a Gentoo environment at Gaikai

Guido Serra <guido@gaikai.com>

Gaikai is internet at scale, Ganeti …&Gentoo

- I got the chance to work on large numbers of machines
- With several networks spreading across the globe
- Have hands-on on a real production Ganeti cloud infrastructure
- Contribute to the project and the community around Ganeti
- Operate a production environment completely Gentoo based
- Work with people acquired from other global internet companies

Gaikai, the gaming cloud company

Gaikai has been acquired by Sony Computer Entertainment in 2012

Initially basing its business as advertising agency for new videogames that users would be able to stream from the Amazon cloud, it had to redesign its solution to provide what today is Playstation NOW

As of now PsNOW is (beta) publicly available in USA and Canada as a way to play PS3 videogames without a PS3 console.

Ganeti is used for the coordination and management of the streaming infrastructure.

The Berlin office

Opened at the beginning of the year to cover the on-call shift of USA WEST coast night hours. We are not only guaranteeing our colleagues a safe sleep but we do also contribute in the tool automation and evolving architecture design of the infrastructure.

Everybody else is in Orange County, Southern California

+some people in the Bay Area

+some other out to Tokyo

Abstract of the talk

- Why did we choose ganeti
- How it is currently implemented
- How we are changing the current setup
- Evaluation on how will we like to change it in the future

Ganeti is well documented (& active)

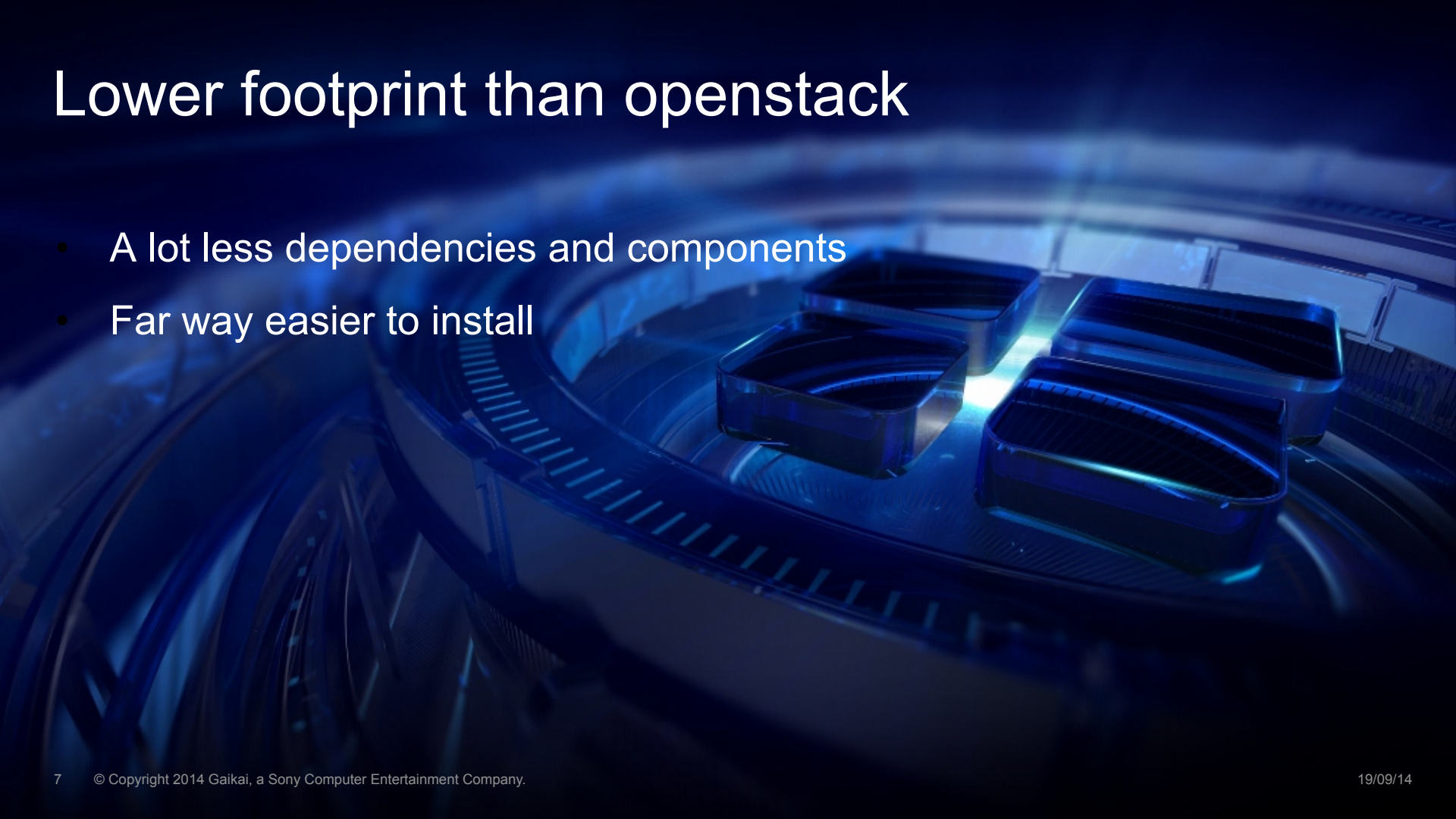
It can get even better, but there is absolutely nothing I can complain.

