gianpietro Sella<sup>5</sup>, Massimo Sgaravatto<sup>1</sup>, Sergio Traldi<sup>1</sup>, Marco Verlatto<sup>1</sup>, and Lisa Zangrando<sup>1</sup>*

<sup>1</sup>INFN - Sezione di Padova, Via Marzolo 8, 35131 Padova, Italy

<sup>2</sup>INFN – Laboratori Nazionali di Legnaro, Viale dell’Università 2, 35020 Legnaro (Padova) Italy

<sup>3</sup>Università di Padova, Padova Neuroscience Center (PNC), Via Orus 2/b, 35131 Padova, Italy

<sup>4</sup>Università di Padova, Dipartimento di Fisica e Astronomia “Galileo Galilei”, Via Marzolo 8, 35131 Padova, Italy

<sup>5</sup>Università di Padova, Dipartimento di Scienze Chimiche (DISC), Via Marzolo 1, 35131 Padova, Italy

**Abstract.** CloudVeneto is a private cloud implemented as the result of merging two existing cloud infrastructures: the INFN Cloud Area Padovana, and a private cloud owned by 10 departments of University of Padova. This infrastructure is a full production facility, in continuous growth, both in terms of users, and in terms of computing and storage resources.

Even if the usage of CloudVeneto is not regulated by a strict pay-per-use model, the availability of accounting information for such infrastructure is a requirement, to detect if the resources allocated to the user communities are efficiently used, and to perform an effective capacity planning.

We present in this paper how the accounting system used in CloudVeneto evolved over time, focusing on the accounting framework being used now, implemented by integrating existing components.

## 1 Introduction

CloudVeneto [1] is the result of the integration, performed in 2018, of two private cloud infrastructures:

- the Cloud Area Padovana [2];
- a Cloud owned by 10 departments of University of Padova [3] (in the rest of this article we will refer to this infrastructure as ‘University cloud’).

The Cloud Area Padovana is a cloud infrastructure owned by INFN (Istituto Nazionale di Fisica Nucleare) in production since the end of 2014. It includes resources spread between two INFN units: ‘Sezione di Padova’ and ‘Laboratori Nazionali di Legnaro’ (LNL). The University cloud was born about one year later than the Cloud Area Padovana. The University cloud and the Cloud Area Padovana had many similarities in their architectures and implementations, but were two independent cloud instances.

To limit the manpower for the operations of these two cloud infrastructures and to avoid the useless duplication of some services, it was decided to integrate the two clouds into a unique infrastructure which was called CloudVeneto. The details of how such integration, finalized in 2018, was performed are discussed in [4].

CloudVeneto is now a computing facility in full production. There are currently about 330 registered users, and the number of people using such infrastructure keeps increasing. Also the amount of computational and storage resources made available to users is constantly growing.

CloudVeneto is not a pay-per-use infrastructure, still, in order to check if resources have been properly allocated among the supported user communities, it is very important to properly track how efficiently the CloudVeneto resources are being used. Resource tracking is also fundamental to be able to effectively perform a capacity planning of the infrastructure and to decide if (and how many) new resources are needed.

The rest of this article is organized as follows. Section 2 describes the CloudVeneto infrastructure: it discusses about its architecture and its implementation, providing also some information about its usage. Section 3 presents the accounting system used in CloudVeneto, and how it evolved over the last years. Section 4 discusses about some potential future activities. Section 5 concludes the paper.

## 2 CloudVeneto: design and implementation

CloudVeneto is using OpenStack [5] as Cloud middleware framework (at the time of writing this article, the Train release is deployed). CloudVeneto is exposed to users as a unique Cloud facility but with resources physically deployed in two distinct sites: one in Padova (a data center shared between INFN Padova and the Physics Department of the University) and one at the INFN-LNL, which are about 10 km apart and connected by a 10 Gbps dedicated network link.

The Openstack software has been augmented with some in-house developments, in particular to manage user authentication (through the integration with the INFN and University of Padova identity providers) and user registrations [2] [4].

CloudVeneto basically implements a IaaS (Infrastructure as a Service) cloud, allowing users to create instances sized, customized and configured for their specific use cases. However users are also provided with some higher level services, such as scalable batch systems and kubernetes clusters, eventually enriched with a set of services for big data analytics.

Table 1 summaries the resources currently integrated in CloudVeneto, provided by the different organizations.

