University of Zurich[UZH]

GC3: Grid Computing Competence Center

elasticluster - virtual HPC clusters on public and private clouds

Start, Setup, Resize and Stop your Cluster from Command Line

Nicolas Bär <nicolas.baer@uzh.ch>

29/04/2013

Enable scientific computing on a cloud infrastructure

Possible use cases

- Scientists: use cases already running on HPC clusters
- Sysadmins: easily test new batch configurations without the need of a real HPC cluster

Start and manage your SLURM or Grid Engine clusters on OpenStack or Amazon within just a few steps

Providing commands to manage your cluster

elasticluster is written in python and offers the following features:

- user-friendly command line interface
- simple configuration file to manage cluster setup
- start virtual clusters for scientific computing
- automated setup of queuing and monitoring system
- grow the cluster to fit your needs

Mix and match your cluster to fit your needs

Make your own cluster with:

- OS: Ubuntu, Debian, CentOS
- Batch system: SLURM / GridEngine / TORQUE
- Monitoring: Ganglia

... and more to come

Uses ansible to keep your cluster setup flexible

"Ansible is the easiest way to deploy, manage, and orchestrate computer systems you've ever seen. You can get started in minutes." <www.ansible.cc>

- Management over SSH
- Requires no bootstrapping or daemons
- Works on "vanilla" images
- Provides an easy-to-use language to define system management (playbooks)

... elasticluster provides predefined playbooks for setting up your cluster - you can modify the predefined playbooks or write your own

Demonstration

```
# start cluster
elasticluster start --name mycluster
--compute-nodes 10 slurmcluster

# grow cluster
elasticluster resize mycluster +3

# ssh to frontend
elasticluster ssh mycluster

#stop cluster
elasticluster stop mycluster
```

Relies on stable frameworks while offering a flexible design

Frameworks:

- boto: python interface to Amazon Web Services
- ansible: advanced system orchestration

Pluggable Design:

- Defined abstraction to provide alternatives to ansible (Puppet, Chef, CFEngine) or boto (gcelib)
- Provide your own ansible playbooks to use different cluster setups

Future Work

- Build an API to enable users to integrate elasticluster with their python code
- Integrate elasticluster with GC3Pie
- Resize the cluster based on workload
- Enhance functionality to further support scientific computing in the cloud

Thank You

Give it a try with

```
pip install elasticluster
```

and get more information on GitHub

https://github.com/gc3-uzh-ch/elasticluster

Command List

```
# start cluster
elasticluster start slurmcluster

# start different clusters with same configuration
elasticluster start --name mycluster
--compute-nodes 10 slurmcluster

# grow cluster
elasticluster resize mycluster +3

# list clusters
elasticluster list

# list nodes within a cluster
elasticluster list-nodes mycluster

#stop cluster
elasticluster stop mycluster
```

Why develop alternatives to StarCluster or VirtualCluster

Evaluation of StarCluster:

- Setup is bound to pre-configured images
- Not compatible with OpenStack (uses spec. Amazon functionality to identify clusters)

Evaluation of VirtualCluster:

- Setup is bound to pre-configured images
- Makes many assumptions about the underlying OpenStack setup
- Not sure about codebase maintenance