It has a wide & active community. Plus companies actively investing into it: with dedicated resources (people) contributing code.

Here YOU are. I just stepped in, with the intention of doing more.

Lower footprint than openstack

- A lot less dependencies and components
- Far way easier to install



Ganeti scales DOWN

One rack per environment. Multiple environments per continuous integration automation.

Smallest setup is 1 box for ancillary services like NFS, tftp, configuration management and DHCP; and 4 hypervisor machines

Very quick to install -> try ganeti vagrant

My laptop was my first playground thanks to that project.

TNX to the people here in this room that contributed to that.

Yes, Ganeti is very fast&easy to install.

How it is currently implemented: requirements

Distributed machines having same role in a load-balanced service pool.

- That is why we have a 1:1 association with each hypervisor

  (I do have heard about “labels”, that is part of the future for us)

No persistency. We are not using DRBD. Apart of special cases.

- Reboostrap from scratch on upgrade => roll&deploy

We run (prod)2.7.1, (test environment)2.9.1

We will run 2.11.5 => we aim to the latest stable release.

In particular, for the following features:

- Pool users, for VMs privileges
- Non root users to ganeti's daemons

Isolated failure domains: One cluster per rack

Good for current upgrade from 2.7.1 to 2.11.5

- Hypervisor rebootstrap is involved, and services multiple restarts
- All the nodes in the cluster are being involved in the process.

(probably not so good for dynamic VMs distribution) since I have to talk to multiple RAPI endpoints

- I would like to investigate on “labels” and “iAllocator”, and how one affects the other on distributing the VMs.

Image based system

Also for the bare metal. We use prebuilded OS images everywhere.

Installing from sources would take too long on entire stacks upgrade.

…and it is highly risky, since you depend on sources that are distributed across the internet. And in case anyone of them goes down, the deployment is affected. A proxy would only mitigate that.

But this also means that /root/.ssh/id_dsa is not preserved, neither /var/lib/gentoo … yes, we are working on that. Thankfully they are still there after the OS upgrade, just in an inactive partition.

Current challenges: slow instance creation

Especially when triggering sequentially via RAPI the creation of multiple VMs that reside in the same hypervisor, and maybe also on the same LVM partition and/or drive. [will the opportunistic locking available > 2.7 solve that?]

UPDATE (SOLVED): discussing the issue during the conference

…mkfs.ext4 exits without waiting the termination of the journaling

Orchestration via RAPI calls

On rollouts everything is handled externally from Ganeti,
via a custom configuration management & deployment tool.

How would that integrate with the "labels"? If we are going to use them.

- Is it going to be an issue the fact that we have one cluster per rack?
- We do state the rack number in the VMs' fqdn
- And we update the DNS system before triggering the RAPI calls.
  - How are other people doing?

Security: privileges separation for daemons and instances

Each daemon can have its own user, uid!=0 (apart of “noded”).

Since the naming of such users/groups is stated at compile/config time, a Makefile generated script is produced and distributed with the binary version. (TNX Michele Tartara, from Google, for providing it).

Noded is still running as root, it can be changed but that requires providing a user enough (potentially harmful) root privileges. That went beyond what I had in scope for the upgrade to 2.11.5

I ported to Gentoo such script, as it was Debian/Ubuntu specific.

Thanks to Patrick McLean, Gentoo/GAIKAI

We have several official Gentoo contributors onboard at Gaikai: he is one of them. He published last Friday (August 26th) my patches to Ganeti's Gentoo ebuild/package.

Check it out, and let us know.

published ebuild 2.11.5 to Gentoo upstream

GanetiCON

PlayStation

app-emulation/ganeti

Ganeti is a virtual server management software tool

```
*ganeti-2.11.5 (29 Aug 2014)
29 Aug 2014; Patrick McLean (chutzpah) +ganeti-2.11.5.ebuild,
+files/ganeti-2.11-daemon-util-tests.patch,
+files/ganeti-2.11-dont-nest-libdir.patch,
+files/ganeti-2.11-dont-print-man-help.patch,
+files/ganeti-2.11-useradd.patch, metadata.xml:
Version bump. Add multiple-users USE flag that enables ganeti's multiuser
support (contributed by Guido Serra). Add monitoring USE flag to enable
ganeti's monitoring daemon.
```

	alpha	amd64	arm	hppa	ia64	ppc	ppc64	sparc	x86
2.11.5		~							~
2.11.2-r3		~							~
2.10.5-r3		~							~
2.9.5		~							~
2.7.2		~							~
2.4.5-r1		+							+

app-emulation | Homepage | GPL-2 | ChangeLog Metadata | Similar | Bugs | Forums

security part2: VMs' users pool

Each kvm/qemu process running with its own user, and not root.

We could have published a helper there too, but we decided that it is up to whoever install the system to decide the way to provision such users pool.

Using it is just matter of a configuration change at Ganeti's

Instance migration between clusters

No specific USE case yet.

(question) Is it going to bring any benefit to our current setup?

Or… having a single cluster? What benefits will it bring?