



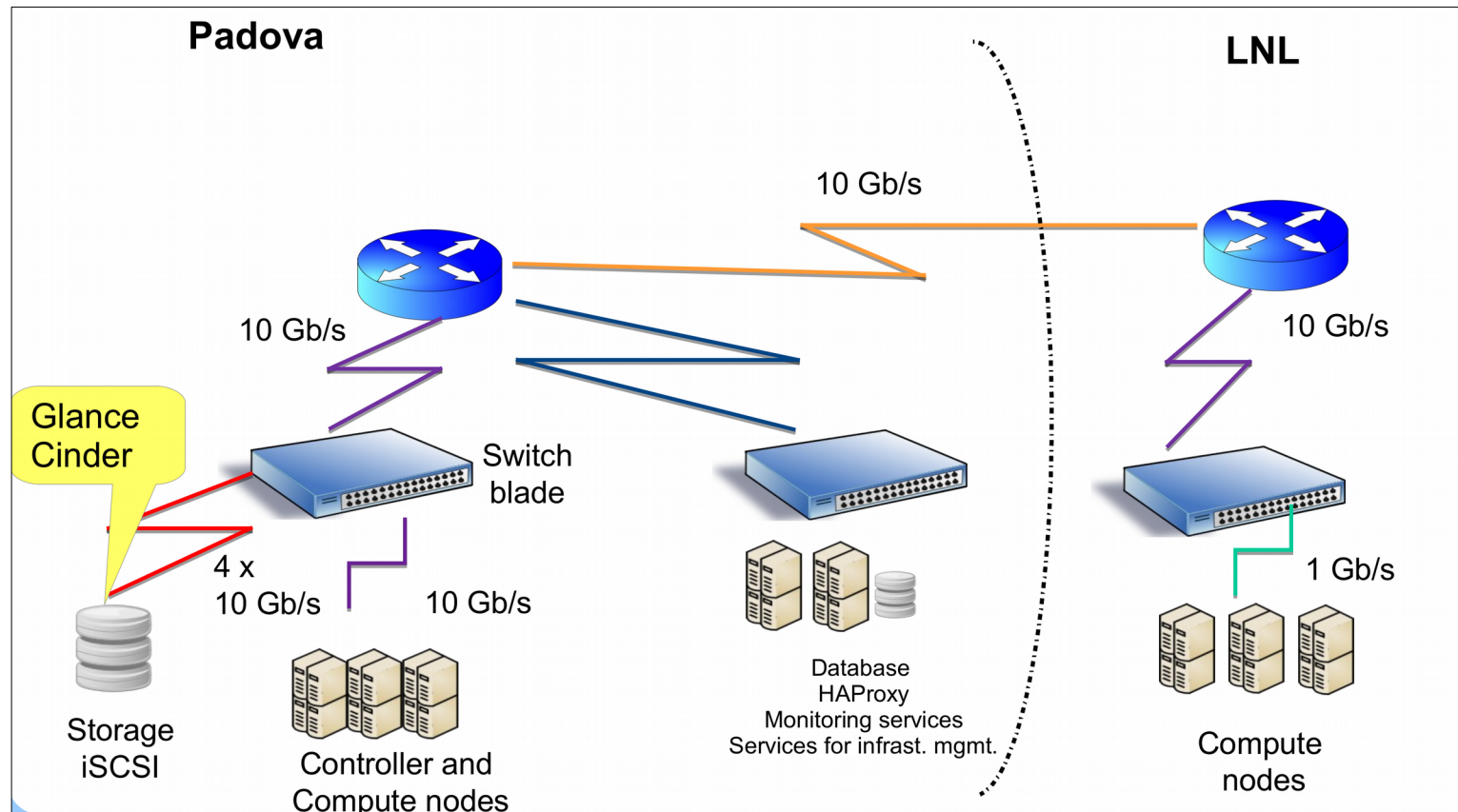
F. Chiarello
INFN Padova

On behalf of the “**Cloud Area Padovana**”
team



CAOS: OpenStack accounting management at Cloud Area Padovana

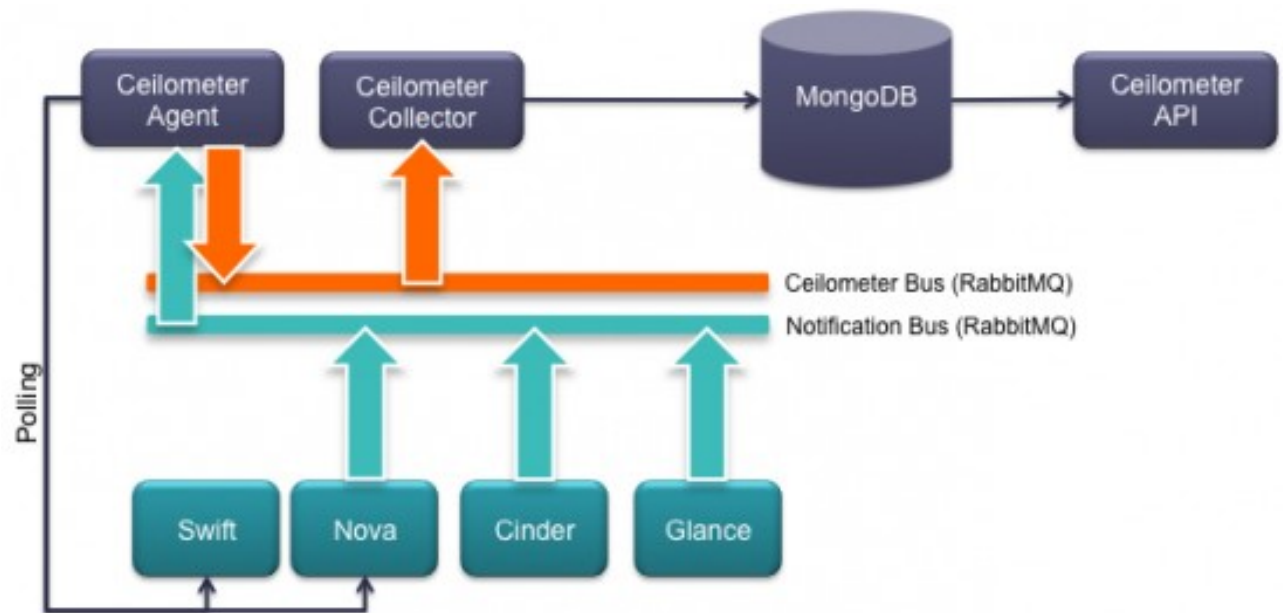
- OpenStack Mitaka (upgrade to Ocata being prepared)
- 28 compute nodes between PD and LNL (~1100 cores, ~4400 VCPUs)
- ~14 Mh cpu used over last year
- ~30 projects
- ~100 users



- Raw accounting information collected by the OpenStack telemetry service (ceilometer)
- An in-house developed tool (CAOS, Control Application for OpenStack) to:
 - collect and handle accounting information
 - resource usage by projects
 - usages on compute nodes
 - quotas
 - easy management and presentation of accounting data
 - actual usages
 - usages in a specified time range
 - custom metrics not provided by ceilometer

Ceilometer is the OpenStack component responsible to collect and manage accounting information:

- collects measurements within OpenStack
- monitoring and metering for
 - billing,
 - benchmarking,
 - scalability,
 - alarming
- API access
 - data query
 - aggregation
 - alarms



However, we found some limiting problems related to the way it handles information...

- inefficient storage of information
 - each measurement is accompanied by redundant metadata
- trade-off between history retention and storage
 - ~150-200 GB in 3 months, depending on cloud activity
 - might be solved in the future using gnocchi db

```
{
  "id": "577e73252360a744a97f6c8f",
  "counter_name": "cpu",
  "counter_volume": "2691000000",
  "timestamp": "2016-07-07T15:20:05Z",
  "project_id": "b38a0dab349e42bdbb469274b20a91b4",
