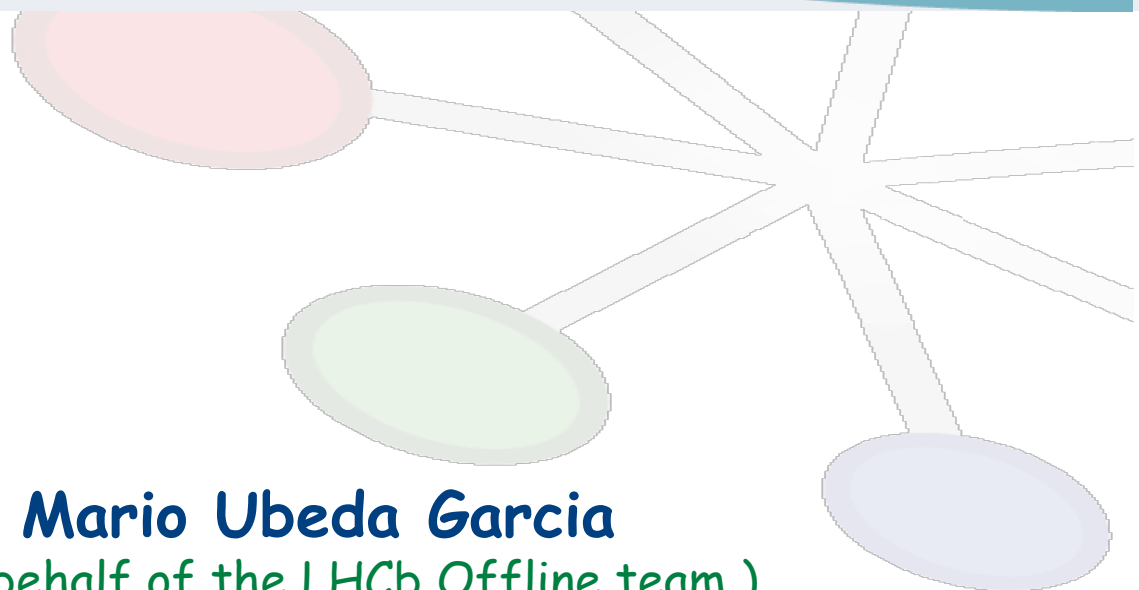
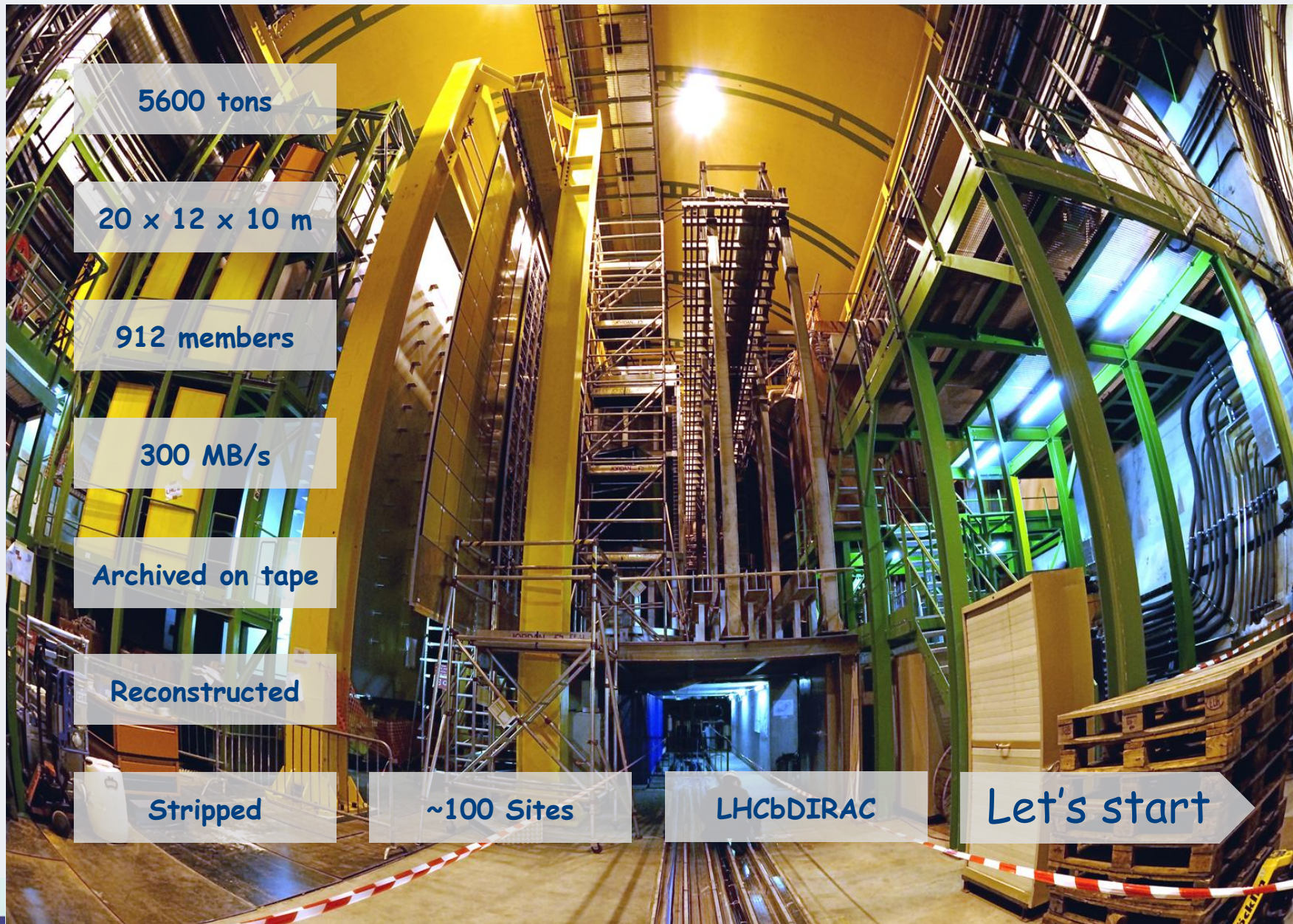




Integration of Cloud resources in the LHCb Distributed Computing



Mario Ubeda Garcia
(On behalf of the LHCb Offline team)



5600 tons

20 x 12 x 10 m

912 members

300 MB/s

Archived on tape

Reconstructed

Stripped

~100 Sites

LHCbDIRAC

Let's start



Problem
description

Use cases

Possible solutions

Architecture

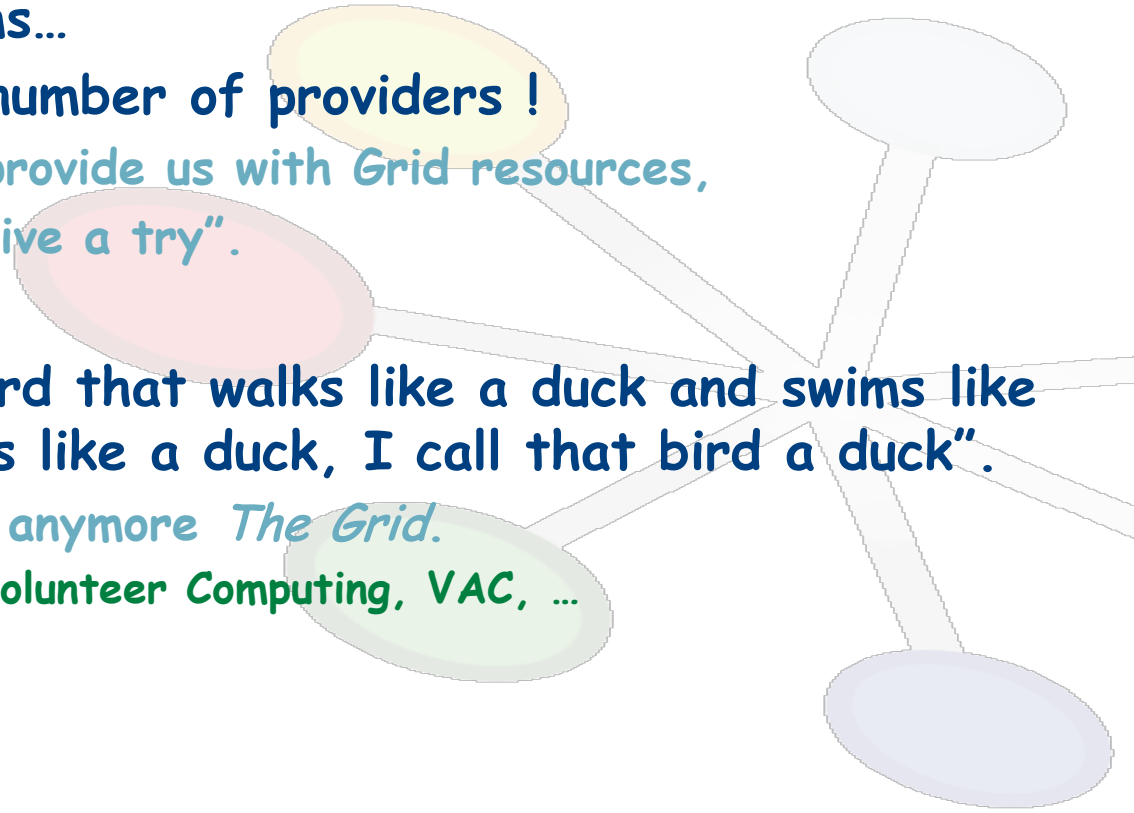
Related work

Next steps





- ❖ There is a growing concern regarding the usage of new types of platforms...
- ❖ ... and a growing number of providers !
 - Some of them provide us with *Grid* resources, and asked to “give a try”.
- ❖ “When I see a bird that walks like a duck and swims like a duck and quacks like a duck, I call that bird a duck”.
 - The *Grid* is not anymore *The Grid*.
 - Grid, Cloud, Volunteer Computing, VAC, ...