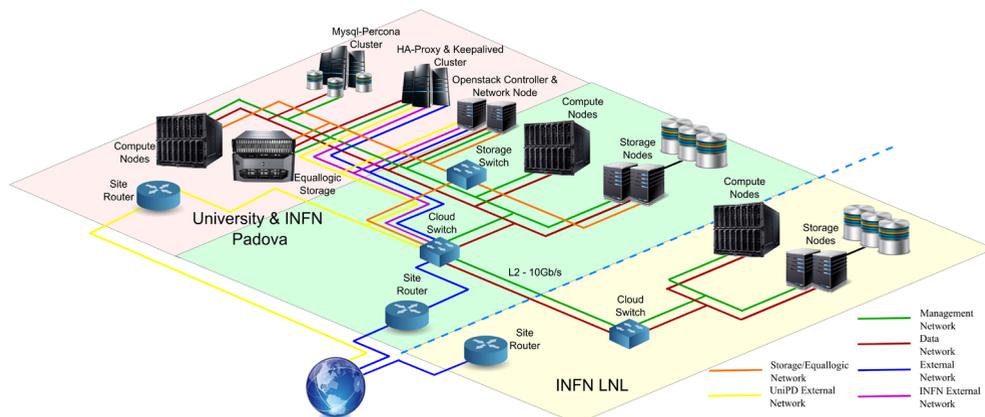
**Table 1.** CloudVeneto resources provided by the different organizations

Organization	Compute nodes	Cores	GPUs	RAM (GB)	Storage (TB)
INFN-Padova	23	1104	14	4832	430
INFN-LNL	20	736	0	2784	0
Unipd	22	1024	6	4448	240
Totals	65	2864	20	12064	670

The architecture of CloudVeneto is shown in figure 1.

Compute nodes (the hypervisors where the Cloud virtual machines are instantiated) are present in both sites (with respect to table 1, the resources listed as “INFN-LNL” are physically installed in the Legnaro INFN-LNL site; all the others are instead in the Padova computing room). Cloud services and storage are instead installed only in Padova.

For what concerns storage, the OpenStack block storage service (Cinder) is configured to use multiple backends:



**Figure 1.** The CloudVeneto architecture

- a ceph [6] pool using replica 3;
- a ceph pool configured using erasure code 5+2;
- an iSCSI storage server, configured using the Cinder NFS driver (this backend is now being dismissed);
- a DELL EqualLogic system, configured using the Cinder EqualLogic driver.

The first 3 backends are available only to INFN users, while the EqualLogic storage serves only University users.

The ceph cluster is also used for the Openstack image service (Glance), and to provide an object storage system (accessible through the S3 and swift protocols).

A similar usage model was also implemented for the computational resources: as imposed by the management, by default compute nodes owned by the University should be usable only for instances created by University users, and vice versa. This was implemented relying on host aggregates and on the AggregateMultiTenancyIsolation OpenStack scheduler filter, which allow to easily manage also different allocation models, if needed.

CloudVeneto also provides 20 Nvidia GPUs of different models (Tesla T4, V100, Quadro RTX 6000, Titan XP, GeForce GTX). They were integrated in the OpenStack cloud through the PCI passthrough feature.

Networking in CloudVeneto is implemented through the OpenStack Neutron service, configured to use Open vSwitch and GRE [10] (Generic Router Encapsulation).

The cloud services were deployed in CloudVeneto in two controller-network nodes in an active-active configuration, therefore providing a high reliable system. A HAProxy [12] - Keepalived [13] cluster, composed of three instances, is used to distribute the load among the two instances and to manage the fault tolerance of the services. HAProxy is also used to expose OpenStack services through an SSL compliant interface, and as frontend interface to the database system, implemented through a three-instance Percona XtraDB cluster [11]. The messaging system (needed by the cloud services) was also configured in high availability, relying on a RabbitMQ cluster composed of three nodes.

CloudVeneto is currently used by about 330 users, organized in about 80 projects (approximately equally distributed between INFN and University). This infrastructure is mostly

used to support a large range of different scientific and engineering disciplines (some of them are reported in [7]) but it is also used for other use cases (e.g. to support University teaching activities, to provide resources for the development and testing of new software, to support training events, etc).

### 3 The CloudVeneto accounting system

CloudVeneto (and previously the Cloud Area Padovana) used to rely on the OpenStack ceilometer and gnocchi [9] services to manage accounting information (in more detail, gnocchi was used to store and manage the accounting metrics collected by ceilometer).

To overcome some issues related on how the accounting information were handled by ceilometer and to be able to present the accounting data according to our specific needs, we implemented a new service, called CAOS [8]: by gathering data from the ceilometer service, CAOS has been used for some years to manage and present in a consistent way data concerning resource usage.