  "user_id": "521300c330f45e3a8e9ec945970cc230",
  "resource_id": "87b1c046-9393-4974-814a-d2471a028714",
  "recorded_at": "2016-07-07T15:20:05.654Z",
  "message_signature": "f0ce5b42dc52b8f2360cce99b2dbaf69bb7d4386879696a2c93b42997f1c3",
  "message_id": "488a05be-4456-11e6-b25f-44a84225d06f",
  "source": "openstack",
  "counter_unit": "ms",
  "resource_metadata": {
    "status": "active",
    "cpu_number": 1,
    "ephemeral_gb": 0,
    "display_name": "segatta-test1",
    "name": "instance-000749a3",
    "disk_gb": 20,
    "kernel_id": null,
    "image": {
      "id": "d67f6702-9c70-4972-9bc5-2adea2e2ca85",
      "links": [
        {
          "href": "http://192.168.60.40:8774/1b2caeedb3e2497b935723dc6e142ec9/images/d67f6702-9c70-4972-9bc5-2adea2e2ca85",
          "rel": "bookmark"
        }
      ],
      "name": "CentOS7",
      "ramdisk_id": null,
      "vcpus": 1,
      "memory_mb": 2048,
      "instance_type": "12",
      "host": "fcaf16d02b798416d8f2a0c382589dc01c8d395dcd9af3bcc04bc894",
      "root_gb": 20,
      "image_ref": "d67f6702-9c70-4972-9bc5-2adea2e2ca85",
      "flavor": {
        "name": "c1dareapd.small",
        "links": [
          {
            "href": "http://192.168.60.40:8774/1b2caeedb3e2497b935723dc6e142ec9/flavors/12",
            "rel": "bookmark"
          }
        ],
        "ram": 2048,
        "ephemeral": 0,
        "vcpus": 1,
        "disk": 20,
        "id": "12"
      }
    },
    "OS-EXT-AZ:availability_zone": "nova",
    "image_ref_url": "http://192.168.60.40:8774/1b2caeedb3e2497b935723dc6e142ec9/images/d67f6702-9c70-4972-9bc5-2adea2e2ca85",
    "counter_type": "cumulative"
  }
}
```

