
Google ganeti

Ganeti

Ganeti Core Team - Google
LISA '13 - 5 Nov 2013

Commandline Overview

Using the cli

- Guido Trotter <ultrotter@google.com>
- Helga Velroyen <helgav@google.com>

Latest version of these slides

Please find the latest version of these slides at:

https://code.google.com/p/ganeti/wiki/LISA2013

Ganeti Commands...

- ... all start with `gnt-`
  - `gnt-cluster`
  - `gnt-node`
  - `gnt-instance`
  - ...
- ... all take a subcommand
  - `gnt-cluster info`
  - `gnt-node list`
  - `gnt-instance list`

```
gnt-[subsystem] [verb] --flags [noun]
```

Three ways to get help

- The man page
  - `man gnt-instance`
- List of subcommands
  - `gnt-instance`
  - `gnt-instance --help`
- Help on individual subcommands:
  - `gnt-instance list --help`

Help

```
# gnt-instance list --help
Usage
=====
  gnt-instance list [<instance>...]

Lists the instances and their status. The available
fields can be shown using the "list-fields" command
(see the man page for details). The default field list
is (in order): name, hypervisor, os, pnode, status,
oper_ram.

Options
=======
--no-headers            Don't display column headers
--separator=SEPARATOR   Separator between output fields ...
...
```

Commands run on the master

Commands must run on the master. Any other node will give you a friendly message. Scripts can use "getmaster" to know the right place.

```
# gnt-node list
Failure: prerequisites not met for this operation:
This is not the master node, please connect to node
'gnta2.example.com' and rerun the command
# gnt-cluster getmaster
gnta2.example.com
# ssh gnta2.example.com
WARNING:
This machine is part of a ganeti cluster.
# gnt-node list
Node              DTotal DFree MTotal MNode MFree Pinst Sinst
gnta1.example.com   3.6T  3.1T  64.0G 1023M 15.0G     5     3
...etc...
```

gnt-cluster

Cluster-wide operations

```
gnt-cluster info
gnt-cluster modify [-B/H/N ...]
gnt-cluster verify
gnt-cluster master-failover
gnt-cluster command ...
gnt-cluster copyfile ...
```

gnt-cluster example

```
# gnt-cluster verify
Submitted jobs 285450, 285451
Waiting for job 285450 ...
Sat Oct 27 19:14:08 2012 * Verifying cluster config
Sat Oct 27 19:14:08 2012 * Verifying cluster certificate files
Sat Oct 27 19:14:08 2012 * Verifying hypervisor parameters
Sat Oct 27 19:14:08 2012 * Verifying all nodes belong to an existing group
Waiting for job 285451 ...
Sat Oct 27 19:14:08 2012 * Verifying group 'default'
Sat Oct 27 19:14:08 2012 * Gathering data (3 nodes)
Sat Oct 27 19:14:10 2012 * Gathering disk information (3 nodes)
Sat Oct 27 19:14:11 2012 * Verifying configuration file consistency
Sat Oct 27 19:14:11 2012 * Verifying node status
Sat Oct 27 19:14:11 2012 * Verifying instance status
Sat Oct 27 19:14:11 2012 * Verifying orphan volumes
Sat Oct 27 19:14:11 2012 * Verifying N+1 Memory redundancy
Sat Oct 27 19:14:11 2012 * Other Notes
Sat Oct 27 19:14:11 2012   - NOTICE: 1 offline node(s) found.
Sat Oct 27 19:14:12 2012 * Hooks Results
```

gnt-node

Per-node operations

```
gnt-node list
gnt-node info
gnt-node remove node4
gnt-node modify \
  [ --master-candidate yes|no ] \
  [ --drained yes|no ] \
  [ --offline yes|no ] node2
gnt-node evacuate/failover/migrate node3
gnt-node powercycle node1
```

gnt-node examples

```
# gnt-node list
Node              DTotal DFree MTotal MNode MFree Pinst Sinst
gnta1.example.com   3.6T  3.1T  64.0G 1023M 15.0G     5     3
gnta2.example.com   3.6T  3.1T  64.0G 1023M 22.9G     4     4
gnta3.example.com      *     *      *     *     *     0     0
gnta4.example.com   3.6T  3.1T  64.0G 1023M 21.0G     4     6

# gnt-node info gnta1
Node name: gnta1.example.com
  primary ip: 172.15.155.15
  secondary ip: 172.99.199.1
      ...etc...
  primary for instances:
    - ginny.example.com
      ...etc...
  secondary for instances:
    - webcsi.example.com
      ...etc...
  node parameters:
    - oob_program: default (None)
    - spindle_count: default (1)
      ...etc...
```