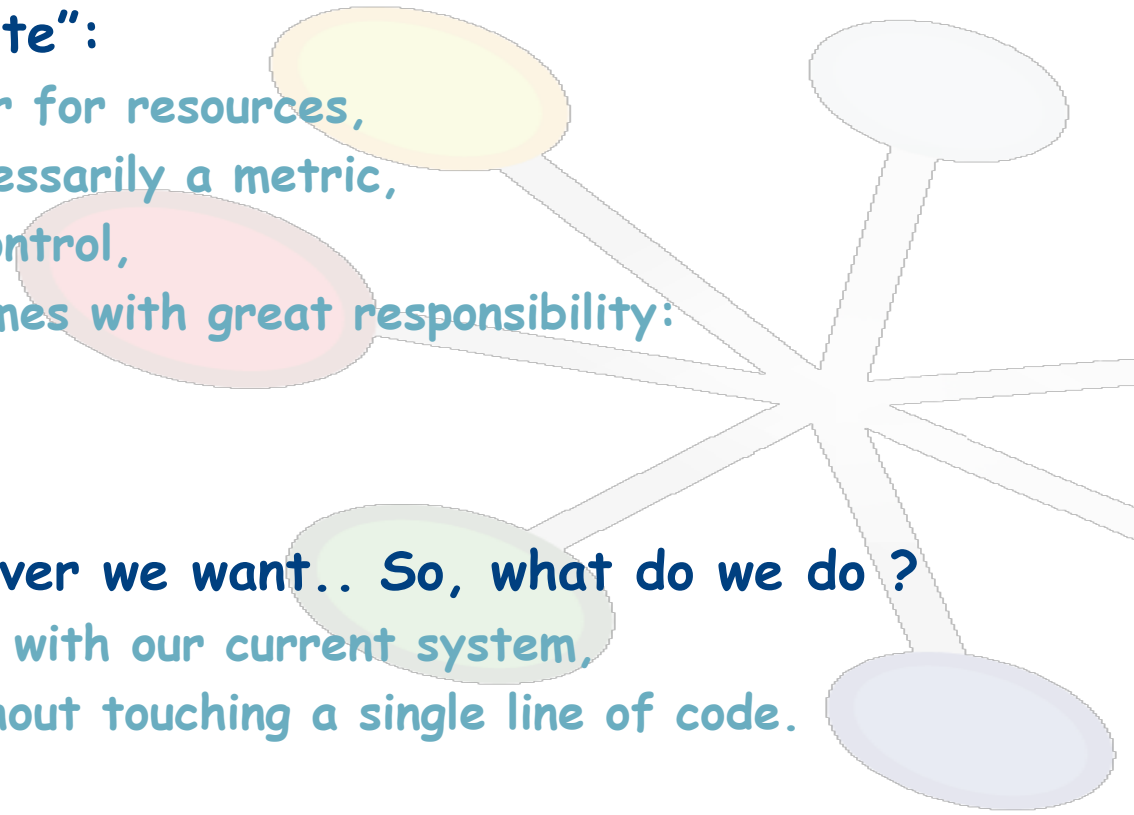


❖ Idea of “Cloud Site”:

- virtual container for resources,
- locality not necessarily a metric,
- VO gains full control,
- great power comes with great responsibility:
 - reliability,
 - monitoring,
 - security.

❖ We can do whatever we want.. So, what do we do ?

- Integrate them with our current system,
- if possible, without touching a single line of code.





Besides the scope of this talk (UC-0), we have identified three major use cases for the Cloud resources (UC-{1..3})

- ❖ UC-0: Cloud Sites.
- ❖ UC-1: Playground (prototypes, future developments...).
- ❖ UC-2: Continuous Integration testing.
- ❖ UC-3: LHCbDIRAC Server Infrastructure.



❖ No big numbers. In contrast, very solid infrastructure.

❖ Running on *Production* at:

➤ CERN (OpenStack):

- CLOUD.CERN.ch,
- CLOUD.CERNMP.ch.

➤ PIC (OpenNebula):

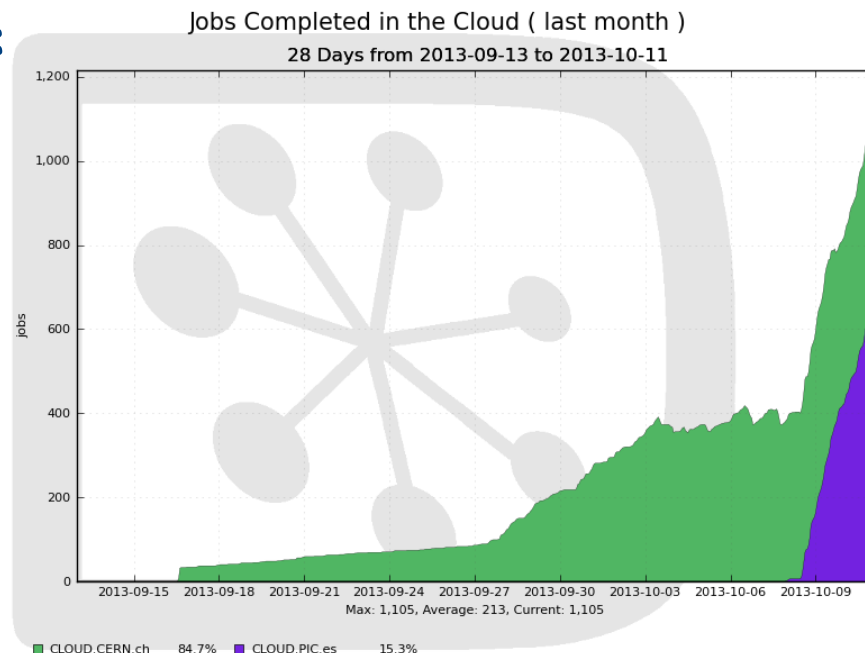
- CLOUD.PIC.es.

➤ RAL (StratusLab):

- VERY SOON !

❖ Jobs we run:

- MC,
- Data processing.



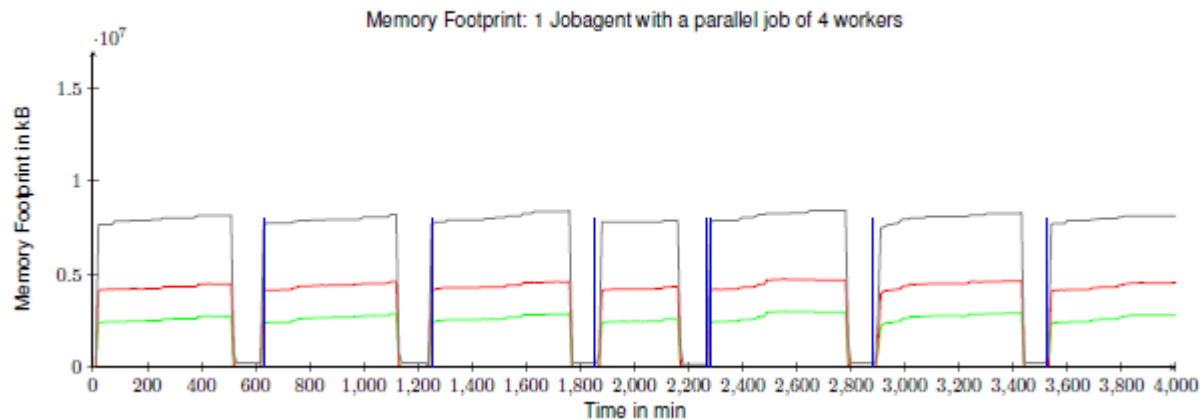
Generated on 2013-10-11 06:05:45 UTC

Cloud Site	Count	RAM *	Disk *	VCPUs
CLOUD.CERN.ch	26	4	40	2
CLOUD.CERNMP.ch	4	8	80	4
CLOUD.PIC.es	75	2	10	1
ΣTotal	105	286	2.11K	143

* (GB)



- ❖ Sometimes, we do not find what we need on *The Grid*:
 - binaries (e.g. KSM requires to be supported),
 - kernel (e.g. x32-ABI requires recompiling the kernel).
- ❖ Proud to show N. Rauschmayr's work on memory footprint reduction with GaudiMP:
 - Preparing the Gaudi-Framework and the DIRAC-WMS for Multicore Job Submission (CHEP'13; Id50).
- ❖ Soon on your favorite multicore Cloud Sites.





UC-2: Continuous Integration testing

Integration .. Cloud :: LHCb

- ❖ QA, Unit, Regression, Integration and System tests.
- ❖ Extremely handy validating SW on different OS/SW.
- ❖ At the moment:
 - Master LHCb Jenkins Server (S.Lohn et al):
 - Systematic profiling to monitor and specify the software refactoring process of the LHCb experiment (CHEP'13; Id339).
 - (virtual) slaves with multiple OS/SW configurations to test and certify our Grid middleware.

Type of test	Count	RAM *	Disk *	VCPUs
Regression, Integration & System	3	4	40	2
QA, Unit	1	8	80	4

Try it out, works like a dream !