What we need

```
"counter_name": "cpu",
"timestamp": "2016-07-07T15:20:05Z",
"counter_volume": "2691000000",
"project_id": "b38a0dab349e42bdbb469274b20a91b4"
```

- metrics refer to single resources (e.g. single VM)
- no data pre-aggregation
 - has to be done on the fly
 - too slow due to inefficiencies
- real case: cpu time of a project in a particular day:

```
# time ceilometer sample-list -m cpu -q 'project=e9ddc1e36c1348fa812c56f1922371b5;  
timestamp>=2016-10-06;timestamp<=2016-10-07'
```

```
...  
real    2m12.483s
```

(OpenStack Kilo)

```
# time ceilometer sample-list -m cpu -q 'project=e9ddc1e36c1348fa812c56f1922371b5;  
timestamp>=2017-04-15;timestamp<=2017-04-16'
```

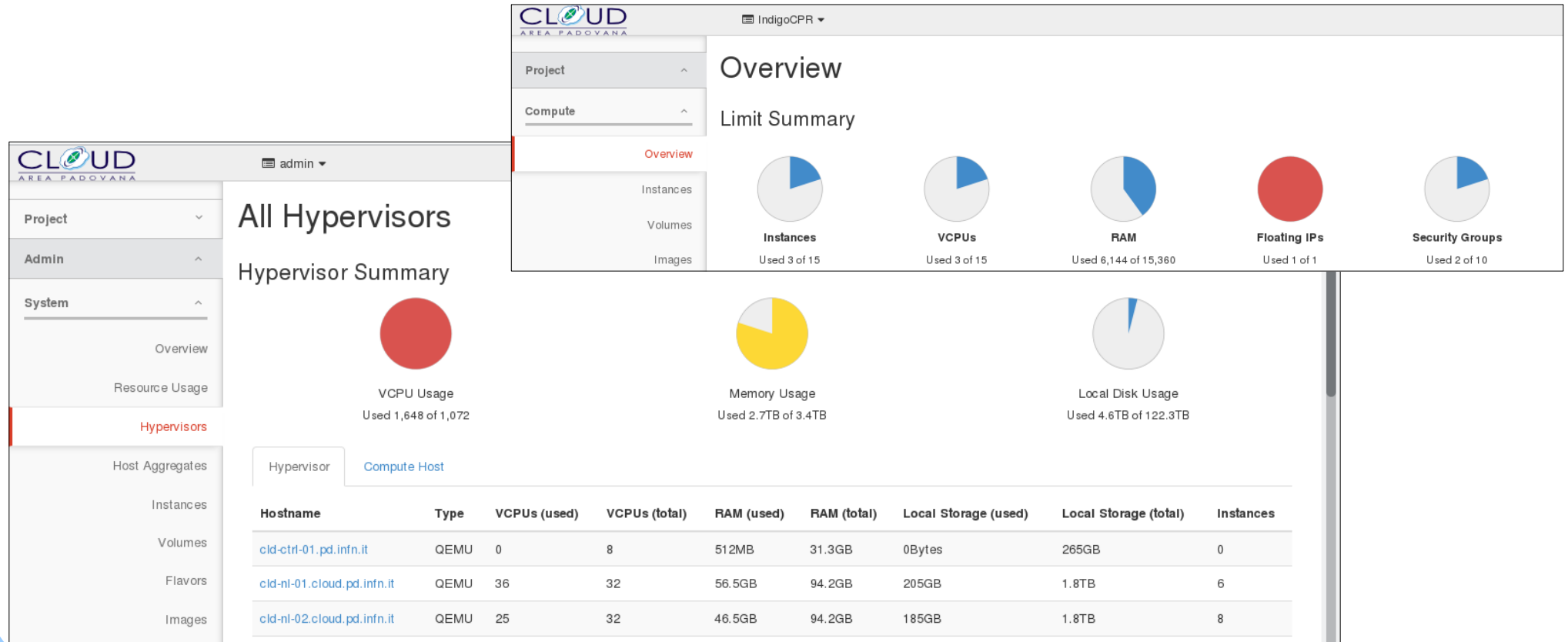
```
...  
real    0m57.153s
```

```
# time ceilometer statistics -m cpu -q 'project=e9ddc1e36c1348fa812c56f1922371b5;  
timestamp>=2017-04-15;timestamp<=2017-04-16' -p 3600 -a avg
```

