



Cloud Monitoring

Speaker: Sara Vallero

The present work is partially funded under contract 20108T4XTM of Programmi di Ricerca Scientifica di Rilevante Interesse Nazionale (Italy).



Istituto Nazionale
di Fisica Nucleare



BARI: Vincenzo Spinoso

CATANIA: Emidio Giorgio

LNGS: Stefano Stalio

PADOVA: Sara Bertocco

Cloud multi regione INFN

PERUGIA: Gian Mario Bilei, Livio Fanò,
Matteo Pergolesi, Daniele Belfiore,
Mirko Mariotti, Lorianò Storchi

MarcheCloud

Open City Platform

TORINO: Andrea Guarise, Stefano Bagnasco,
Stefano Lusso, Sara Vallero

PRIN

Open City Platform

Monitoring

(informazione real-time)

Infrastruttura

Servizi /
applicazioni

Log analysis

Accounting

(storia)



Billing

Monitoring

(informazione real-time)

Infrastruttura

Nagios

Zabbix

ELK stack

ProZaC

Servizi / applicazioni

Log analysis
rsyslog

Accounting

(storia)

FAUST

Graphite

Grafana



Billing

ELK stack

Monitoring della Cloud multi regione INFN

- Nagios versus Zabbix
- Integrazione con OpenStack: ProZaC
- Centralizzazione log: rsyslog

- essenziale e critico in una **struttura distribuita** con **elementi ridondati**
- non facile e non standard
- **non** deve **pesare sulle performance** del sistema
- bisogna trovare il modo di **distribuire i log** in maniera **sicura**

- Monitoring dell'**hardware**
- Monitoring della **rete**
- Monitoring dei **servizi**:
 - Processi attivi
 - Processi realmente funzionanti
- Analisi dei **log**:
 - Per la funzionalità (ricerca errori)
 - Per la sicurezza (analisi log VM)
- Metriche per determinare il **livello di servizio** (anche per l'utente)
- Metriche per **valutare l'uso da parte dell'utente** o di gruppi di utenti
(per determinare eventuali politiche di billing)

- Infrastruttura di monitoring **Nagios** (server VM a LNGS)
 - Plugin standard (check-ping, check-disk, ecc.)
 - Plugin custom (check remoti via nrpe e check locali usando OpenStack-clients, es. cinder-client)

Nagios dashboard



Nagios®

- General
- Home
- Documentation
- Current Status
- Tactical Overview
- Map
- Hosts
- Services
- Host Groups
 - Summary
 - Grid
- Service Groups
 - Summary
 - Grid
- Problems
 - Services (Unhandled)
 - Hosts (Unhandled)
 - Network Outages
- Quick Search:
- Reports
 - Availability
 - Trends
 - Alerts
 - History
 - Summary
 - Histogram
 - Notifications
 - Event Log
- System
 - Comments
 - Downtime

	the object ring.						error[s] while checking hosts.
swift-proxy-node-pd	check if process swift-proxy-server is alive		OK	2014-12-12 11:40:08	7d 0h 48m 26s	1/4	PROCS OK: 7 processes with args 'swift-proxy-server'
	try to upload download and delete a file in a Swift container to check that it works correctly. Read input parameters from configuration file		OK	2014-12-12 11:41:39	2d 10h 11m 55s	1/4	Upload+download+delete of 595 KiB file in container sara
	uses swift-dispersion tools to report dispersion analysis and checks that all copies of objects are OK		OK	2014-12-12 11:37:29	0d 18h 26m 5s	1/4	100.00% objects copies found, 100.00% containers copies found
	uses swift-recon to query all clusters servers and ensure they all have the same copy of the object ring.		OK	2014-12-12 11:42:04	0d 18h 21m 30s	1/4	6/6 hosts matched, 0 error[s] while checking hosts. 0/6 hosts matched, 1 error[s] while checking hosts.
swift-storage-node-lngs-1	check if process rsync is alive		OK	2014-12-12 11:39:05	8d 18h 49m 29s	1/4	PROCS OK: 1 process with args 'rsync'
	check if swift process account-auditor is alive		OK	2014-12-12 11:41:11	8d 18h 57m 47s	1/4	PROCS OK: 1 process with args 'account-auditor'
	check if swift process account-reaper is alive		OK	2014-12-12 11:40:35	8d 18h 57m 59s	1/4	PROCS OK: 1 process with args 'account-reaper'
	check if swift process account-replicator is alive		OK	2014-12-12 11:40:36	8d 18h 57m 58s	1/4	PROCS OK: 1 process with args 'account-replicator'
	check if swift process account-server is alive		OK	2014-12-12 11:41:09	8d 20h 42m 25s	1/4	PROCS OK: 6 processes with args 'account-server'
	check if swift process container-auditor is alive		OK	2014-12-12 11:42:31	8d 18h 58m 22s	1/4	PROCS OK: 1 process with args 'container-auditor'
	check if swift process container-replicator is alive		OK	2014-12-12 11:38:28	8d 18h 55m 6s	1/4	PROCS OK: 1 process with args 'container-replicator'
	check if swift process container-server is alive		OK	2014-12-12 11:38:36	0d 10h 9m 58s	1/4	PROCS OK: 7 processes with args 'container-server'
	check if swift process container-sync is alive		OK	2014-12-12 11:42:14	8d 18h 57m 51s	1/4	PROCS OK: 1 process with args 'container-sync'
	check if swift process container-updater is alive		OK	2014-12-12 11:42:05	8d 18h 56m 29s	1/4	PROCS OK: 1 process with args 'container-updater'
	check if swift process object-auditor is alive		OK	2014-12-12 11:42:43	8d 20h 40m 51s	1/4	PROCS OK: 3 processes with args 'object-auditor'
	check if swift process object-replicator is alive		OK	2014-12-12 11:38:16	8d 18h 57m 36s	1/4	PROCS OK: 1 process with args 'object-replicator'
	check if swift process object-server is alive		OK	2014-12-12 11:42:35	8d 20h 40m 59s	1/4	PROCS OK: 8 processes with args 'object-server'
	check if swift process object-updater is alive		OK	2014-12-12 11:38:36	8d 18h 54m 58s	1/4	PROCS OK: 1 process with args 'object-updater'
	check used disk space on /srv/node/vdb1. Warning means free space less than 20pct. Critical means free space less than 10pct		OK	2014-12-12 11:41:20	8d 20h 40m 29s	1/4	DISK OK - free space: /srv/node/vdb1 3756334 MB (98% inode=99%):
	check used disk space on /srv/node/vdc1. Warning means free space less than 20pct. Critical means free space less than 10pct		OK	2014-12-12 11:38:39	8d 20h 39m 55s	1/4	DISK OK - free space: /srv/node/vdc1 3747449 MB (98% inode=99%):
	check used disk space on /srv/node/vdd1. Warning means free space less than 20pct. Critical means free space less than 10pct		OK	2014-12-12 11:39:30	8d 20h 39m 4s	1/4	DISK OK - free space: /srv/node/vdd1 3750384 MB (98% inode=99%):
	check used disk space on root partition. Warning means free space less than 20pct. Critical means free space less than 10pct		OK	2014-12-12 11:41:21	8d 20h 42m 13s	1/4	DISK OK - free space: / 694 MB (37% inode=48%):
swift-storage-node-lngs-2	check if process rsync is alive		OK	2014-12-12 11:42:26	8d 18h 41m 8s	1/4	PROCS OK: 1 process with args 'rsync'
	check if swift process account-auditor is alive		OK	2014-12-12 11:42:23	8d 18h 35m 34s	1/4	PROCS OK: 1 process with args 'account-auditor'