Unfortunately we later found some scalability problems operating the ceilometer service. Such issues, added to the uncertainty about the future of the ceilometer and gnocchi services, made us decide to look for alternative solutions for accounting.

With the main goal of limiting as much as possible new home-made developments and instead relying on tools and services that have a good chance of being supported by the community at least for the short-medium future, we decided to assess the accounting system used in the operations of the EGI Federated Cloud [14]. Such accounting system, outlined in figure 2, foresees that each cloud instance composing the federation is instrumented with an accounting reporter, which for OpenStack is provided by the cASO [15] service.

cASO queries the OpenStack nova service, and produces accounting records which are sent, using the STOMP protocol, to a message broker. Such messages are then read and digested by an APEL [18] server, which stores and processes accounting data from all sites of the federation. The APEL server creates summaries on a daily basis from all the received accounting records. A web view of these summarized accounting data is then provided through an accounting portal.

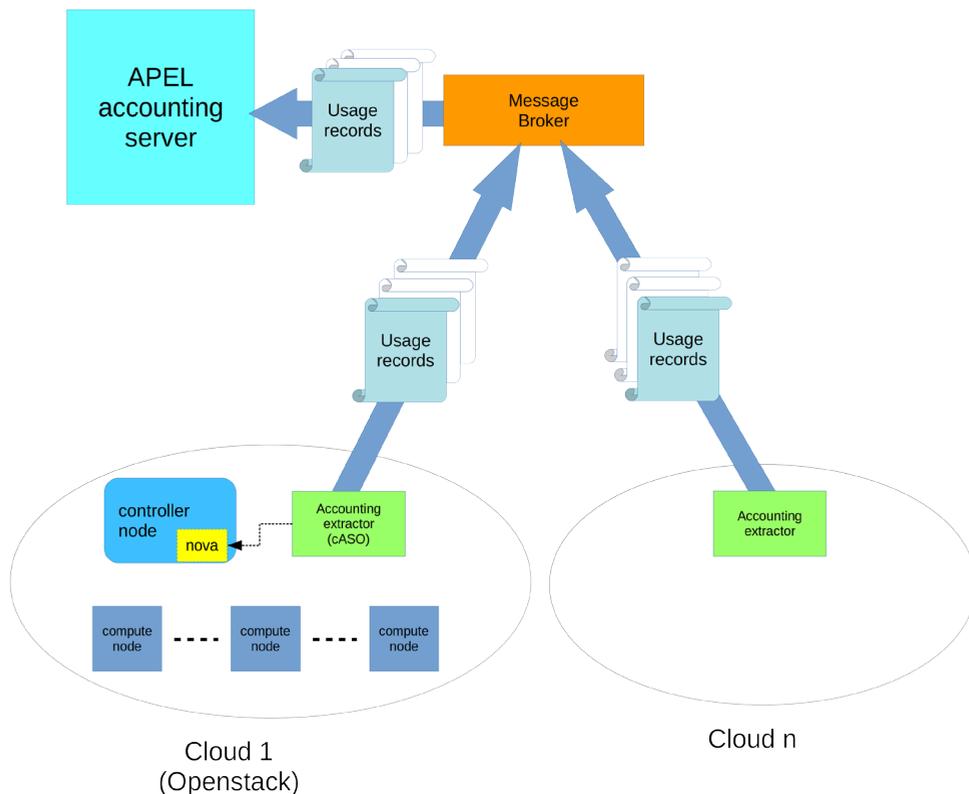
Accounting data collected by cASO are compliant with a Cloud Usage Record agreed among the EGI partners, which inherits from the OGF Usage Record [16]. An accounting record defines the resources used by a cloud instance, in terms of consumed computing power, memory, disk, network, etc. In particular such usage record includes the WallDuration (the wall clock time) and the CpuDuration (the CPU time consumed by the instance), which are the two metrics most relevant for the CloudVeneto stakeholders.

Unfortunately the CpuDuration metric collected by the cASO software, configured using the nova extractor, is not the real CPU time consumed by the instance, but it is simply calculated as the product of multiplying the WallDuration by the number of CPUs allocated by that instance.

To overcome this problem, we modified the accounting system as shown in figure 3 (the changes with respect to the EGI accounting system are highlighted in red).

Therefore each Openstack compute node is instrumented with the collectd [17] daemon, configured with the virt plugin, which allows to collect different metrics related to the virtual instances hosted on the hypervisor, including the real CPU time consumed by each cloud instance.

Collectd was then configured to send these data to a Influxdb instance.