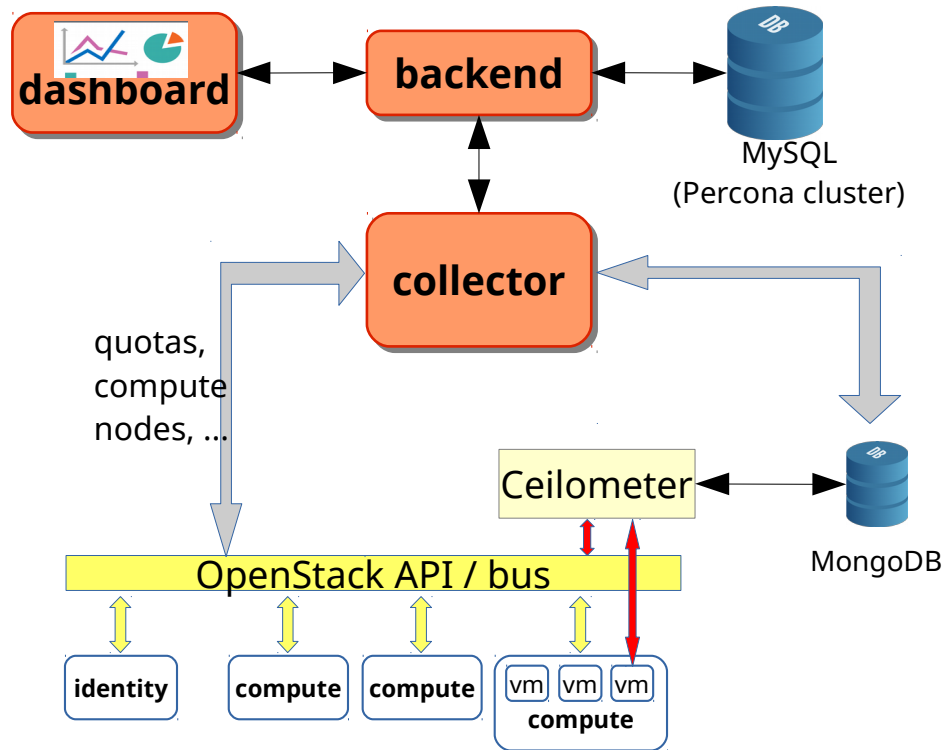
```
...  
real    2m49.561s
```

(OpenStack Mitaka)

- important metrics are not tracked by ceilometer
 - quotas, compute node usages, ...
 - available via OpenStack APIs
 - actual value only, no history preservation



- no support for operations between metrics
 - e.g. quota usage , CPU usage efficiency
 - “by hand”, by querying 2+ metrics (2+ slow)
 - no handling of time misalignment or missing data
- very limited support for custom metrics
 - require modifications to controller nodes
- inefficient and slow queries
- there exists some known bugs
 - e.g. cputime can become negative
 - post-processing required to correct the data



collector:

- gathers data at regular intervals
- workaround to some Ceilometer issues
 - directly query its backend (MongoDB)
 - query the OpenStack API (e.g. for quotas)
 - correction of wrong data (e.g. negative cpu time)
- data pre-aggregation at coarser granularity (hourly, daily)
- computations across different metrics (cpu efficiency)

backend:

- provides a time series framework for writing and reading metrics
- data aggregation and downsampling at given resolution/time range