* (GB)



UC-3: LHCbDIRAC Server Infrastructure

Integration .. Cloud .. LHCb

❖ At the moment:

➤ resources:

- 10 machines at CERN,
- 6 machines distributed among our T1s.

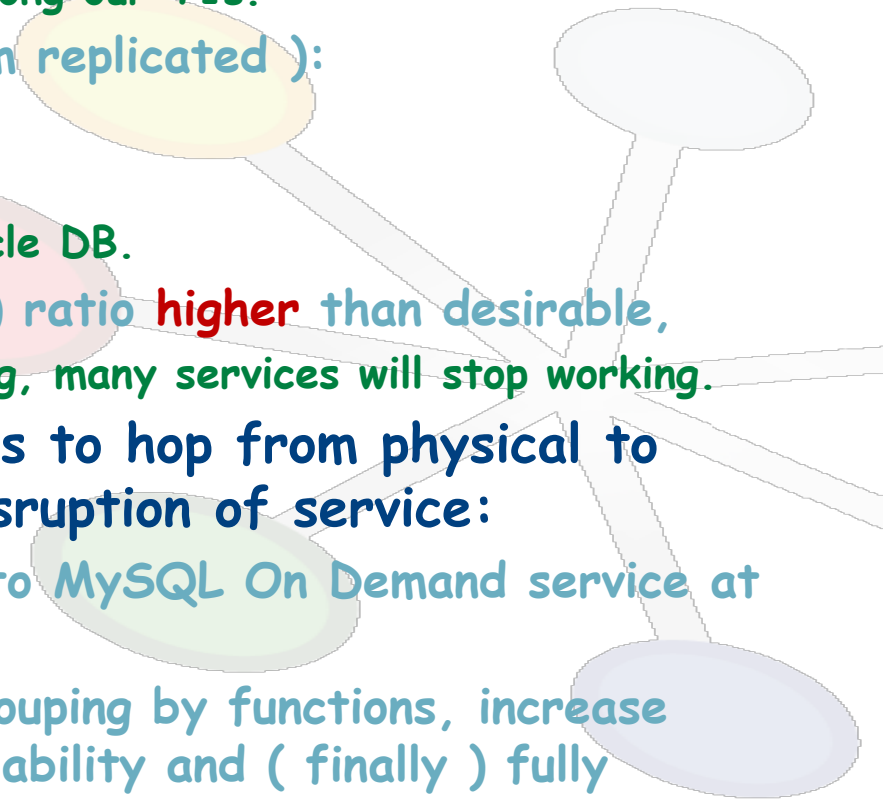
➤ components (some of them replicated):

- ~40 different services,
- ~40 different agents,
- ~20 MySQL DBs & 1 Oracle DB.

- (components / resources) ratio **higher** than desirable,
 - when a machine goes wrong, many services will stop working.

❖ Our infrastructure allows us to hop from physical to virtual machines without disruption of service:

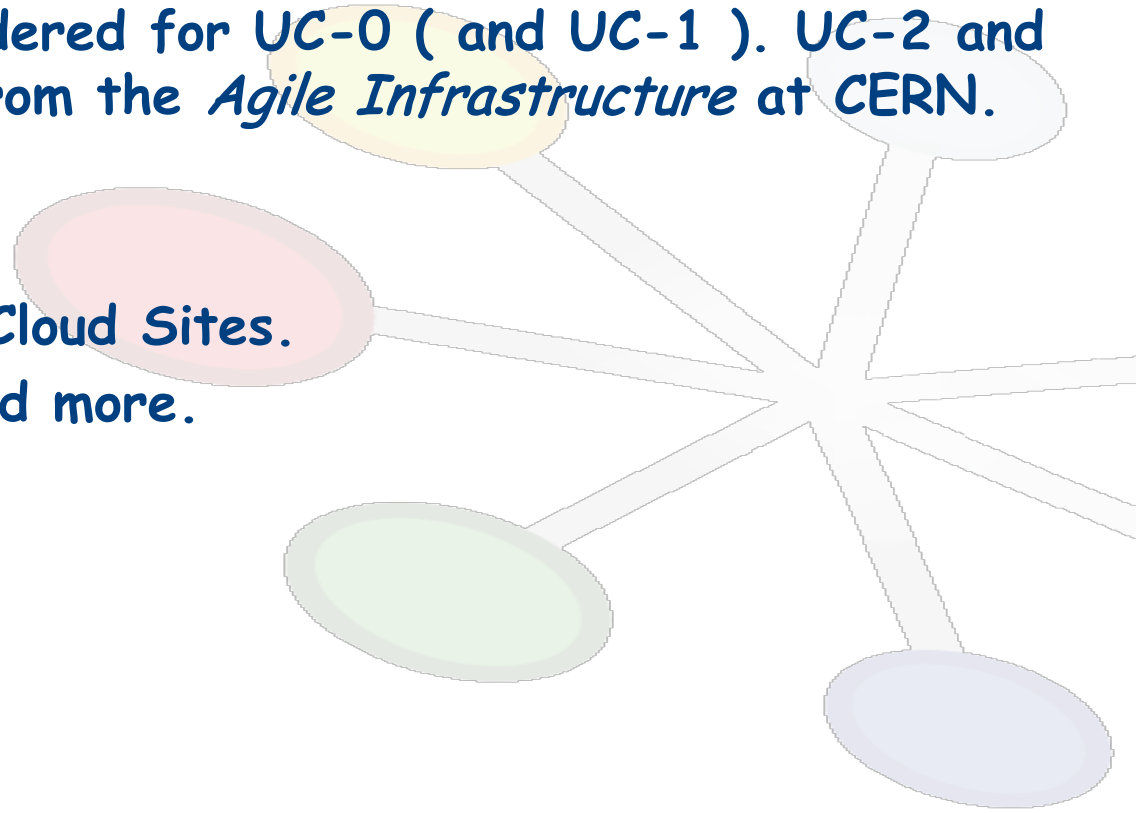
- MySQL DBs to be moved to MySQL On Demand service at CERN.
- Components isolation, regrouping by functions, increase availability, horizontal scalability and (finally) fully distributed.





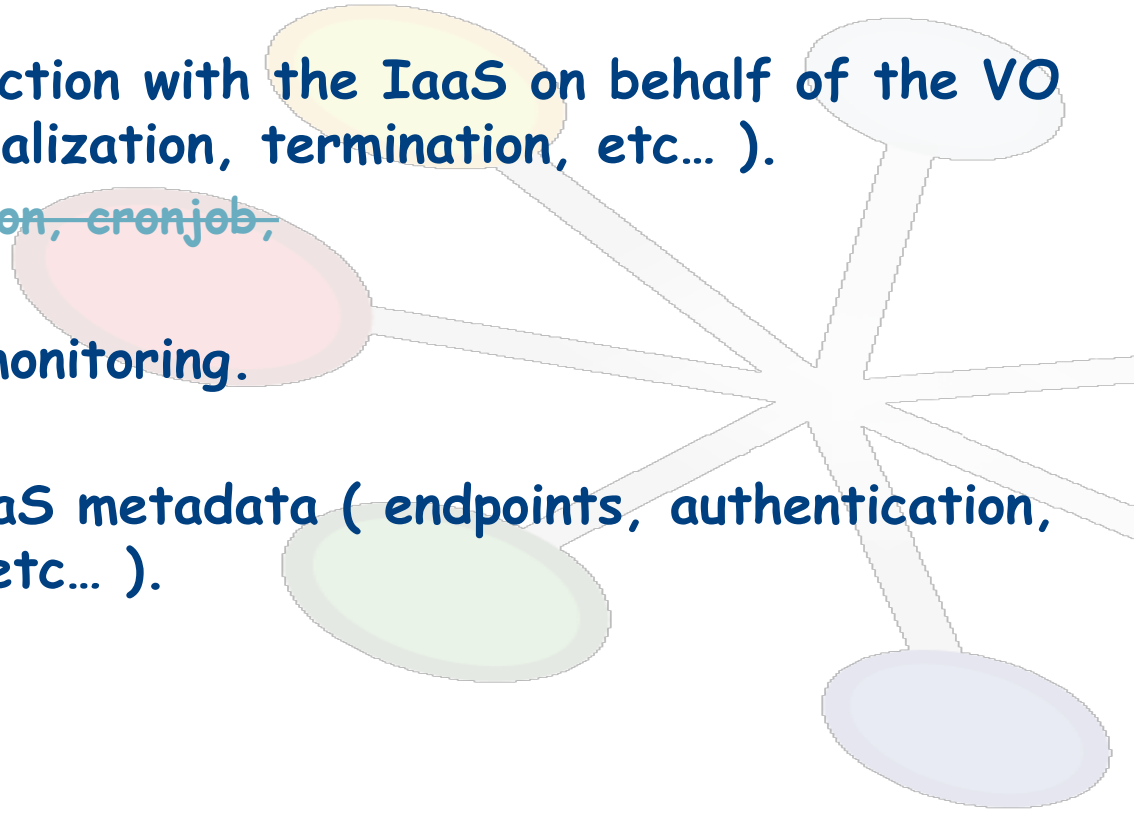
Solutions we considered for UC-0 (and UC-1). UC-2 and UC-3 profited from the *Agile Infrastructure* at CERN.

- ❖ PS-1: Core.
- ❖ PS-2: IaaS and Cloud Sites.
- ❖ PS-3: Goodies and more.