**Figure 2.** Architecture of EGI FedCloud accounting system

The last needed change was modifying the cASO code, to get the CpuDuration metric querying the Influxdb database: this was a very simple modification.

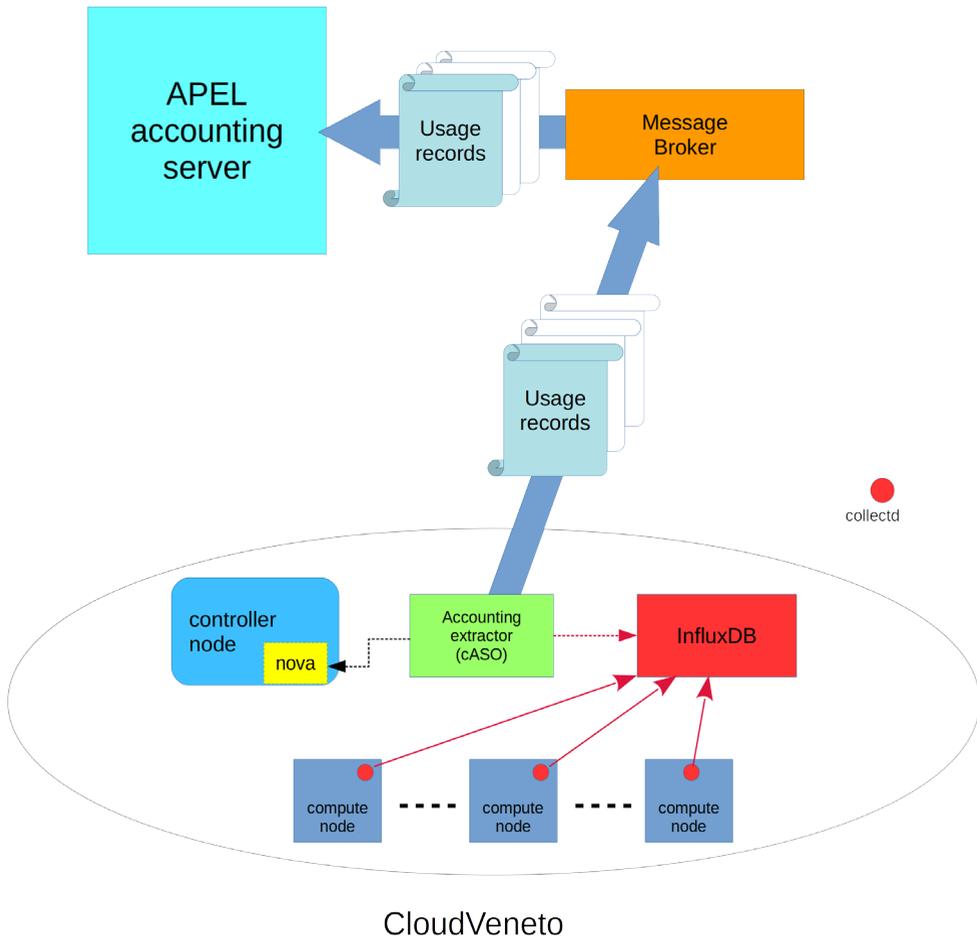
By mimicking the EGI architecture, cASO was configured in CloudVeneto as if the accounting records were pushed by two different sites: the CloudVeneto INFN partition and the CloudVeneto University partition. This allows to easily produce separate accounting reports for the two CloudVeneto stackholders, as requested by the management.

To graphically show the accounting data, Grafana was interfaced with the relational database of the APEL server storing the summarized accounting information. This allows to produce graphs showing the consumed CPU Time and Wall Clock Time over a certain period, with the possibility to select specific partitions (University and INFN) and/or specific user communities (which map to OpenStack projects) and/or specific users. An example of a graph that can be produced through this Grafana dashboard is shown in figure 4

## 4 Future work

There are several foreseen future developments. The first one is trying to push upstream the changes done in the cASO code so that it won't be needed to manually modify the code (even if this change is very simple).