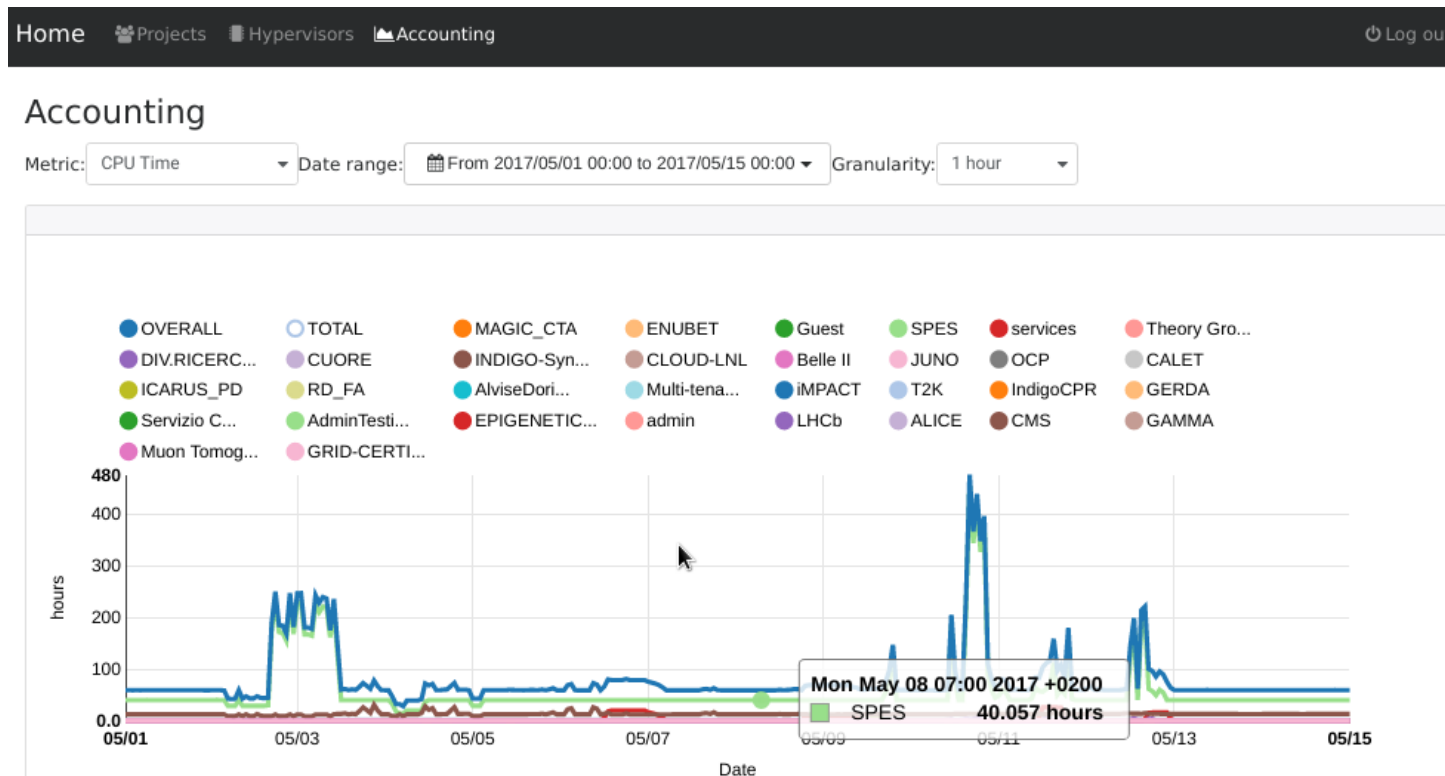
dashboard:

- frontend to easily get resource usage information
- setting of time ranges/resolution parameters
- display data in graphs and tabular forms

metrics aggregated per project and per compute node

- cpu time, wall clock time, cpu efficiency
- projects quotas and their usages (RAM, DISK, number of VMs, ...)
- current and past usages easy to determine
- fast aggregation over selected time range and granularity

- Comprehensive graphs of accounting data
 - on-the-fly generation
 - cover any of the tracked metrics
 - customizable time range / resolution
 - navigation through the usage history



- Accounting data can also be aggregated and reported in tabular form

Aggregates from 2017-05-01 00:00 CEST to 2017-05-15 00:00 CEST

Name ⚙	Wall Clock Time [hours] ▾	Percentage w.r.t overall ⚙
OVERALL	510.394,89	100,00%
SPES	242.294,55	47,47%
CMS	102.227,36	20,03%