- ❖ Registry: metadata service (lifetime, configuration, status, etc...).
 - ~~Notebook, provider registries, custom DB,~~
 - VMDIRAC.
- ❖ Scheduler: interaction with the IaaS on behalf of the VO (quotas, VM initialization, termination, etc...).
 - ~~Manual submission, cronjob,~~
 - VMDIRAC.
- ❖ Accounting: job monitoring.
 - DIRAC.
- ❖ Configuration: IaaS metadata (endpoints, authentication, flavors, images, etc...).
 - DIRAC,
 - ~~none,~~
 - ¿ BDII ?.





❖ Image store:

- ¿ any ?,
- ~~none.~~

❖ OS:

- ~~SLC*~~,
- CernVM, ¿ μ CernVM ?,
- ~~Others (Fedora, CentOS).~~

❖ SW distribution:

- ~~tarballs,~~
- ~~AFS,~~
- CVMFS.
 - ¿ Local proxies ?.

❖ Security:

- ¿ AAA ?,
 - ¿ Credentials ?.
- ¿ logging ?.

❖ Monitoring:

➤ VMs:

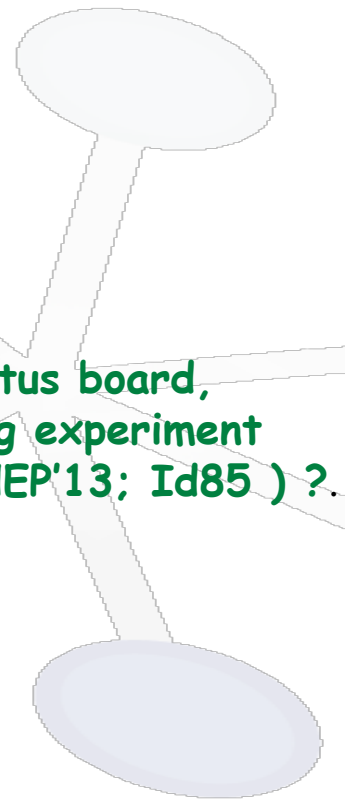
- Ganglia,
- ~~Lemon,~~
- VMDIRAC.

➤ IaaS providers:

- ¿ GocDB ?,
- e-mail, IT status board,
- ¿ Self managing experiment resources (CHEP'13; Id85) ?.

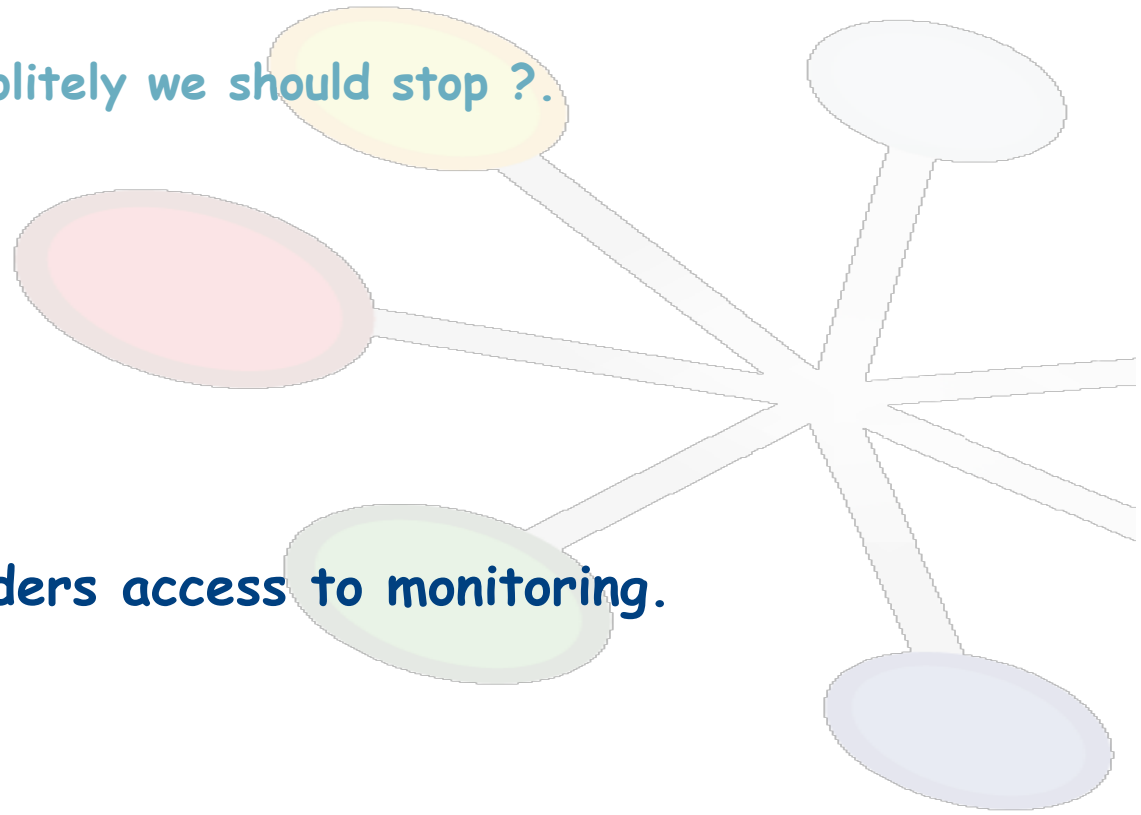
❖ Contextualization:

- Amiconfig,
- ~~Cloud-init.~~



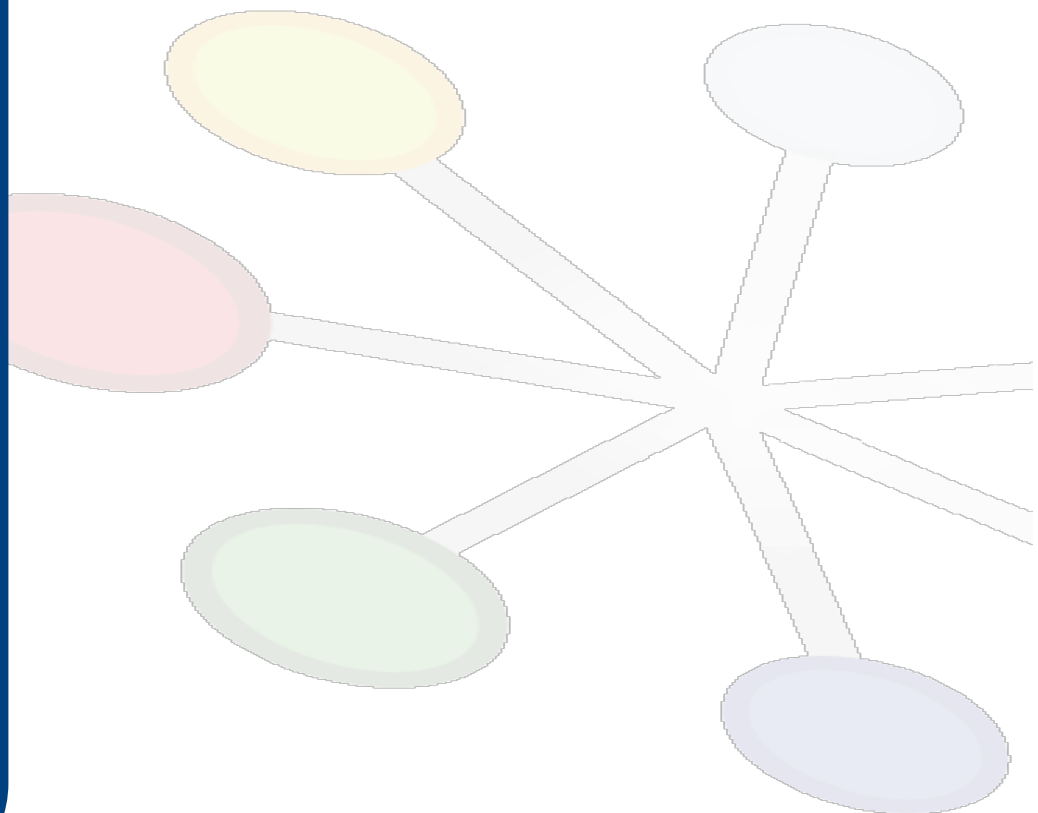
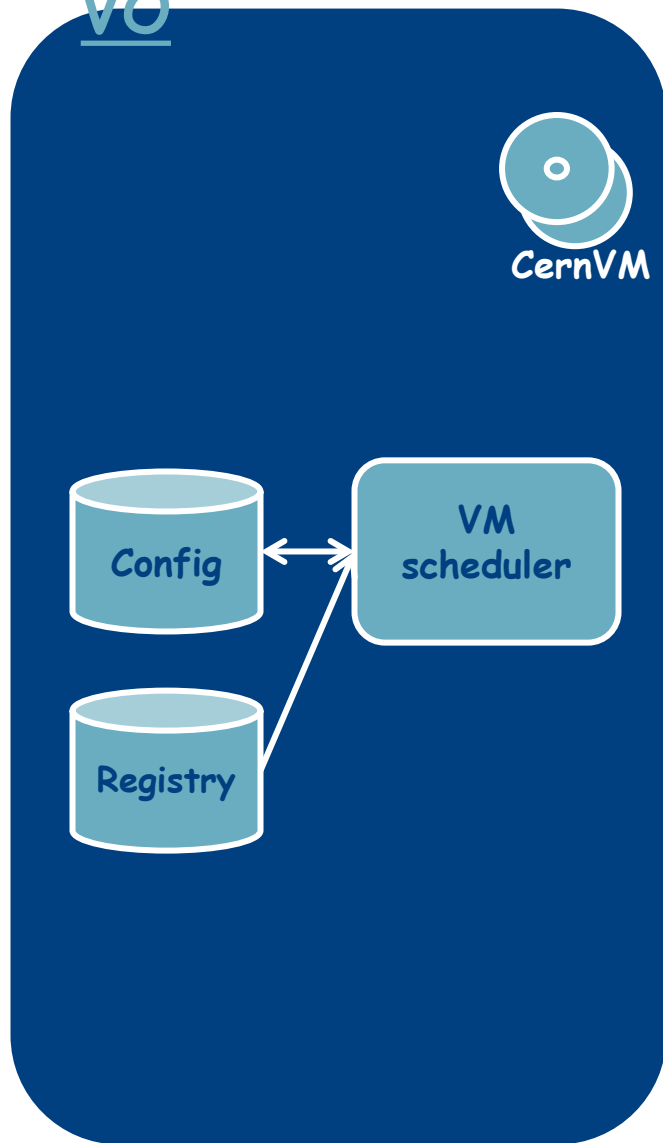