Nagios → Zabbix



Stiamo migrando da Nagios a Zabbix perché:

- Zabbix presenta già nativamente una serie di vantaggi
 - pilotabile via **API**
 - **registrazione** e configurazione **automatica** degli **host**
 - **mass update** dei check con un solo comando
 - interfaccia web molto espressiva
- molto lavoro è già stato fatto su Zabbix nell'ambito del progetto PRISMA ed è riutilizzabile
- a Catania esiste già un'applicazione per realizzare l'**integrazione Zabbix-Ceilometer**
- il lavoro già fatto su Nagios è riutilizzabile in Zabbix senza difficoltà

- Infrastruttura di monitoring **Zabbix** (server VM a Bari)
 - **creati** tutti i **gruppi**, sia per i siti, sia per gli host
 - **creati** tutti i **template**
 - la maggior parte dei **sensori sui servizi** (check del servizio sulla macchina) è stata configurata
 - la procedura per l'**auto-registrazione** dai siti è funzionante e disponibile
 - la migrazione dei check da Nagios è in corso

Zabbix dashboard

- Interfaccia molto ricca
- Configurabilità completa via web
- Comoda funzione di auto-registrazione degli host
- Critica la scelta dei gruppi per avere un display *pratico*
- Display degli allarmi mirato sui casi di errore

The screenshot displays the Zabbix web interface. At the top, there are navigation tabs for Monitoring, Inventory, Reports, Configuration, and Administration. Below this, a breadcrumb trail shows the current path: Home » Overview » Web » Latest data » Triggers » Events » Graphs » Screens » Maps » Discovery » IT services. The main content area is titled 'PERSONAL DASHBOARD' and is divided into several sections:

- Favourite graphs, screens, and maps:** Each section shows 'No [graphs/screens/maps] added.' with a 'Graphs/Screens/Maps »' button.
- Status of Zabbix:** A table showing system health metrics.

Parameter	Value	Details
Zabbix server is running	Yes	localhost:10051
Number of hosts (enabled/disabled/templates)	56	10 / 0 / 46
Number of items (enabled/disabled/not supported)	109	104 / 0 / 5
Number of triggers (enabled/disabled [problem/ok])	59	59 / 0 [3 / 56]
Number of users (online)	2	1
Required server performance, new values per second	1.40	-
- System status:** A table showing the number of hosts in different states across various host groups.

Host group	Disaster	High	Average	Warning	Information	Not classified
Bari	0	2	0	0	0	0
GlusterFS	0	0	0	0	0	0
Gran Sasso	0	0	0	0	0	0
HAProxy	0	1	0	0	0	0
Keystone	0	1	0	0	0	0
Neutron	0	0	0	0	0	0
Padova	0	0	0	0	0	0
Swift	0	1	0	0	0	0
Zabbix servers	0	0	0	0	0	0
- Host status:** A table showing the distribution of hosts by problem status.

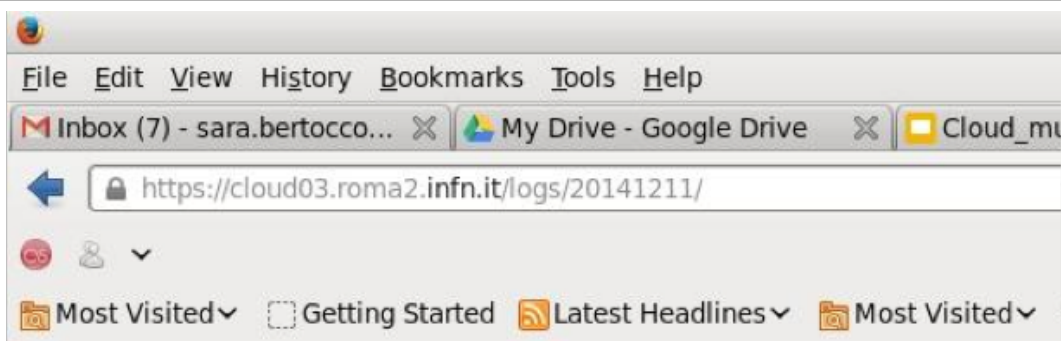
Host group	Without problems	With problems	Total
Bari	2	2	4
GlusterFS	1	0	1
Gran Sasso	1	0	1
HAProxy	0	1	1
Keystone	0	1	1
Neutron	1	0	1
Padova	3	0	3
Swift	4	1	5
Zabbix servers	1	0	1
- Last 20 issues:** A table showing recent alerts.

Host	Issue	Last change	Age	Info	Ack	Actions
swift2.be.infn.it	Cannot ping swift2.be.infn.it	2014-12-11 16:24:00	11m 46s		No	
haproxy.cloud.be.infn.it	Cannot ping haproxy.cloud.be.infn.it	2014-12-11 16:22:00	13m 46s		No	
keystone2.cloud.be.infn.it	Cannot ping keystone2.cloud.be.infn.it	2014-12-11 16:22:02	13m 50s		No	

- **Proxy** tra **Z**abbix e **C**eilometer :
 - Pull delle metriche delle VM da Ceilometer → push su Zabbix
 - Aggiornamento automatico dello stato dell'infrastruttura (nuove VM/tenants) attraverso threads in ascolto sui message broker
- <https://github.com/egiorgio/ProZaC>
 - Fork da <https://github.com/clmarques/ZabbixCeilometer-Proxy>
 - + supporto QPID / token renewal / Message Broker distinti fra Nova e Keystone / raffinamento metriche su Zabbix / logging / init script
 - Lavoro analogo in OCP (con analoghe modifiche)
 - avviati contatti per convergenza nei futuri sviluppi

- Centralizzazione dei log con **rsyslog** (server VM a Roma2)
 - Accesso remoto via Web al server centrale
 - Display dei log tipo *tail*
 - Log inviati **criptati**

Log display



Index of /logs/20141211/cloud02.roma2.