LHCb

Belle II

AdminTesting

Muon Tomogr

EPIGENETICS

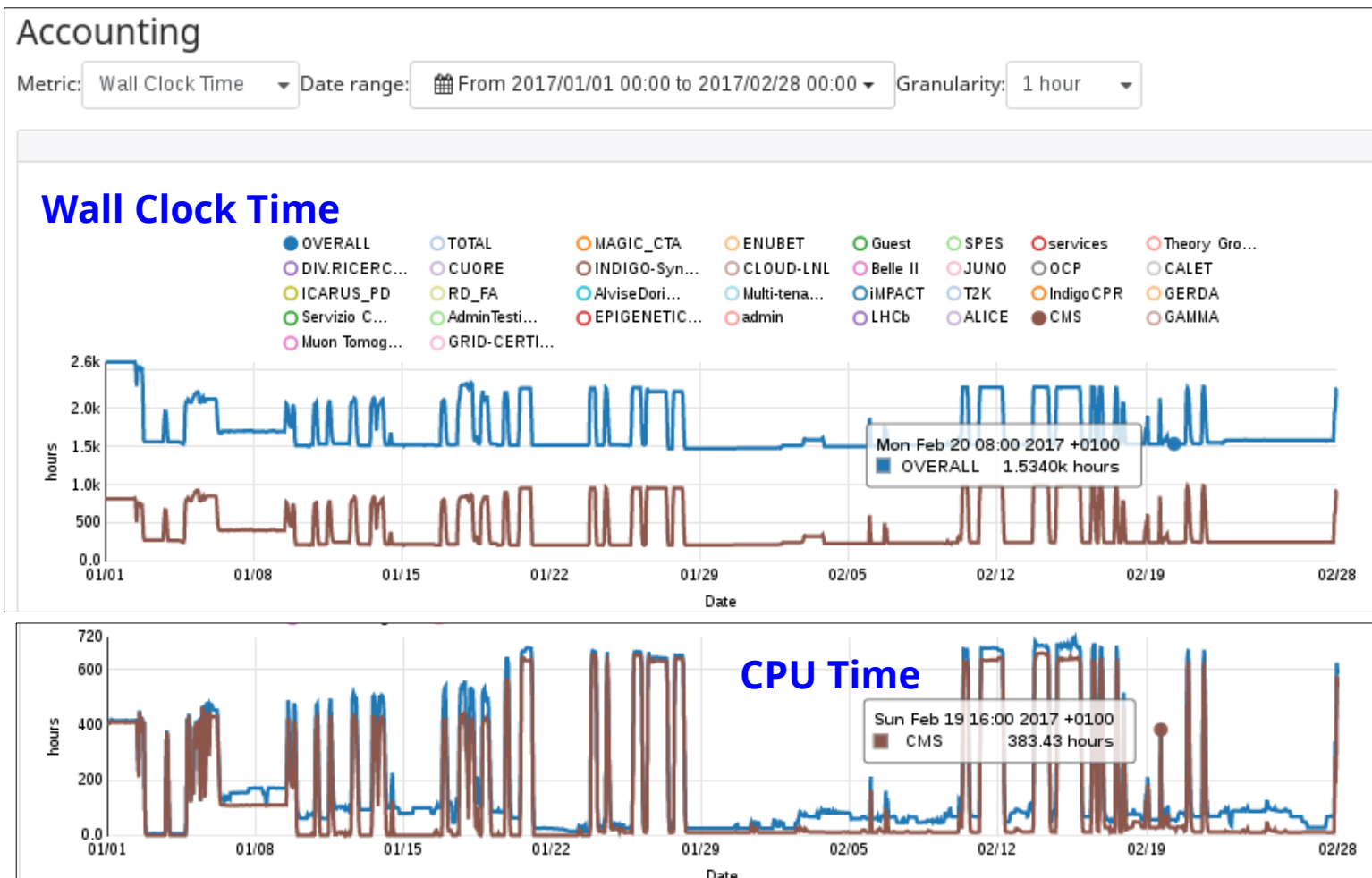
OCP

Aggregates from 2017-05-01 00:00 CEST to 2017-05-15 00:00 CEST

Name ⚙	CPU Time [hours] ▾	Percentage w.r.t overall ⚙
OVERALL	26.656,97	100,00%
SPES	19.300,04	72,40%
CMS	4.381,45	16,44%
EPIGENETICS	697,93	2,62%
Multi-tenant Services	659,82	2,48%