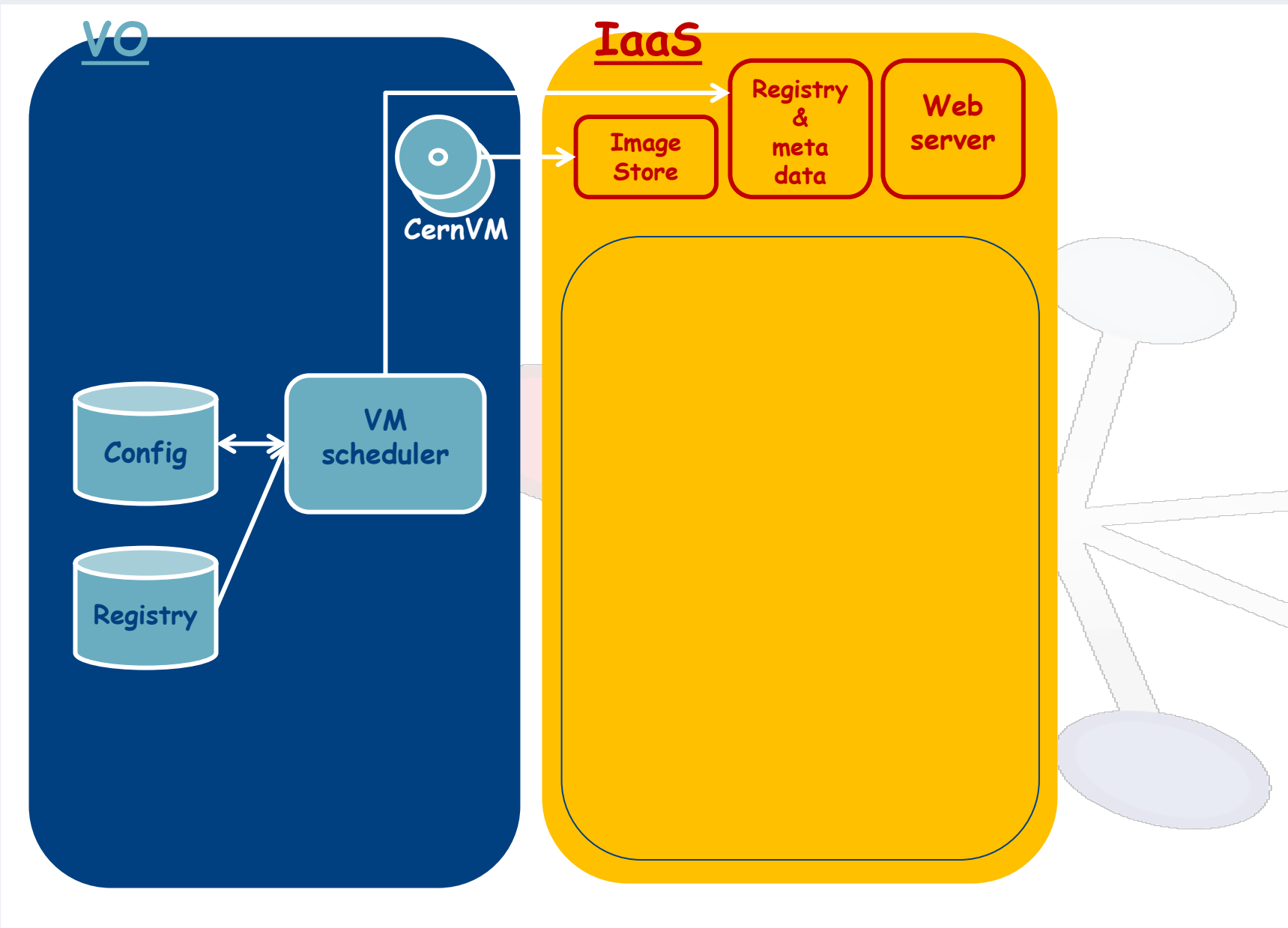
- ❖ **VMs lifetime (see Scheduler):**
 - run forever,
 - load based.
 - credit based,
 - ¿ let us know politely we should stop ?.
- ❖ **Web portal:**
 - ~~Horizon,~~
 - ~~Sunstone,~~
 - VMDIRAC.
- ❖ **API:**
 - DIRAC.
- ❖ **Grant IaaS providers access to monitoring.**