Name	Last Modified	Size	Type
Parent Directory/			Directory
nova-err.log	2014-Dec-11 11:12:36	1.31MB	LOG File
nova.log	2014-Dec-11 11:12:43	1.97MB	LOG File

Index of /logs/20141211/

Name	Last Modified	Size	Type
Parent Directory/			Directory
cloud01.roma2.infn.it/	2014-Dec-11 00:12:29	--	Directory
cloud02.roma2.infn.it/	2014-Dec-11 00:12:48	--	Directory
cloud03/	2014-Dec-11 09:12:39	--	Directory
proxy-node.pd.infn.it/	2014-Dec-11 11:12:17	--	Directory
storage-node-01.pd.infn.it/	2014-Dec-11 00:12:10	--	Directory
storage-node-02.pd.infn.it/	2014-Dec-11 00:12:01	--	Directory

Index of /logs/20141211/storage-node-01.pd.

Name	Last Modified	Size	Type
Parent Directory/			Directory
swift-err.log	2014-Dec-11 11:12:35	3.11MB	LOG File
swift.log	2014-Dec-11 11:12:42	22.80MB	LOG File

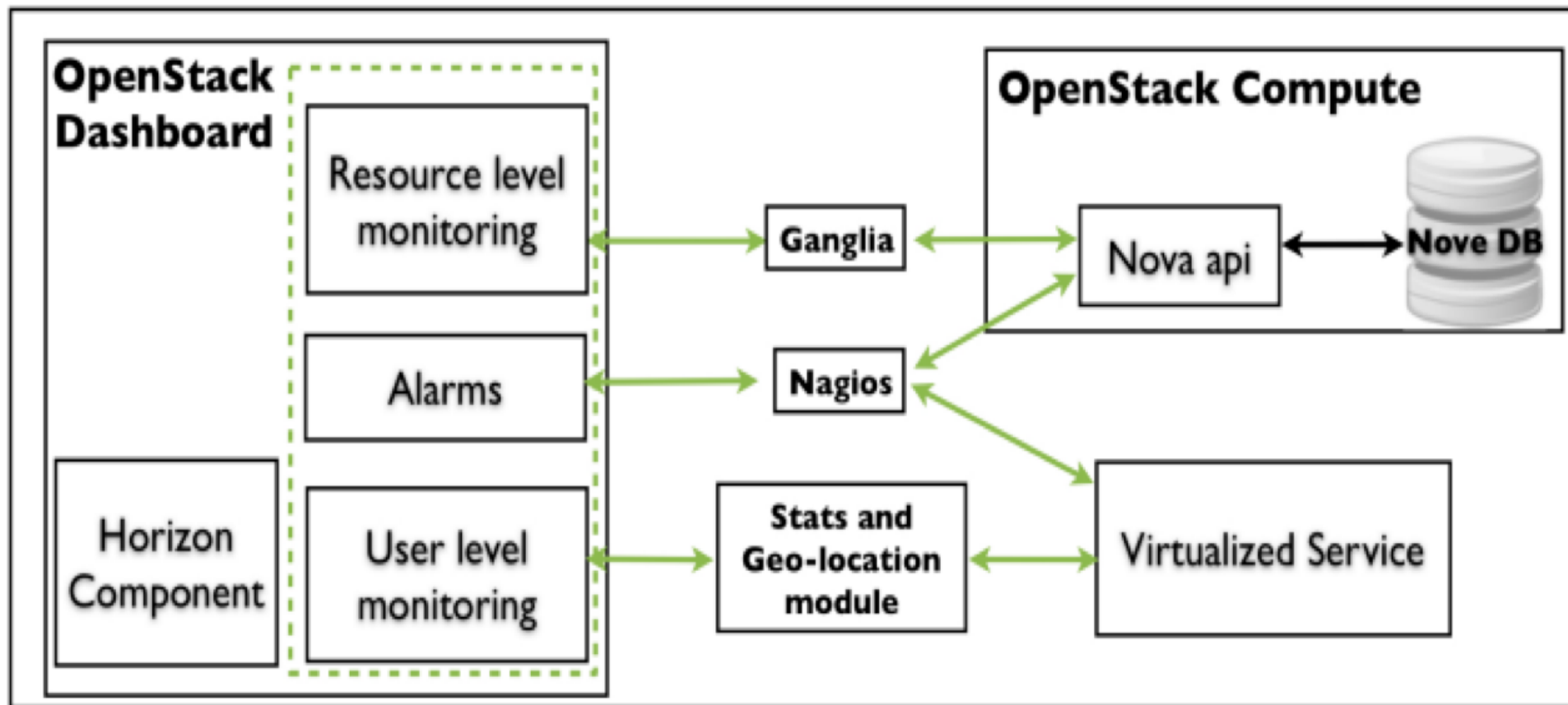
```
2014-12-11T11:05:20+01:00 storage-node-01 object-replicator: rsync: rcv_generator: mkdir "/vdb1/objects/197126/8f2" (in object) failed: No space left on device (28)
2014-12-11T11:05:20+01:00 storage-node-01 object-replicator: *** Skipping any contents from this failed directory ***
2014-12-11T11:05:20+01:00 storage-node-01 object-replicator: rsync error: some files/attrs were not transferred (see previous errors) (code 23) at main.c(1183) [sender=3.1.0]
2014-12-11T11:05:20+01:00 storage-node-01 object-replicator: Bad rsync return code: 23 <- ['rsync', '--recursive', '--whole-file', '--human-readable', '--xattrs', '--itemize-changes', '--
2014-12-11T11:05:20+01:00 storage-node-01 object-replicator: Error syncing with node: {'replication_port': 6000, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.17', 'region': 3, 'port': 6000}
2014-12-11T11:05:21+01:00 storage-node-01 container-replicator: ERROR reading HTTP response from {'replication_port': 6001, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.17', 'region': 3, '
2014-12-11T11:05:23+01:00 storage-node-01 object-replicator: message repeated 19 times: [ Error syncing with node: {'replication_port': 6000, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.1
2014-12-11T11:05:27+01:00 storage-node-01 object-replicator: rsync: rcv_generator: mkdir "/vdb1/objects/221917/9c1" (in object) failed: No space left on device (28)
2014-12-11T11:05:27+01:00 storage-node-01 object-replicator: *** Skipping any contents from this failed directory ***
2014-12-11T11:05:27+01:00 storage-node-01 object-replicator: rsync error: some files/attrs were not transferred (see previous errors) (code 23) at main.c(1183) [sender=3.1.0]
2014-12-11T11:05:27+01:00 storage-node-01 object-replicator: Bad rsync return code: 23 <- ['rsync', '--recursive', '--whole-file', '--human-readable', '--xattrs', '--itemize-changes', '--
2014-12-11T11:05:28+01:00 storage-node-01 object-replicator: Error syncing with node: {'replication_port': 6000, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.17', 'region': 3, 'port': 6000}
2014-12-11T11:05:30+01:00 storage-node-01 object-updater: ERROR with remote server 90.147.66.17:6001/vdb1: Connection refused
2014-12-11T11:05:32+01:00 storage-node-01 object-updater: message repeated 66 times: [ ERROR with remote server 90.147.66.17:6001/vdb1: Connection refused]
2014-12-11T11:05:32+01:00 storage-node-01 object-updater: Exception dumping recon cache: #012Traceback (most recent call last):#012 File "/usr/lib/python2.7/dist-packages/swift/common/u
2014-12-11T11:05:35+01:00 storage-node-01 object-replicator: message repeated 57 times: [ Error syncing with node: {'replication_port': 6000, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.1
2014-12-11T11:05:35+01:00 storage-node-01 object-replicator: Error syncing with node: {'replication_port': 6000, 'zone': 3, 'weight': 0.0, 'ip': '90.147.66.17', 'region': 3, 'port': 6000}
2014-12-11T11:05:35+01:00 storage-node-01 container-replicator: message repeated 646 times: [ ERROR reading HTTP response from {'replication_port': 6001, 'zone': 3, 'weight': 0.0, 'ip':
```


