Stefano Bagnasco, Dario Berzano, Andrea Guarise, Stefano Lusso, Massimo Masera, Sara Vallero

**STOA LHC ACTIVITIES IN TORINO**

Insert funny picture here
Focus on the **computational** part
- collaboration with other sites in data access activities
- adapting the tools to **different user groups**

Studies on **automatic elasticity**

Development of a **monitoring and accounting** platform

Definition of the “**Elastic Virtual Farm**” on-demand provisioning model

Development of a **self-service portal**

Gaining production-level **operational experience**!
The Virtual Analysis Facility

What is the VAF?

- A cluster of CernVM virtual machines: one head node, many workers
- Running the HTCondor job scheduler
- Capable of growing and shrinking based on the usage with elastiq

- Configured via a web interface: cernvm-online.cern.ch
- Entire cluster launched with a single command
- User interacts only by submitting jobs
- Elastic Cluster as a Service: elasticity is embedded, no external tools
- PoD and dynamic workers: run PROOF on top of it as a special case

Dario.Berzano@cern.ch - A grounds-up approach to High-Throughput Cloud Computing in High-Energy Physics
How elasticity works: elastiq

elastiq is a Python app monitoring the queue to make it elastic

cloud controller exposing EC2 API

Jobs waiting too long will trigger a scale up

Start new VMs

Shutdown idle VMs

Supports minimum and maximum quota of VMs

You deploy only the master node: minimum quota immediately launches VMs automatically

Integrated in CernVM 3

source: github.com/dberzano/elastiq

Dario.Berzano@cern.ch - A grounds-up approach to High-Throughput Cloud Computing in High-Energy Physics
If **new VMs need to be instantiated**, workers deploy time ranges from 2.5 min to 3.5 min. If **VMs are already available**, deploy time ranges from 16s to 3 min. The “golden number” of **30 workers** is reached in 2.5 min in the first case and 25 s in the latter.
- **Grid CEs are not cloud-aware**
- A Grid CE resources are almost always saturated (if they’re not, you’re doing something wrong)
- Idiosyncratic and complex middleware means starting a contextualized worker node from a fresh minimal OS image is slow and unreliable
BESIII Tier2
200 cores

How to use a tool inappropriately...

**OneFlow** is an optimal solution for stateless applications: i.e. load balancing (you do not choose which VM to undeploy)

But we want to handle a **batch farm**:

- jobs received in bunches with similar running time (from few hours to a day)
- allocated resources idle most of the time
- stateful application
- not cloud-aware (LRMS is Torque)
• application is never idle
• go above pledged if Cloud is free
• heterogeneous job duration (unpredictable)
• magic number of 6 cores per VM

Bunch of VMs that (tries to) **expand and shrink** (drain and shutdown worker nodes) **periodically**:
  • HTCondor
  • smaller VMs (some in draining mode all the time)

⇒ OneFlow with scheduled elasticity policy
The Virtual Farm

- **Virtual Router**
  - cloud-gw-216.to.infn.it
  - 193.205.66.216
  - 172.16.261.254

- **Virtual Private Network**
  - iSCSI
  - NFS

- **Persistent Storage**

- **Services**

- **Workers**
  - VirtualFarm001: 172.16.261.1
  - VirtualFarm002: 172.16.261.2
  - VirtualFarm003: 172.16.261.3

- **LAN**
  - Configurable port forwarding
Main user is ALICE

This is also for ALICE
Can submit to the Torino Cloud

CernVMVM images only!
SELF SERVICE PROVISIONING

EC2 and local credentials

Head node configuration

Custom user-data for context

Worker configuration

Much more not shown...
**VAF Monitoring with the ELK Stack**

**ELK stack**
- **logstash**
- **elasticsearch**
- **Kibana**

**MySQL DB**
- Also accounting INFN Grid services
- Dedicated DB tables

**TProofMon SenderSQL**

**Collect monitoring and accounting data from both IaaS and application**

**Investigation of the ELK stack to handle heterogeneous and unstructured data sources**

**Possible solution for Monitoring-as-a-Service providing uniform extendable monitoring platform to applications**
Example of ES analysis on a complex string
CONCLUSIONS AND OUTLOOK

- The **EVF model** derived from the VAF is practical and economic for many uses.
- Adding elasticity to **non cloud-aware applications** is hard (but we knew that).
- The **concurrent elasticity** of several applications in a saturated regime is still an unresolved issue.
- There is ongoing activity in the context of the **INDIGO-DataCloud** EU project.
- Also new developments will come from the **C3S project** of the University of Torino.


5. S. Bagnasco, D. Berzano, S. Lusso, M. Masera and S. Vallero, “Managing competing elastic Grid and Cloud scientific computing applications using OpenNebula”, accepted for publication in the Proceedings of CHEP2015, Okinawa, Japan


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