Another foreseen work is try to augment the produced accounting records also with some benchmark (e.g. HepSpec) of the instances, which is something already foreseen in the usage



**Figure 3.** Architecture of CloudVeneto accounting system

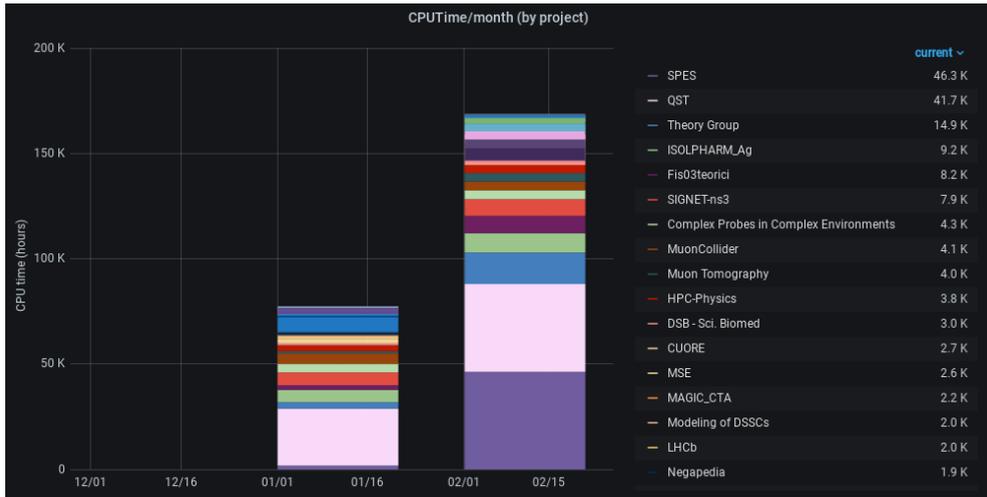
record. This should be quite straightforward, since the InfluxDB database records the name of the hypervisor where each cloud instance ran.

New developments will be needed also to track usage of GPU resources, which is something already being discussed also in the EGI community.

## 5 Conclusions

We presented here an accounting system implemented by integrating existing components, and which required very few modifications. The integrated components are widely used in other important infrastructures and have therefore a good chance to be kept supported and maintained for the foreseeable future.

The accounting information obtained by the implemented system allow us to detect if the resources are being used efficiently, or if some reallocations of the resources should be done.



**Figure 4.** Example of accounting information visualized through Grafana

The system described in this paper is also being implemented as accounting framework for the INFN-Cloud infrastructure [19].

## 6 Acknowledgements

We thank the APEL and cASO developers for their support.

## References

- [1] CloudVeneto home page, <http://cloudveneto.it>
- [2] P. Andreetto et al, **The Cloud Area Padovana: from pilot to production**, J. Phys.: Conf. Ser. 898 052007
- [3] P. E. Mazzon, M. Menguzzato, G. Sella, M. Sgaravatto, **Progetto CloudVeneto.it - Status Report**, Technical Report, May 2016
- [4] P. Andreetto et al, **Merging OpenStack based private clouds: the case of CloudVeneto.it**, DOI: <https://doi.org/10.1051/epjconf/201921407010>, in Proceedings of CHEP2018 conference.
- [5] Home page for the OpenStack project, <http://www.openstack.org>
- [6] Home page for the Ceph project, <http://ceph.com/>
- [7] <http://www.cloudveneto.it/testimonials.php>
- [8] F. Chiarello, **CAOS: a tool for OpenStack accounting management**, DOI: <https://doi.org/10.1051/epjconf/201921407006>, in Proceedings of CHEP2018 conference.
- [9] Gnocchi – Metric as a Service, <https://gnocchi.xyz/>
- [10] Hanks S et al., **Generic routing encapsulation (GRE)** (2000). <http://tools.ietf.org/html/rfc2784.html>
- [11] Percona XtraDB Cluster, <https://www.percona.com/software/mysql-database/percona-xtradb-cluster>

- [12] Tarreau W, **HAProxy-The Reliable, High-Performance TCP/HTTP Load Balancer**, <http://haproxy.1wt.eu>
- [13] Cassen A, **Keepalived: Health checking for LVS and high availability**, (2002), <http://www.linuxvirtualserver.org>
- [14] The EGI Federated Cloud, <https://www.egi.eu/federation/egi-federated-cloud/>
- [15] López García, Álvaro; Fernández-del-Castillo, Enol, **cASO: OpenStack Accounting extractor**; DIGITAL.CSIC; <http://doi.org/10.20350/digitalCSIC/8595>
- [16] R Mach et al., **Usage Record - Format Recommendations**, Open Grid Forum GFD-R-P.098.
- [17] Collectd official website, <https://collectd.org>
- [18] M. Jiang, C. Del Cano Novales, G. Mathieu, J. Casson, W. Rogers, J. Gordon, **An APEL Tool Based CPU Usage Accounting Infrastructure for Large Scale Computing Grids**. In: Data Driven e-Science. Springer, New York, NY
- [19] INFN-Cloud home page, <https://www.cloud.infn.it/>