- Completare la migrazione a Zabbix
- Studiare un'architettura organica che tenga conto di tutti gli aspetti
- Ridondare la VM che ospita il server
(monitoring in HA a livello regionale)
- Usare un log server dedicato, anche questo possibilmente ridondato
- Completare la raccolta e scrittura dei check
- Valutare la possibilità di mettere i log su swift
- Migliorare l'analisi dei log via web (search)
- Basare il sistema di logging su CA non self-signed com'è ora (Terena?)

Monitoring: attività INFN Perugia

- MarcheCloud
- Open City Platform

- Progetto MarcheCloud infrastruttura di monitoring basata su **Nagios** e **Ganglia**. Infrastruttura tutt'ora in funzione presso Cloud Regione Marche
- **Integrazione** di dati di monitoring **nella dashboard di OpenStack**:
 - Resource Level Monitor e Alarms: in pratica accesso diretto ai portali di Nagios e Ganglia
 - User Level Monitor: aggrega in viste specifiche i dati di uso di risorse relativi ad utenti OpenStack

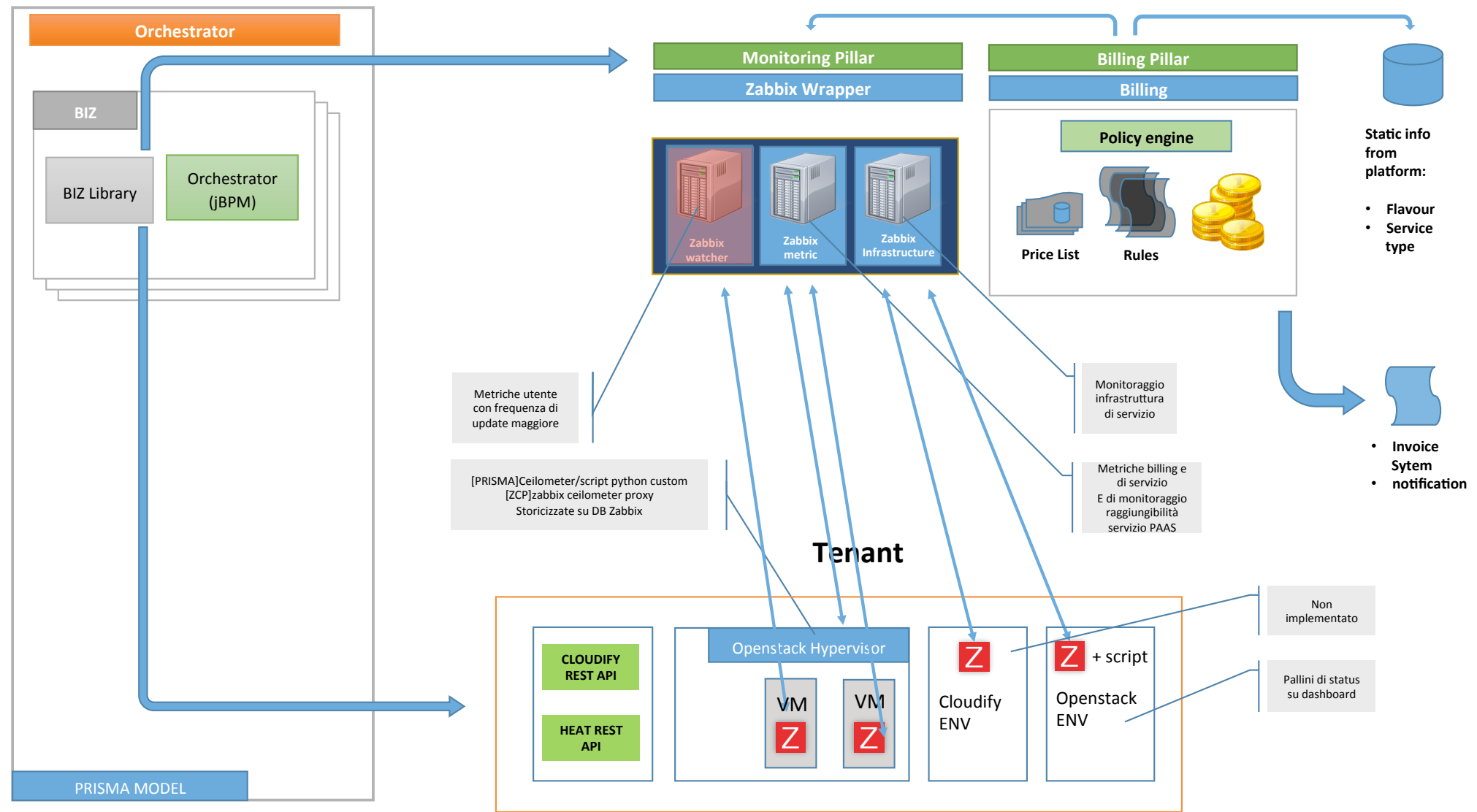


- Sono stati sviluppati diversi **plugins ad-hoc** per il monitoring dell'infrastruttura MCloud.
- Oltre a questo le **componenti server** di Ganglia e Nagios sono state **integrate nella dashboard di OpenStack**.