gnt-instance

Instance operations

```
gnt-instance start/stop myinstance
gnt-instance modify ... myinstance
gnt-instance info myinstance
gnt-instance list
gnt-instance migrate myinstance
gnt-instance console myinstance
```

gnt-instance examples

```
# gnt-instance list
Instance             Hypervisor OS            Primary_node      Status  Memory
rocker1.example.com  xen-pvm    debian-server gnta2.example.com running   512M
webcsi.example.com   xen-pvm    debian-server gnta3.example.com running   1.0G

# gnt-instance info rocker1
Instance name: rocker1.example.com
UUID: 3244567d-a08a-4663-8349-c68307fab664
Serial number: 2
Creation time: 2012-07-05 20:08:14
Modification time: 2012-07-09 15:33:03
State: configured to be up, actual state is up
  Nodes:
    - primary: gnta2.example.com
    - secondaries: gnta3.example.com
  Operating system: debian-server
  Allocated network port: None
  Hypervisor: xen-pvm
    - ...
```

Job Queue

- Ganeti operations generate jobs in the master
  - (with the exception of queries)
- Jobs execute concurrently
- You can
  - cancel non-started jobs,
  - inspect the queue status, and
  - inspect jobs
- Development in progress: more fine-grained job-queue management

```
gnt-job list
gnt-job info
gnt-job watch
gnt-job cancel
```

gnt-backup

Manage instance exports/backups

```
gnt-backup export -n node1 instance1
gnt-backup import -t plain \
  {-n node3 | -I hail } --src-node node1 \
  --src-dir /tmp/myexport instance1
gnt-backup list
gnt-backup remove
```

gnt-group

Managing node groups

```
gnt-group add
gnt-group assign-nodes
gnt-group evacuate
gnt-group list
gnt-group modify
gnt-group remove
gnt-group rename
gnt-instance change-group
```

Custom output

Customize `gnt-* list` output with `-o`:

```
# gnt-instance list -o name,snodes
Instance             Secondary_Nodes
rocker1.example.com  gnta3.example.com
webcsi.example.com   gnta2.example.com
```

`--no-headers` is useful in shell scripts:

```
# gnt-instance list -o name,snodes --no-headers
rocker1.example.com
webcsi.example.com
```

What are the -o fields?

The `list-fields` subcommand lists all available fields.

```
gnt-group list-fields
gnt-instance list-fields
gnt-job list-fields
gnt-backup list-fields
```

Filtering a list

Filter output of `list` subcommands using the `-F` option:

```
# gnt-instance list -F 'pnode == "gnta1"' --no-headers -o name
ringo.example.com
george.example.com
john.example.com
paul.example.com
luke.example.com
```

More on filtering

Filtering language is described in `man ganeti`

Examples:

```
'(be/vcpus == 3 or be/vcpus == 6) and pnode.group =~ m/^rack/'

'pinst_cnt != 0'

'not master_candidate'
```

Tools: cross-cluster instance migration

Instances can be moved between clusters that share a common secret.

- Operation available via the CLI or RAPI
- CLI tool uses RAPI, and can be seen as an example
- Data is transferred directly between clusters

Setup

Setup common secret and RAPI authentication

BASH

```
ssh root@cluster1 --> root@cluster1:~#
gnt-cluster renew-crypto --new-cluster-domain-secret
cat > /var/lib/ganeti/rapi/users <<EOF
mover testpwd write
EOF

# copy /var/lib/ganeti/cluster-domain-secret to the second cluster

ssh root@cluster2 --> root@cluster2:~#
gnt-cluster renew-crypto --cluster-domain-secret=path_to_domain_secret
# rapi access can be the same or different. in production use hashed passwords.
cat > /var/lib/ganeti/rapi/users <<EOF
mover testpwd write
EOF
```

Execute move

Can be run on a third party machine

BASH

```
PWDFILE=$(mktemp)
echo testpwd > $PWDFILE

# Note: --dst-* defaults to --src-* if not specified
/usr/lib/ganeti/tools/move-instance --verbose \
  --src-ca-file=rapi.pem --src-username=mover \
  --src-password-file=$PWDFILE \
  [--dest-instance-name=new_name --net=0:mac=generate] \
  --iallocator=hail cluster1 cluster2 instance.example.com
```

Bugs:

- Either--iallocator or nodes must be specified manually
- Move is slower than it ought to be
- Instances with file-based disk template not supported.

Thank You!

Questions?

Survey at https://www.usenix.org/lisa13/training/survey