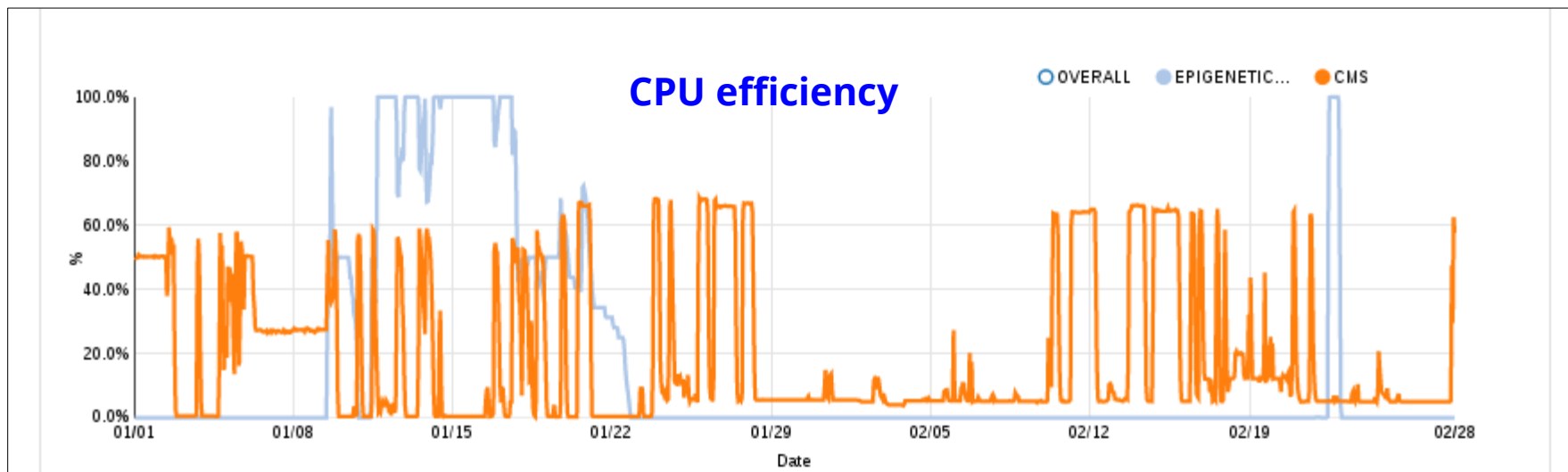
One use case: CMS elastic batch cluster on the Cloud (1/2)

The elastic batch system used by CMS dynamically create and destroy VM based on jobs load: resources are freed when not needed.



CPU usage efficiency is affected by overcommitment

- Elastic clusters share physical resources with the rest of the Cloud, but are dynamically allocated and expanded as needed.
- Other use cases can require reserved nodes or lower overcommitment for higher CPU throughput.



Resource usages by projects

Home	Projects	Hypervisors	Accounting	VCPUs	VMs	RAM	Log out
CLOUD-LNL	367f4c8918fc4d658d49be0b79ac9daf	3 minutes ago	Usage: 0,00 Quota: 15,00	Usage: 0 Quota: 15	Usage: 0,00 Quota: 15,36		
Belle II	36b1ddb5dab8404d7fc359ec95ecf5	3 minutes ago	Usage: 83,00 41.50% Quota: 200,00	Usage: 13 6.50% Quota: 200	Usage: 120,83 59.00% Quota: 204,80		
JUNO	37c6f66ffa0047898811f1c3b9d3e2cb	3 minutes ago	Usage: 7,42 12.37% Quota: 60,00	Usage: 7 7.00% Quota: 100	Usage: 15,21 24.75% Quota: 61,44		
OCP	3beba6dd3f2648378263bc04d9c205fa	3 minutes ago	Usage: 29,00 29.00% Quota: 100,00	Usage: 14 70.00% Quota: 20	Usage: 59,39 90.63% Quota: 65,54		
CALET	3f13d911e5a0448db1ad8363d0d264d5	3 minutes ago	Usage: 4,00 26.67% Quota: 15,00	Usage: 1 6.67% Quota: 15	Usage: 8,19 53.33% Quota: 15,36		
ICARUS_PD	4acc5c73693d4b8f909a5271f3b09a53	3 minutes ago	Usage: 7,38 14.76% Quota: 50,00	Usage: 2 4.0% Quota: 50	Usage: 15,11 29.51% Quota: 51,20		
AlviseDorigo	529769bcecf64277b7447a3c11088f2a	3 minutes ago	Usage: 22,00 27.50% Quota: 80,00	Usage: 10 50.00% Quota: 20	Usage: 45,06 90.11% Quota: 50,00		
Multi-tenant Services	535c148bff524f5ca7ac1cb4dc6ef08b	3 minutes ago	Usage: 9,00 45.00% Quota: 20,00	Usage: 5 25.00% Quota: 20	Usage: 18,43 92.16% Quota: 20,00		
T2K	73986e45bd624099bb4f4c5f03e56105	3 minutes ago	Usage: 0,00 Quota: 40,00	Usage: 0 Quota: 40	Usage: 0,00 Quota: 40,96		
GERDA	8c8f846533a24f2abd632c7f7a8d5fac	3 minutes ago	Usage: 2,00 13.33% Quota: 15,00	Usage: 1 6.67% Quota: 15	Usage: 4,10 26.67% Quota: 15,36		
Servizio Calcolo e Ret	b3280982c25146c8b7dccbde74ba1532	3 minutes ago	Usage: 0,00 Quota: 15,00	Usage: 0 Quota: 15	Usage: 0,00 Quota: 15,36		
AdminTesting	b38a0dab349e42bdbb469274b20a91b4	3 minutes ago	Usage: 61,00 30.50% Quota: 200,00	Usage: 23 46.00% Quota: 50	Usage: 92,16 45.00% Quota: 204,80		
EPIGENETICS	bdd04490175e4462a4c258b2cfed6ca0	3 minutes ago	Usage: 32,00 100.00% Quota: 32,00	Usage: 2 6.25% Quota: 32	Usage: 32,77 100.00% Quota: 32,77		
admin	beaeede3841b47efb6b665a1a667e5b1	3 minutes ago	Usage: 22,15 11.08% Quota: 200,00	Usage: 19 47.50% Quota: 40	Usage: 17,49 28.68% Quota: 60,96		
			Usage: 98,00	Usage: 8	Usage: 102,40		