- **TestBed locale** basata su OpenStack Icehouse
- **Esplorazione Ceilometer** sue funzionalità e sue API
(sviluppo prima applicazione di monitoring di test basata su Ceilometer)
- Test *dashboard* Ceilometer (funzionalità decisamente limitate)
- Superati in parte alcuni problemi prestazionali di Ceilometer
(approfondimenti ulteriori necessari ed in corso vedi infrastruttura CERN)

- **Esplorazione di Zabbix**, verso definizione di architettura mutuata da PRISMA basata su Zabbix.
- **Integrazione Ceilometer con Zabbix** mediante fork locale di ZCP (Zabbix –Ceilometer Proxy)
- Verso Integrazione fork locale con ProZaC con INFN CT
- **Porting su testbed locale** architettura in *discussione* monitoring di OCP

Verso monitoring OCP e Monitoring-as-a-Service

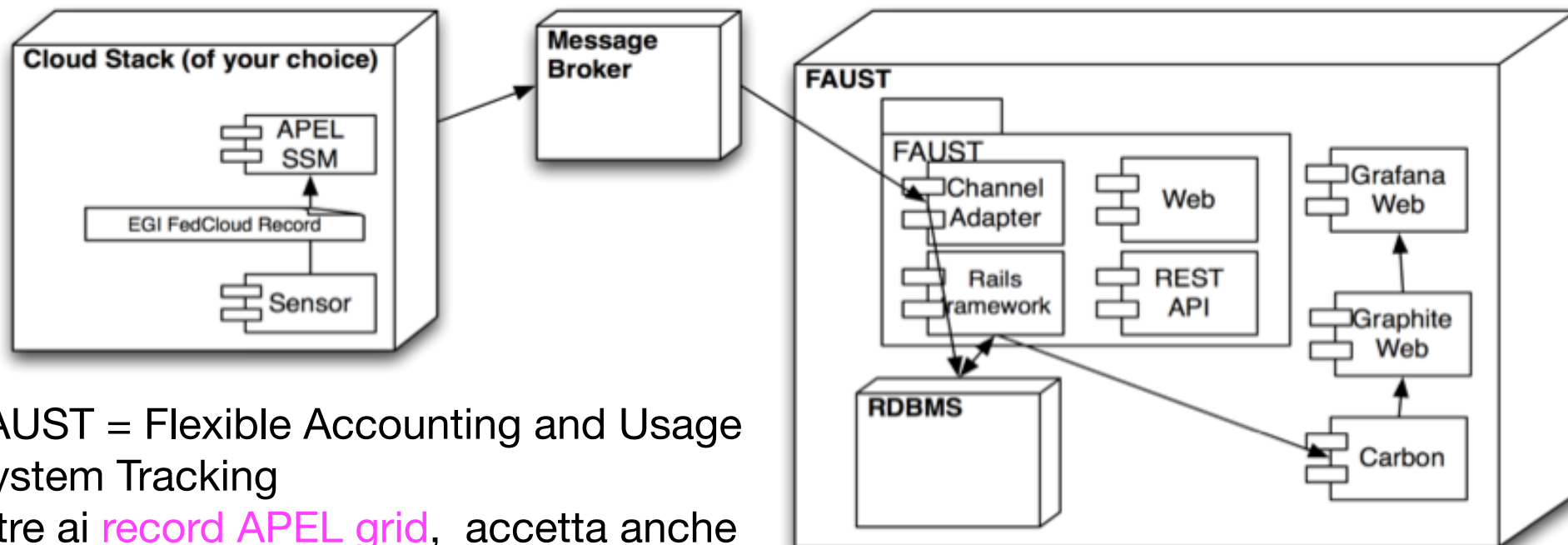


Vedi la presentazione di Giacinto su PRISMA.

Accounting: attività INFN Torino

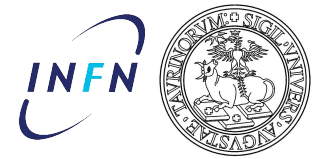
- FAUST
- ElasticSearch - Kibana - Logstash (ELK) stack

