
Google ganeti

Ganeti

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

Ganeti and Networks

- Guido Trotter <ultrotter@google.com>
- Helga Velroyen <helgav@google.com>
- Slides contributed by Dimitris Aragiorgis <dimara@grnet.gr>

Latest version of these slides

Please find the latest version of these slides at:

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

Overview

- Why `gnt-network`?
  - MAC + IP + link + mode = enough?
  - challenges
  - `gnt-network` support
- `snf-network` + `nfdhcpd`
- Hands on `gnt-network`

MAC + IP + link + mode = enough?

NIC configuration

- DHCP: Subnet? IPv6?
- `mode=bridged`. `brctl addif` only? firewall?
- All NICs same MAC prefix. Why?

Management

- Which VMs are on the same collision domain?
- A VM wants an IP. Which one is available?
- One router broke down. Renumber VMs now! How?

Challenges

- easy way to assign IPs to instances
  - If resources are shared in multiple clusters, allocation must be done externally
- provide a way to configure each NIC differently
- find a way to hide underlying infrastructure
- better networking overview

`gnt-network`: Who does what?

masterd: `config.data` integrity

- abstract network infrastructure: network + netparams per nodegroup
- IP uniqueness inside network: IP pool management
  - bitarray, TemporaryReservationmanager, Locking
- encapsulate network information in NIC opjects: RPC

external scripts and hooks: ping `vm1.ganeti.example.com`

- use exported environment provided by `noded`
- `brctl`, `iptables`, `ebtables`, `ip rule`, etc.
- update external dhcp/DNS server entries
- let VM act unaware of the "situation" (`dhclient`, etc.)

`gnt-network` + external scripts

- `gnt-network` alone is nothing more than a nice `config.data`
- `snf-network`: node level scripts and hooks
- `nfdhcpd`: node level DHCP server based on NFQUEUE

snf-network

node level scripts and hooks

- overrides Ganeti default scripts (`kvm-ifup`, `vif-ganeti`)
- looks for specific tag types in NIC's network
- applies corresponding rules
- created `nfdhcpd` binding files
- provides hook to update DNS entries

nfdhcpd

node level DHCP server based on NFQUEUE

- listens on specific NFQEUE
- updates its leases db
  - `inotify` on specific directory for binding files
- mangles DHCP requests and replies based on it's db
- responds to RS and NS for IPv6 auto-configuration

gnt-network

Examples

Create and connect a new network

```
gnt-network add --network 192.168.1.0/24 --gateway 192.168.1.1 --tags nfdhcpd net1
gnt-network connect net1 bridged prv0
```

Create an instance inside this network

```
gnt-instance add --net 0:ip=pool,network=net1 ... inst1
gnt-instance info inst1
gnt-network info net1
```

gnt-network + snf-*

Examples

Use snf-network and nfdhcpd

```
apt-get install snf-network nfdhcdpd
iptables -t mangle -A PREROUTING -i prv+ -p udp -m udp --dport 67 \
   -j NFQUEUE --queue-num 42
ip addr add 192.168.1.1/24 dev prv0
```

Test connectivity

```
gnt-instance reboot inst1
ping 192.168.1.2
```

References

- snf-network: http://code.grnet.gr/git/snf-network
- nfdhcpd: http://code.grnet.gr/git/snf-nfdhcpd

Thank You!

Questions?

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