Usages on compute nodes

Home Projects Hypervisors Accounting Log out				
Active VMs	Allocated VRAM (TB)		Allocated VCPUs	
	Virtual	On bare	Virtual	On bare
259	Virtual: 3,15 62,34% Virtualizable: 5,05	Virtual: 3,15 88,96% Total Bare: 3,54	Virtual: 1926 46,48% Virtualizable: 4144	Virtual: 1926 179,66% Total Bare: 1072

Hypervisors

Name	Last seen	State / Status	VMs	Allocated VRAM (MB)		Allocated VCPUs	
				Virtual	On bare	Virtual	On bare
cld-ctrl-01.pd.infn.it (IP: 192.168.60.40")	4 minutes ago	down / disabled	0	Virtual: 512 1,06% Virtualizable: 48088.5	Virtual: 512 1,60% Total Bare: 32059	Virtual: 0 0,00% Virtualizable: 32	Virtual: 0 0,00% Total Bare: 8
cld-nl-01.cloud.pd.infn.it (IP: 192.168.60.56")	4 minutes ago	up / enabled	10	Virtual: 78848 54,49%	Virtual: 78848 81,73%	Virtual: 59 46,09%	Virtual: 59 184,38%

Hypervisors

Name	Last Updated	State / Status	Running VMs	Workload	Allocated VCPUs	Allocated VRAMs	Allocated VCPUs (w.r.t bare)	Allocated VRAMs (w.r.t bare)	Load
cld-nl-07.cloud.pd.infn.it	11 minutes ago	up / enabled	14	0	60,94% 78 of 128	48,33% 93,18GB of 192,80GB	243,75% 78 of 32	72,50% 93,18GB of 128,53GB	5m 0,37% 10m 0,40% 15m 0,41%
cld-nl-06.cloud.pd.infn.it	11 minutes ago	up / enabled	13	0	51,56% 66 of 128	64,40% 93,18GB of 144,70GB	206,25% 66 of 32	96,59% 93,18GB of 96,47GB	5m 1,07% 10m 1,29% 15m 1,15%
cld-nl-08.cloud.pd.infn.it	11 minutes ago	up / enabled	12	0	64,84% 83 of 128	54,44% 104,96GB of 192,80GB	259,38% 83 of 32	81,66% 104,96GB of 128,53GB	5m 1,40% 10m 1,32% 15m 1,26%

- Features
 - leverages standard OpenStack services
 - resource tracking at different levels (e.g. per project or per compute node)
 - current and past resource usage easily determined, stored and presented
 - no modification required to the cloud installation
- Actually deployed at
 - Cloud Area Padova
 - INFN-PADOVA-STACK instance of the EGI Federated Cloud
 - Cloud of the University of Padova
- Usable on different OpenStack versions: Kilo, Liberty, Mitaka, (planned for Ocata)
 - ongoing development: <https://github.com/CloudPadovana/caos>
- Future plans
 - support for usages of nova-docker containers (no support by ceilometer)
 - provide access to project admins (actually reserved to the cloud admin)
 - evaluation/integration with gnocchi

@ PD

P. Andreetto, F. Chiarello, F. Costa, A. Crescente, A. Dorigo,
F. Fanzago, E. Konomi, M. Segatta, M. Sgaravatto,
S. Traldi, N. Tritto, M. Verlato, L. Zangrando

@ LNL

S. Fantinel

Questions?