FAUST e l'accounting Cloud

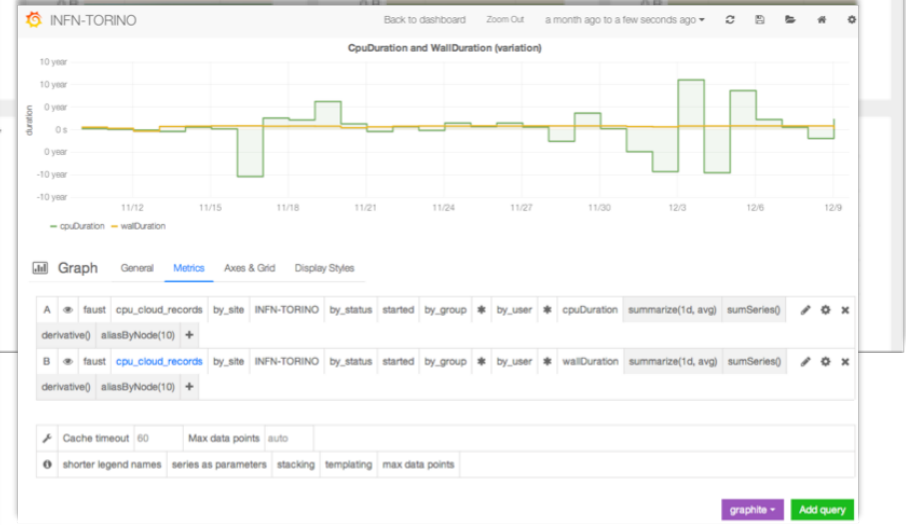
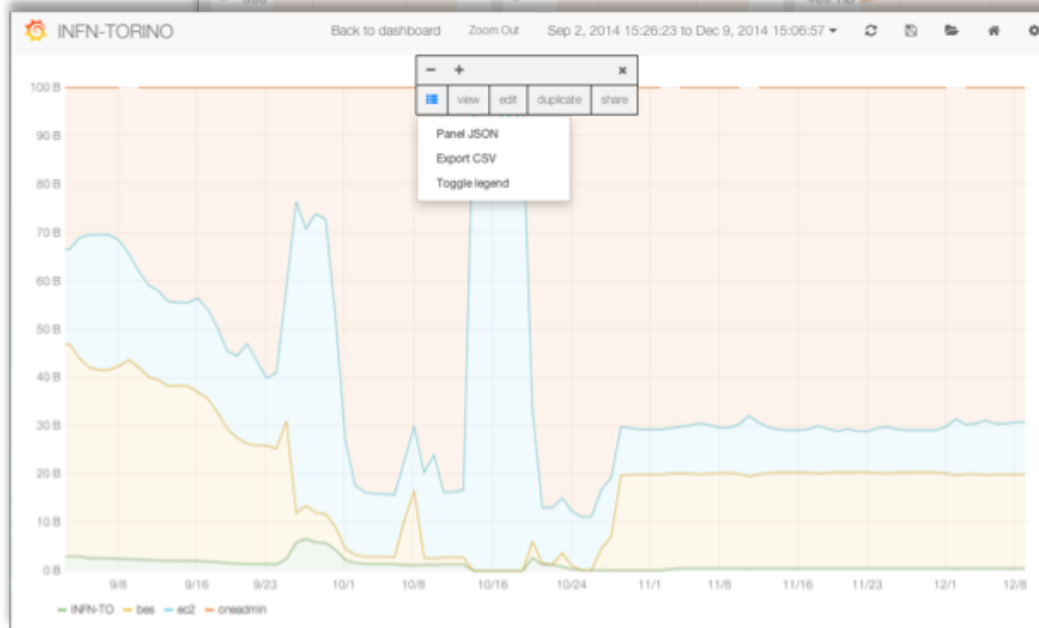
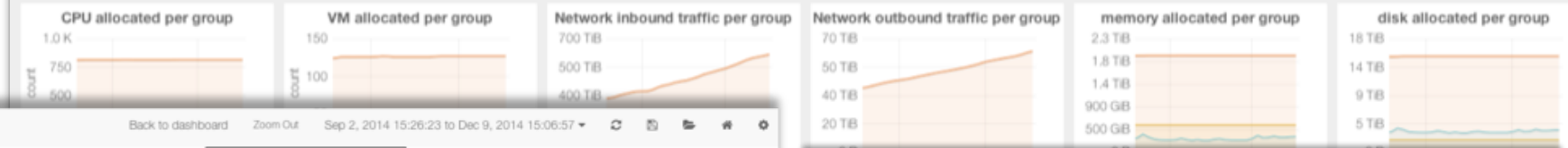
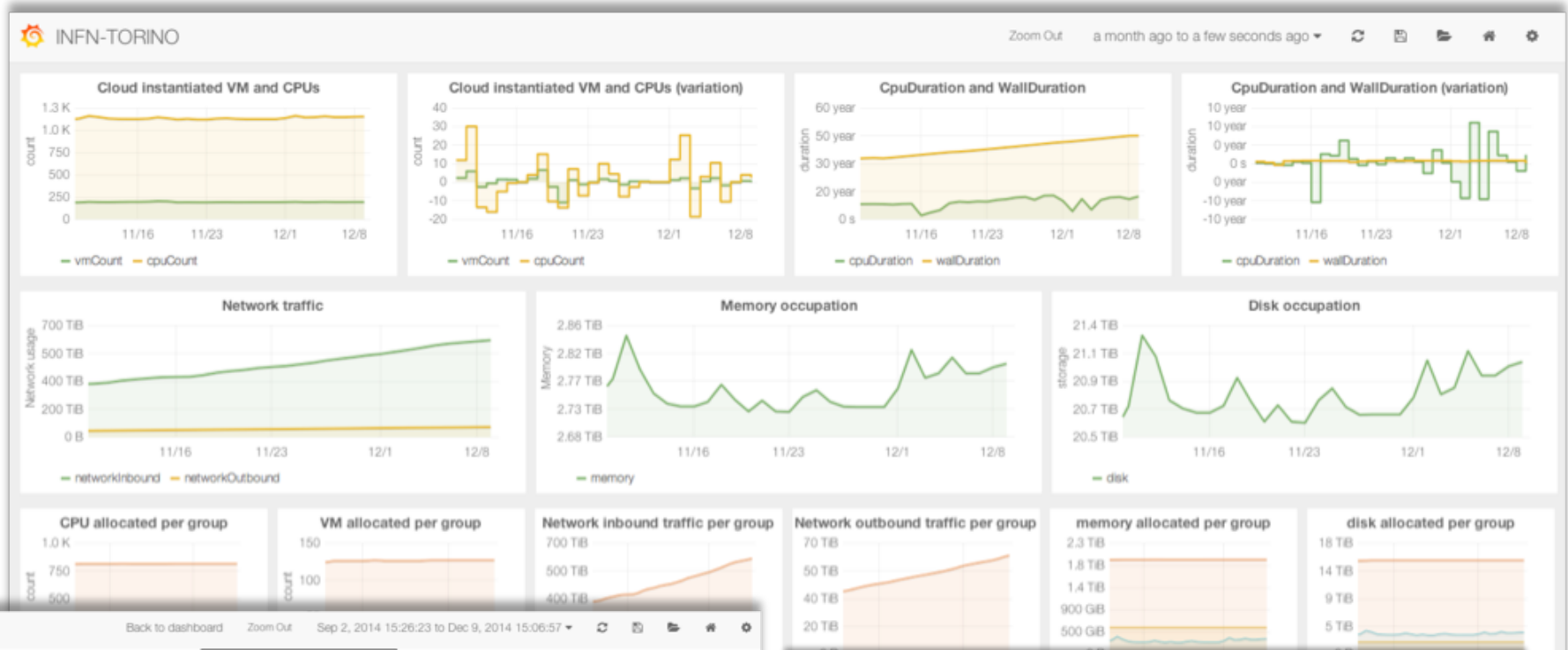