VO

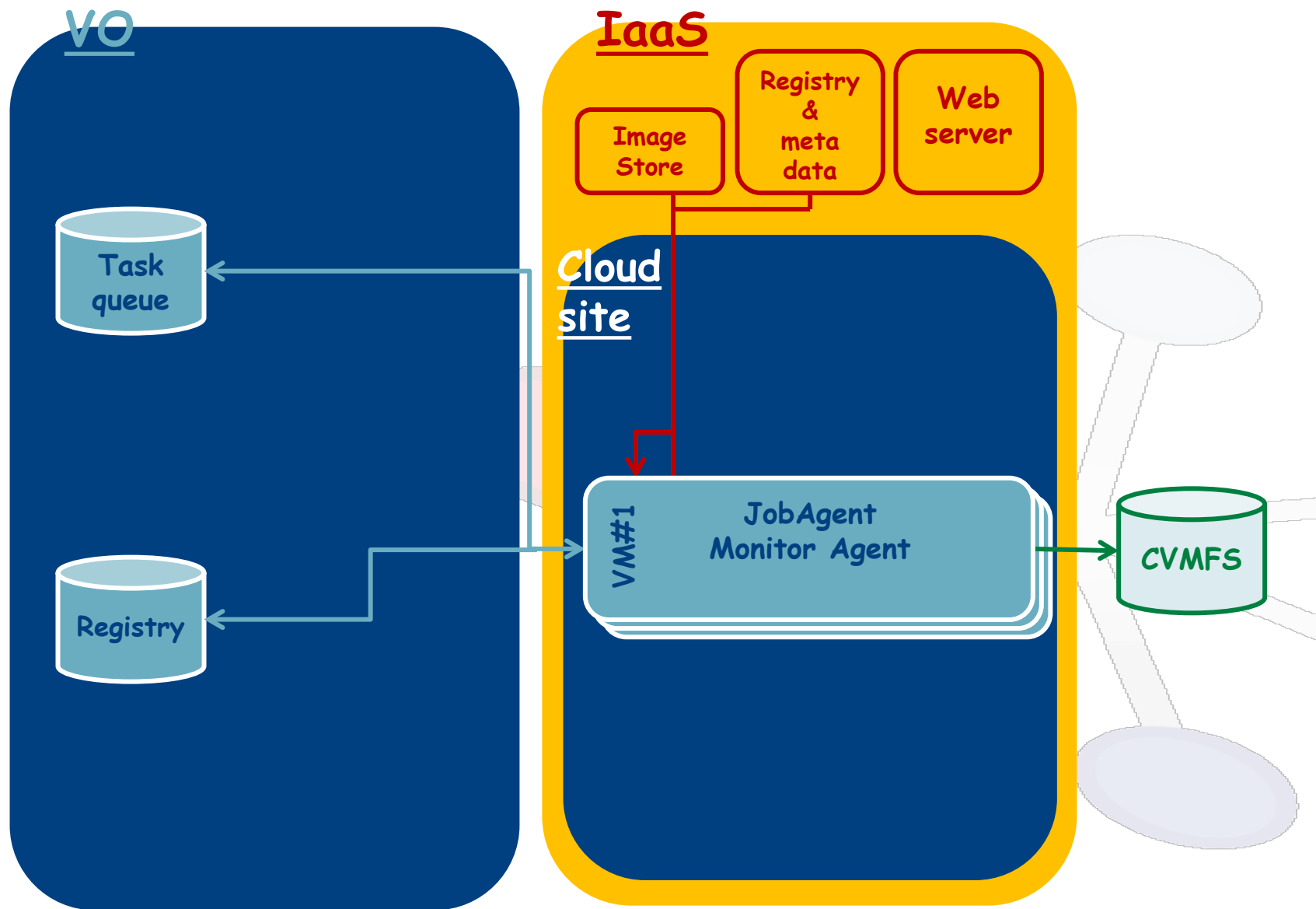






Architecture: Cloud Site with WNs

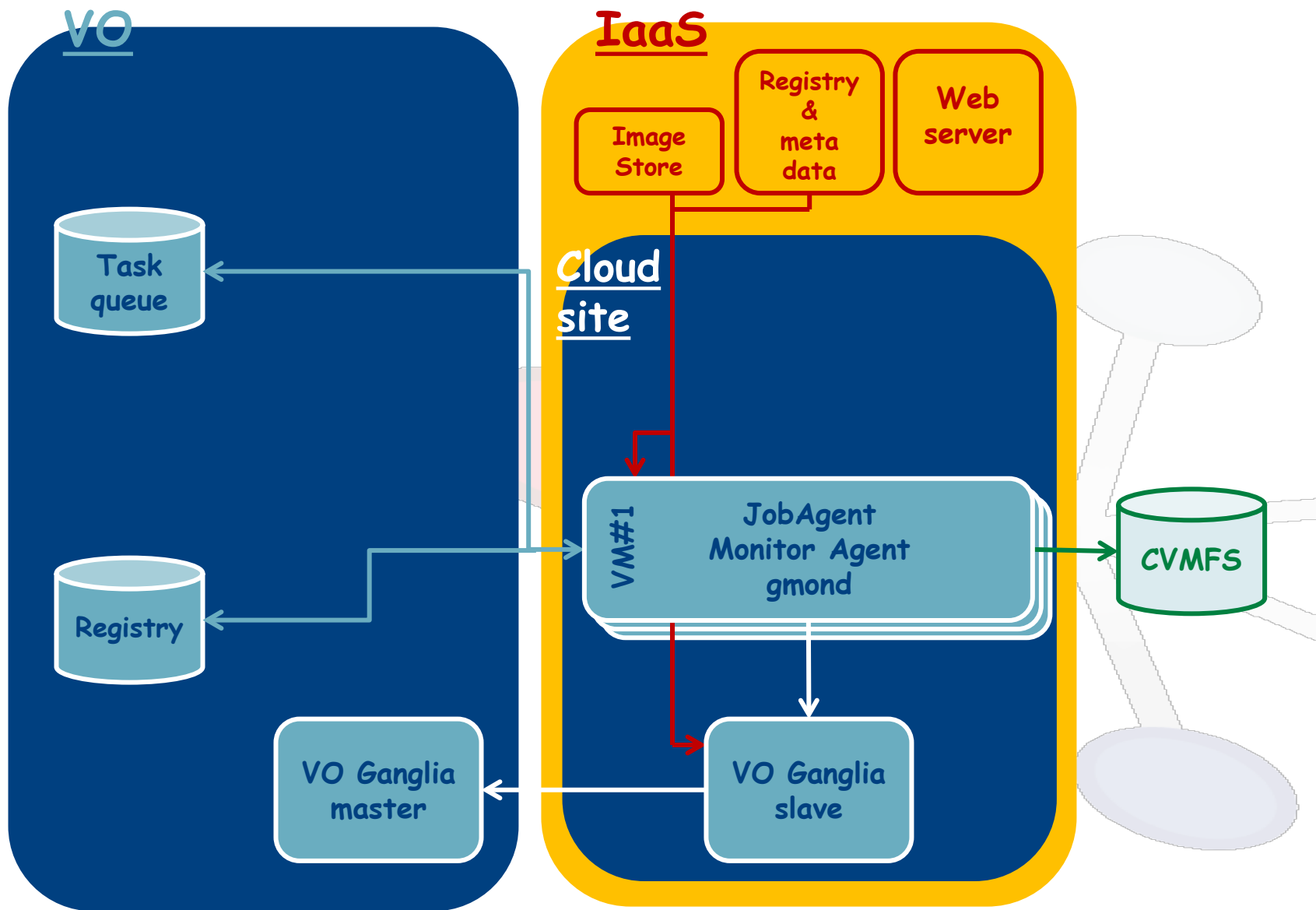
Integration .. Cloud .. LHCb





Architecture: Full Cloud Site with WNs

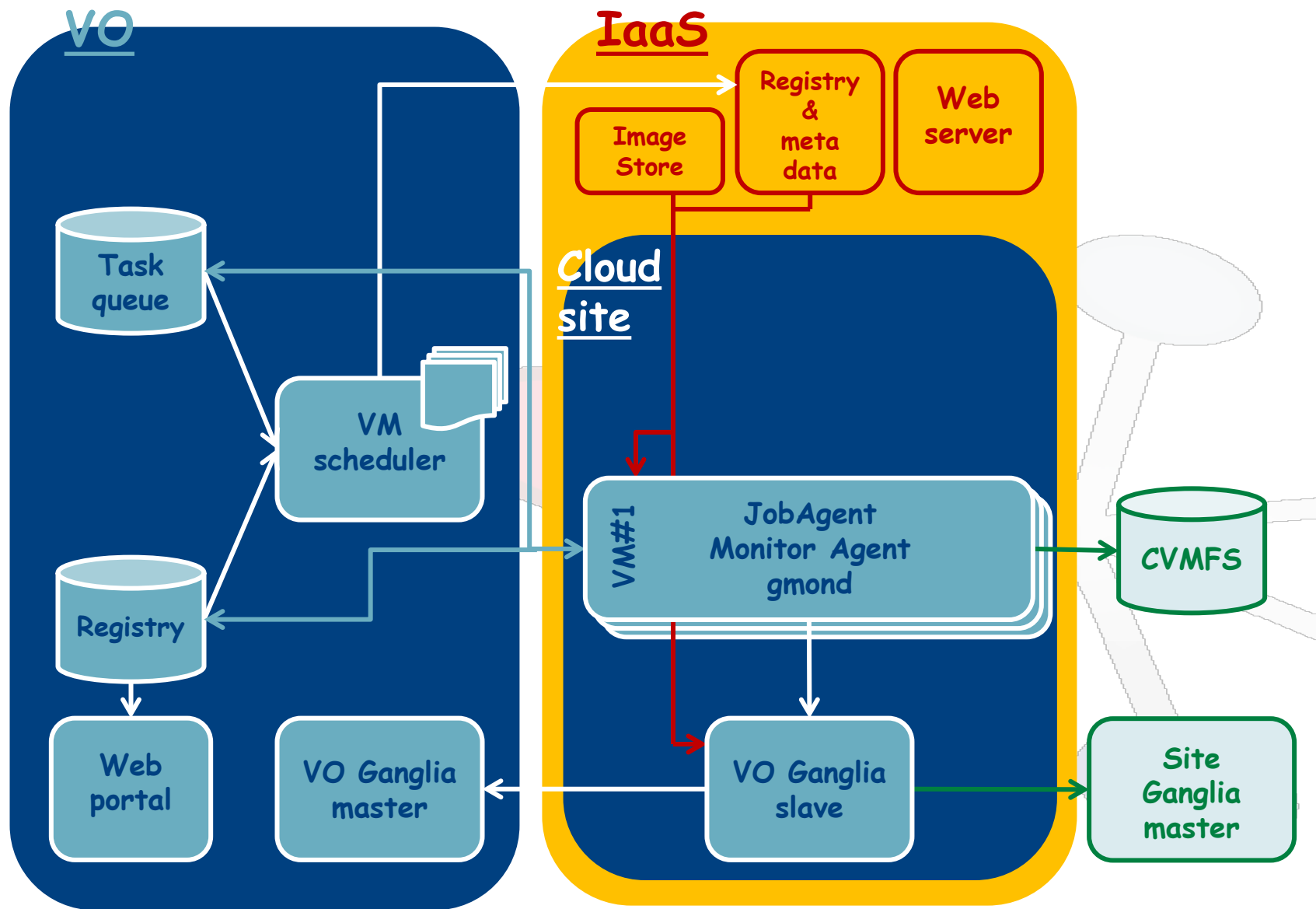
Integration .. Cloud .. LHCb





Architecture: Goodies and more

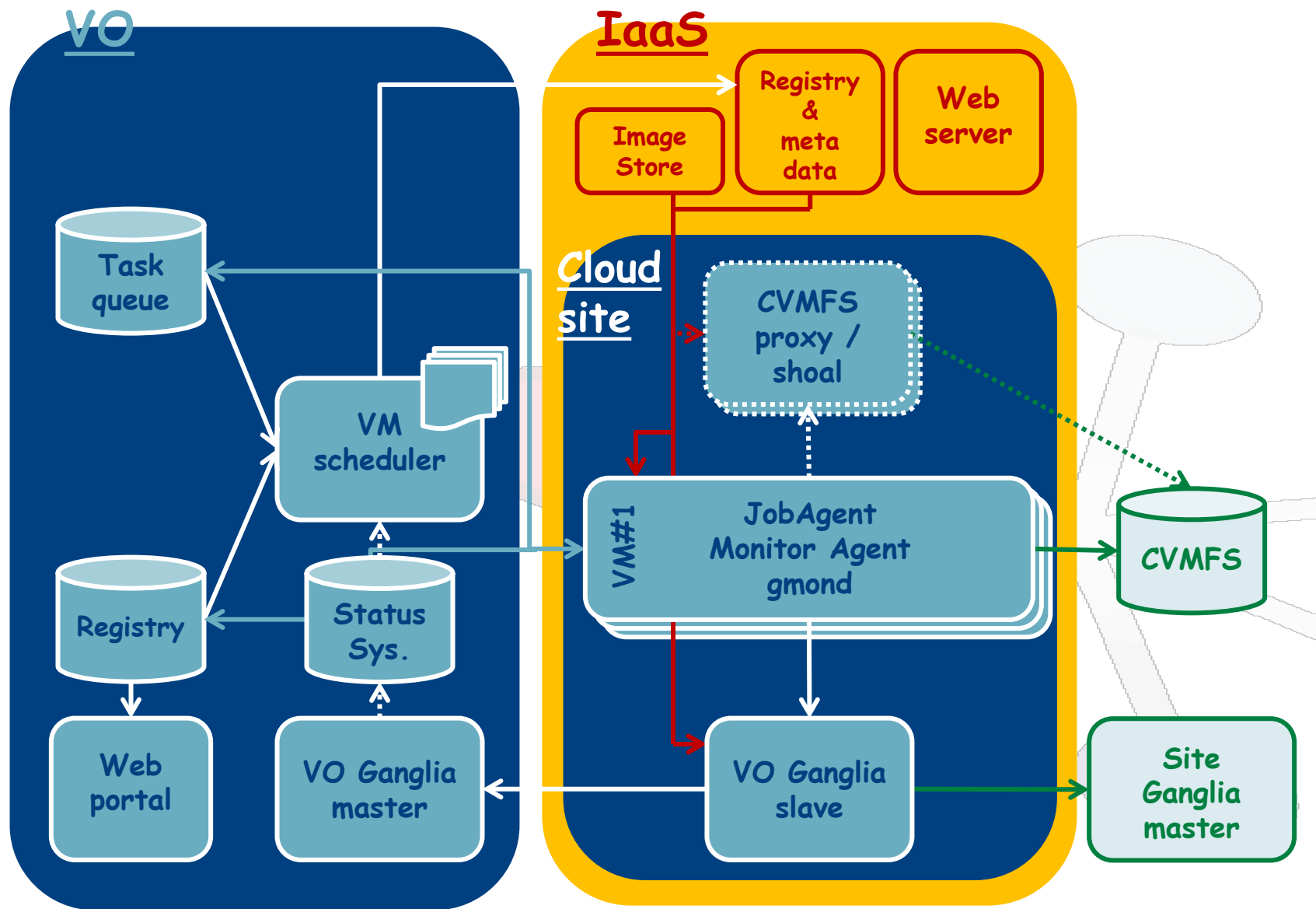
Integration .. Cloud .. LHCb





Architecture: Next steps

Integration .. Cloud .. LHCb



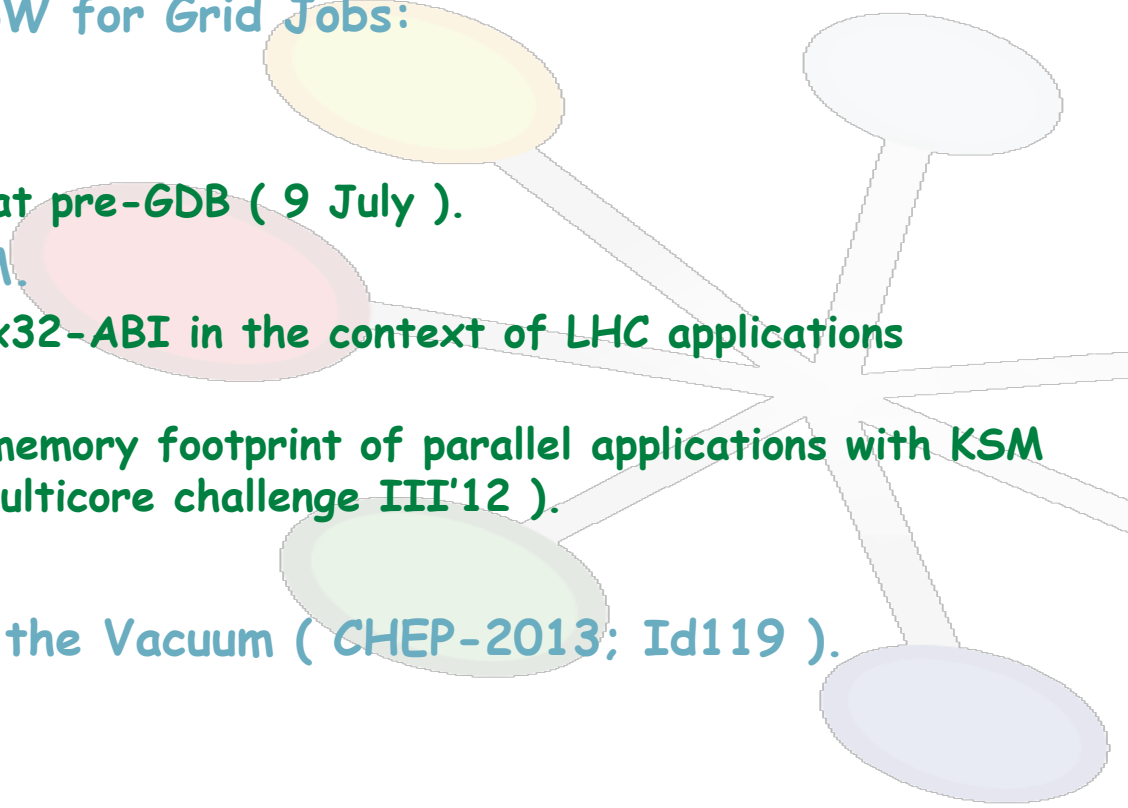


❖ Merging developments:

- Volunteer Computing.
- CVMFS based SW for Grid Jobs:
 - applications,
 - middleware.
 - See F.Stagni at pre-GDB (9 July).
- x32-ABI & KSM.
 - Evaluation of x32-ABI in the context of LHC applications (ICCS'13).