- FAUST = Flexible Accounting and Usage System Tracking
- oltre ai record APEL grid, accetta anche record cloud (IaaS) nel formato definito all'interno dell'EGI FedCloud Taskforce
- Il trasporto è implementato usando un sistema di messaggistica (Apache ActiveMQ) ed il protocollo STOMP
- La definizione dello Usage Record è presente all'URL:
<https://wiki.egi.eu/wiki/Fedcloud-tf:WorkGroups:Scenario4>
- È possibile definire dashboard veloci ed altamente personalizzabili grazie all'adozione di Graphite e Grafana

FAUST - cloud dashboard



INFN Torino



Elasticsearch:

- motore di ricerca ed analisi (Apache Lucene)
- interroga documenti in formato JSON, tutti i campi possono essere indicizzati ed usati in una singola query
- ricerca full-text tra dati non strutturati
- accetta ogni API di tipo RESTful che usi JSON su http
- scalabile orizzontalmente

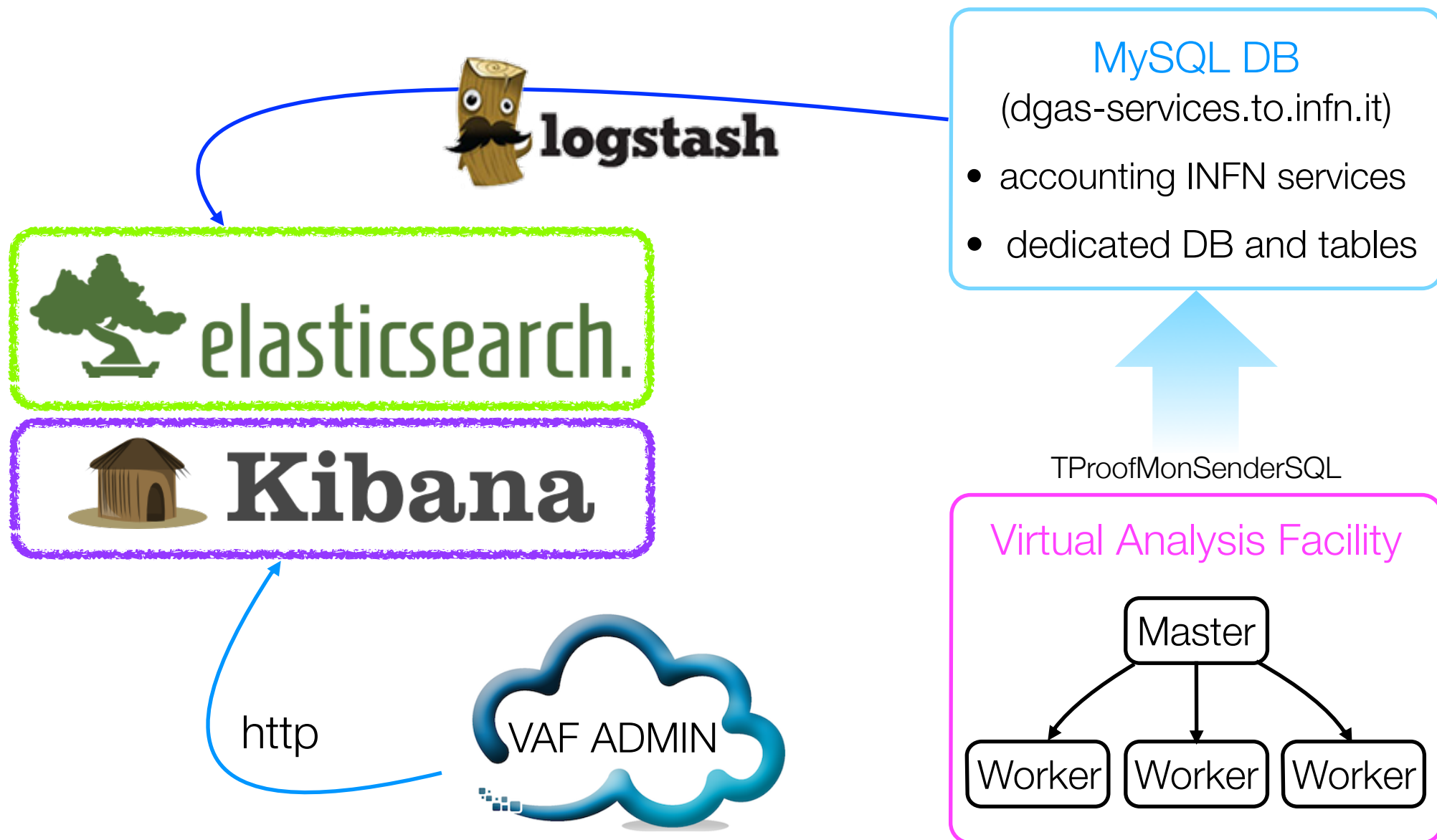
Kibana:

- GUI per visualizzare e fare ricerche sui dati ElasticSearch
- creazione di dashboard interattive con pochi mouse-click

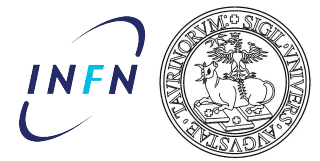
Logstash:

- raccoglie e reindirige log-files ad un sistema centrale
- sviluppato un plugin custom per prendere i dati da MySQL