 - Reducing the memory footprint of parallel applications with KSM (Facing the multicore challenge III'12).

❖ Other activities:

- Running Jobs in the Vacuum (CHEP-2013; Id119).





Related Work: Volunteer Computing

Integration .. Cloud .. LHCb

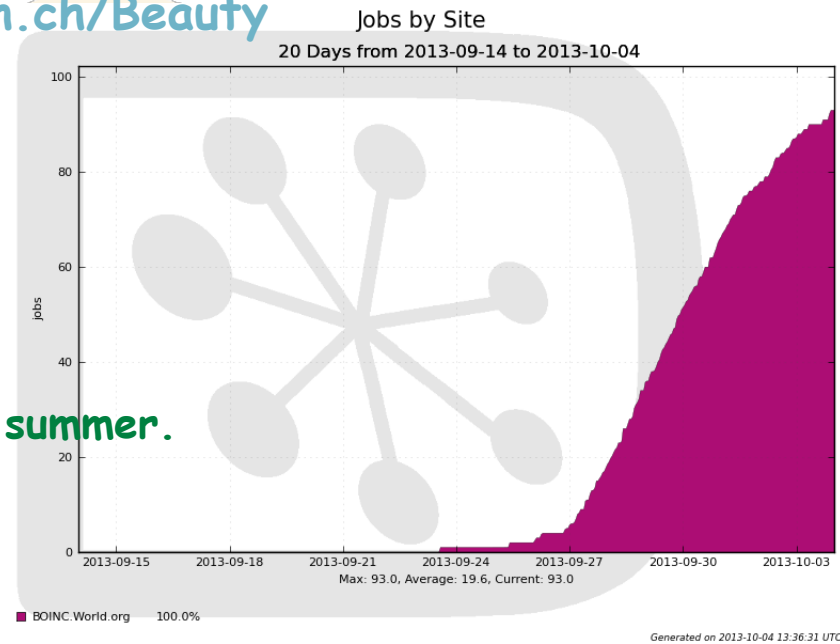
❖ Thanks B.Cabarrou for your CPU Cycles !

❖ If you want to play:

- <http://lhcbathome.web.cern.ch/Beauty>
- lhcb-boinc@cern.ch

➤ Different contributions

- BOINC sites:
 - Building 2 @ CERN,
 - University Student Labs @ summer.
- Single users.
 - There are power users !



Do you leave your desktop ON during the night ?
know how many desktops are there at CERN ?
at the collaboration's universities ?



❖ Short term (gather data & statistics):

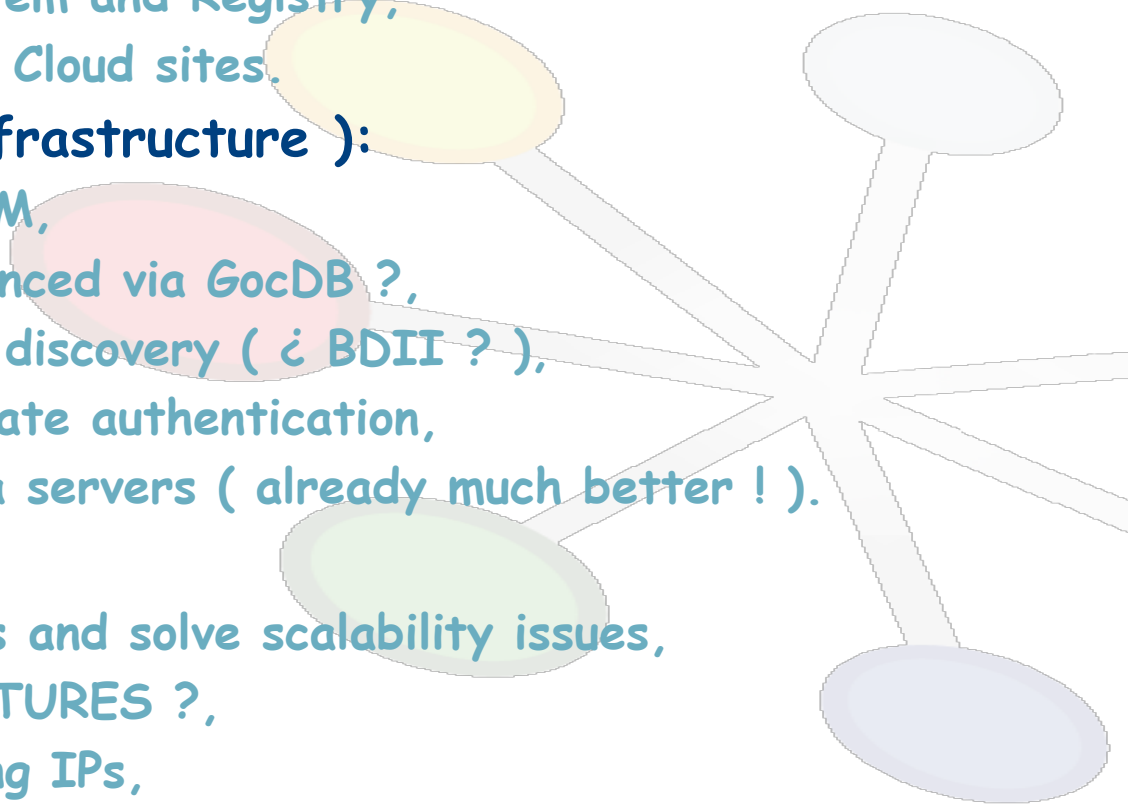
- more IaaS providers,
- read VM monitoring information from Ganglia and feedback the Status System and Registry,
- Run GaudiMP on Cloud sites.

❖ Medium term (infrastructure):

- move to μ CernVM,
- ϵ outages announced via GocDB ?,
- automatic IaaS discovery (ϵ BDII ?),
- proxy / certificate authentication,
- stable metadata servers (already much better !).

❖ Long term:

- Find bottlenecks and solve scalability issues,
- MACHINE_FEATURES ?,
- public vS floating IPs,
- and much more !



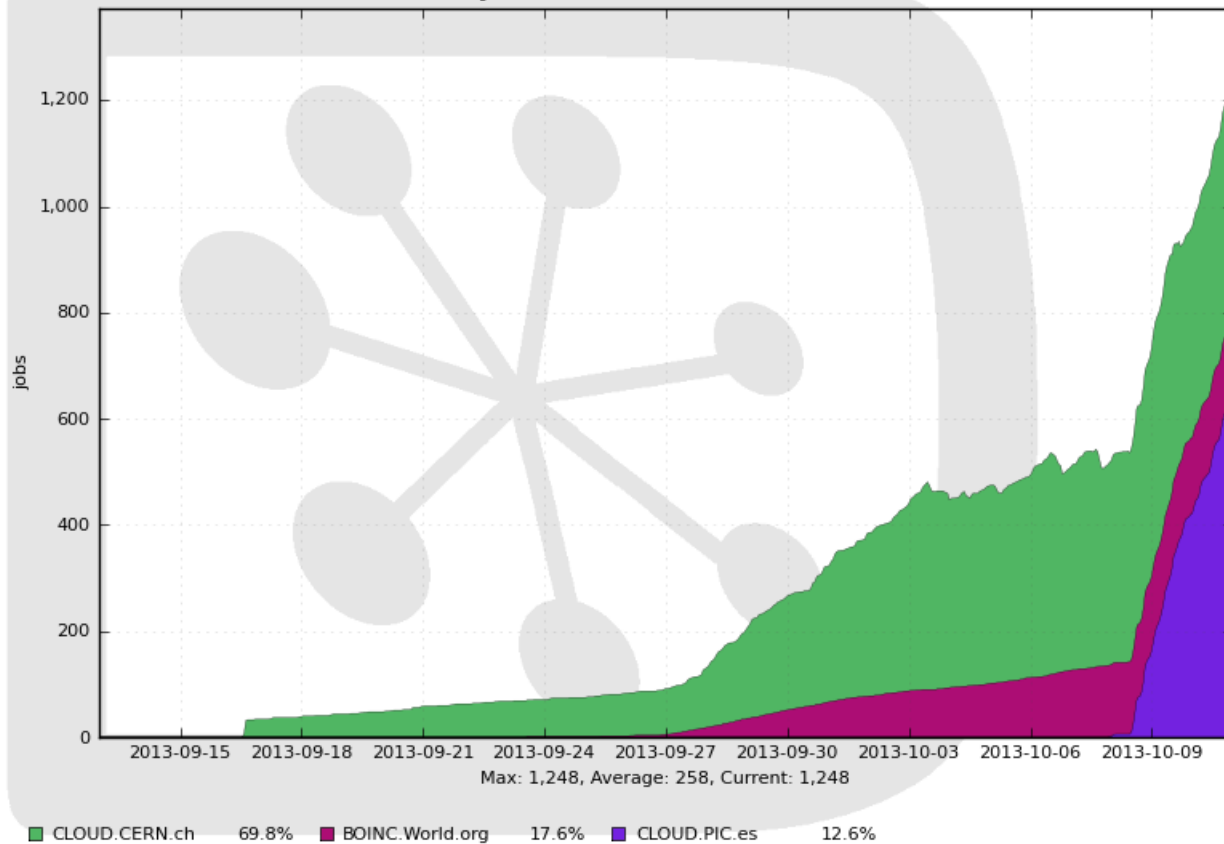


If we put everything together...

...one could say we solved a problem.

Jobs Completed in the Cloud and BOINC (last month)

28 Days from 2013-09-13 to 2013-10-11



Generated on 2013-10-11 08:16:59 UTC

Integration .. Cloud .. LHCb



There are many questions...

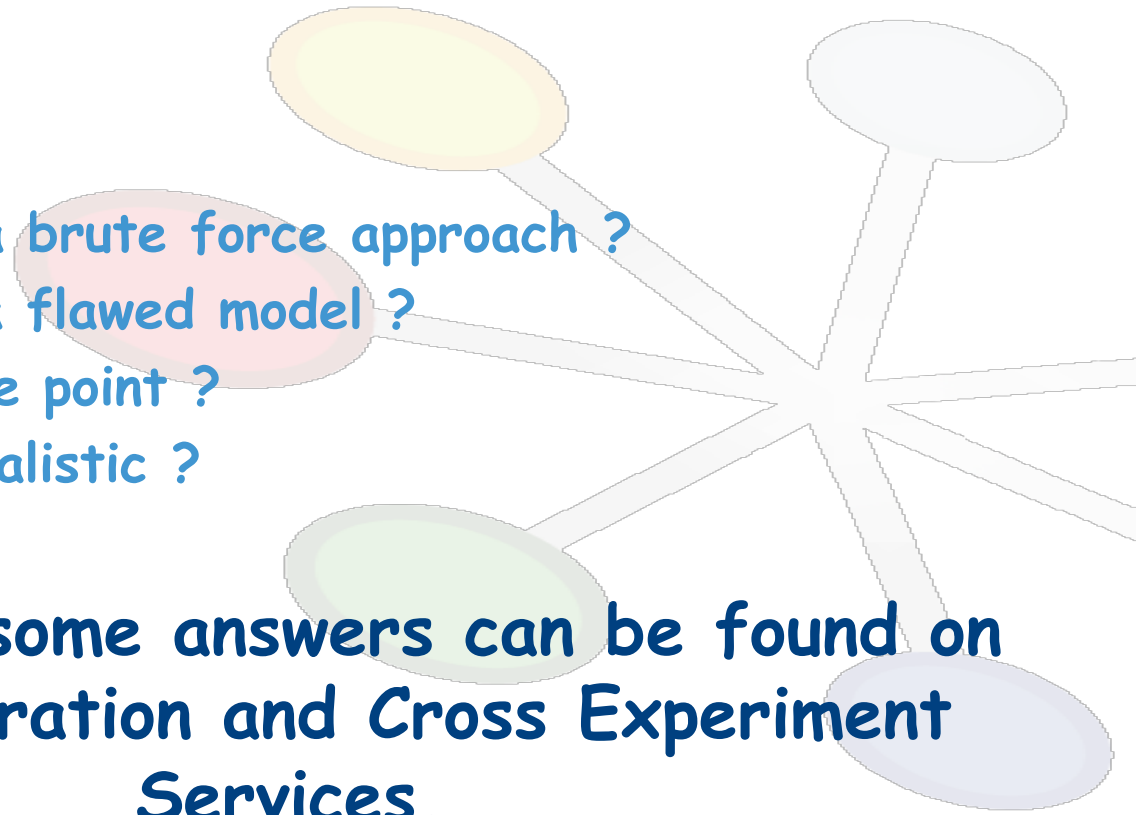
Integration .. Cloud .. LHCb

we formulated ourselves after solving the problems described on this talk.

Are we:

- Following a brute force approach ?
- Imposing a flawed model ?
- Missing the point ?
- Being unrealistic ?

**We believe some answers can be found on
Cloud Federation and Cross Experiment
Services.**





Big thanks

For your time and attention !

The LHCb Offline Team.

We do not forget:

LHCb collaboration for being our guinea pigs testing the BOINC.

OpenStack@CERN and OpenNebula@PIC for their help and patience.

Test4Theory group for their advices and guidance.

Integration .. Cloud .. LHCb



Architecture: Full Picture (WNs)

Integration .. Cloud .. LHCb