ELK stack: example set-up

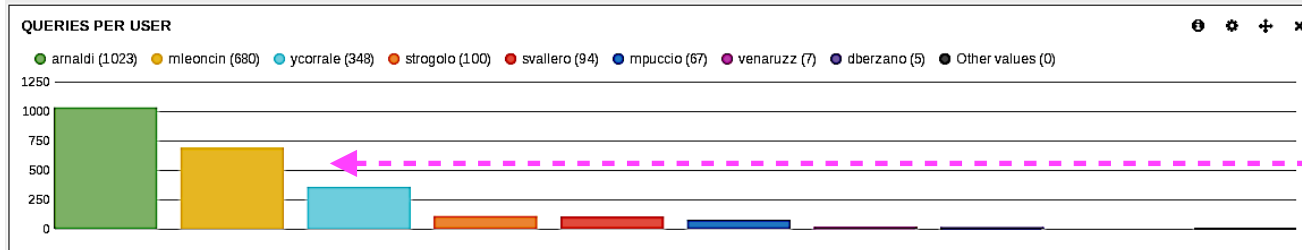


ELK stack: accounting e billing

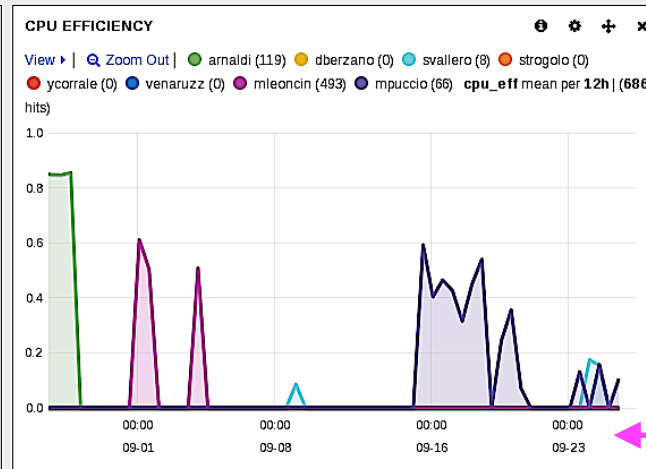
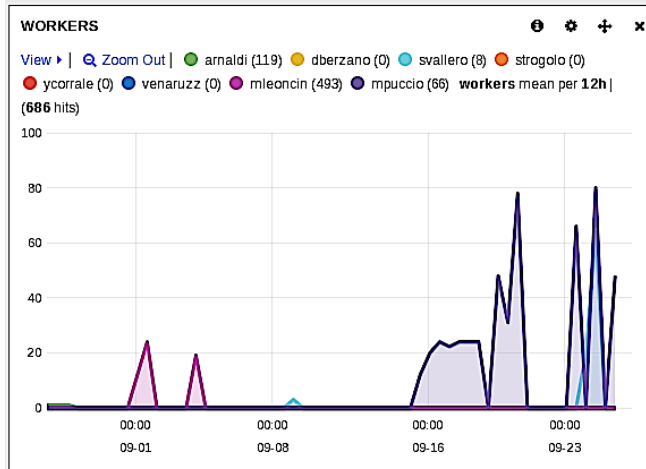


- il sito di Torino ospita una Cloud **multi-tenant** (grid Tier2, analysis facility interattiva, batch farm minori) con parziale automazione dell'allocazione di risorse
- **fonti di informazione eterogenee**:
 - cloud controller (OpenNebula)
 - applicazioni (i.e. Root)
 - altri tool di monitoring (i.e. Zabbix)
- interfaccia di monitoring uniforme per **laaS e applicazioni**
- ma non consente di creare allarmi e triggerare azioni
- una **conoscenza accurata dell'utilizzo di risorse** (CPU, memoria, disco, networking) è il prerequisito per un servizio di **billing** e per delle **quote intelligenti**
→ work in progress... (<https://github.com/svallero/cloud-accounting>)
- verso il **Monitoring-as-a-Service**:
 - l'utente si occupa di sviluppare metriche/sensori e popolare il database (sulla base di un esempio standard customizzabile)
 - semplice creazione di dashboard dinamiche con Kibana

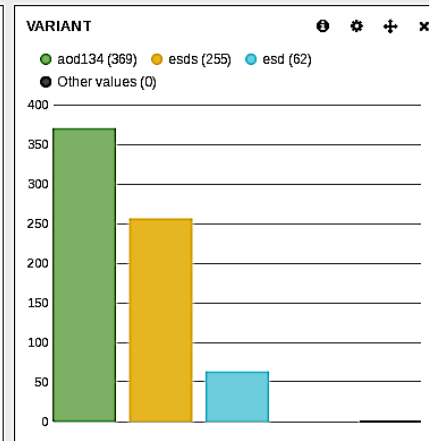
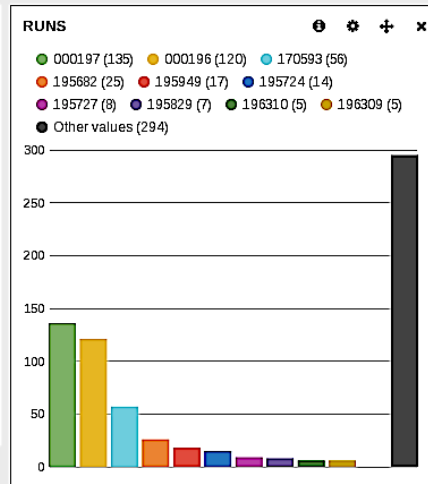
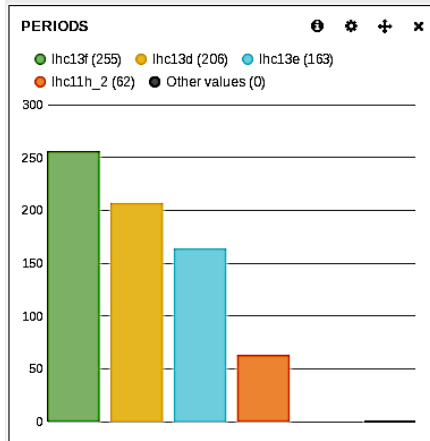
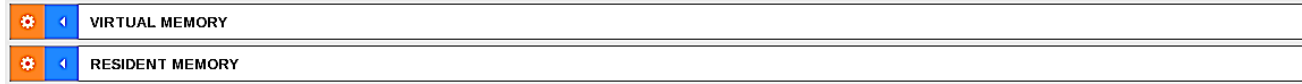
ELK stack: example dashboard



click per selezionare solo le entries relative ad un utente



range temporale dinamico



risultato analisi ES di una stringa complessa

- diverse attività di monitoring/accounting nei vari progetti:
 - parzialmente coordinate
 - utilizzo di **strumenti solidi e standard** (Nagios, Zabbix)
- maturata una buona **esperienza nel monitoring IaaS**, anche in un ambiente *difficile* quale un'infrastruttura distribuita
- lavori preliminari nel monitoring/accounting a livello di applicazione
 - **indagine di strumenti emergenti**, presi in prestito dal mondo dei Big Data (ELK stack)

... un modello di Monitoring-as-a-Service

- uno degli obiettivi espliciti del progetto OCP
- necessità di un **piano d'azione comune per il monitoraggio delle applicazioni**:
 - evitare proliferare di sviluppi verticali
 - evitare duplicati
- **core-framework** comune per la **componente infrastrutturale**
- i proprietari delle **applicazioni sviluppano plugin specifici**, che siano compliant con il core-framework
- due esempi di possibile core-framework:
 - Zabbix (OCP)
 - ElasticSearch (alta flessibilità, queries su dati non strutturati i.e. logfiles, VM templates...)