

Xen 4.8 at Gandi

Vincent Legout

vincent.legout@gandi.net

Xen Developer and Design Summit 2017

July 12, 2017



Gandi



- I have been working for Gandi for ≈ 1 year
- on Simple Hosting (PaaS, i.e. LXC) & Servers (IaaS)

Outline

- Xen history at Gandi
- How we moved from Xen 4.1 to Xen 4.8
- How we use 4.8 features (especially live patching)

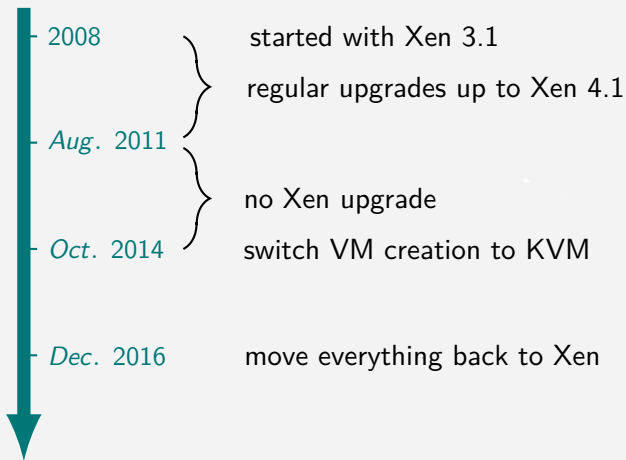
Virtualization at Gandi

- \approx 10000 VMs today (customers & internal)
 - ▶ x86 only (32 & 64 bits)
- Xen 4.8 (almost) everywhere
 - ▶ mostly PV, some PVHVM (e.g. for VF, FreeBSD)
- Home made code to orchestrate VMs
 - ▶ daemon running on each host
 - ▶ makes calls to xl

Xen features usage

- Default scheduler & cpupool (no NUMA)
- Cap of domU according to its number of vCPUs
 - ▶ similar weight for all vCPUs, no pinning
- VMs: only net, block (iSCSI), & console (e.g. no USB, video)
 - ▶ CPU hotplug: attach & detach
 - ▶ no longer any memory ballooning
 - ▶ PCI passthrough for internal VMs
- PV & PVHVM with direct kernel boot (kernel or grub)

Xen versions history



Upgrade from 4.1 to 4.8

Why? Reduce impact for customers, avoid reboot

- Live patching
- Security support

Tooling & Build process

Upgrade to use xl and python bindings (xen.lowlevel)

- Was previously using xend
- Few bugs in xl, patches sent and accepted very quickly
 - ▶ very much appreciated!
 - ▶ allows us to carry almost no patch
- Hypervisor: built from the git stable branch
- Userspace: based on the package in Debian Stretch

Live migration from 4.1 to 4.8

- Not officially supported
 - ▶ special tool to convert migration stream
- Test migrations went well but issues after a few days/weeks
 - ▶ disk corruptions, crashes
 - ▶ might be due to domU kernel (Linux suspend & resume?)
- Reboot was needed
 - ▶ XSAs arrived just in time, a good excuse to reboot all VMs!

Refresh VMs & disks

- Disks: from no partition or MBR/DOS to GPT
 - ▶ plan transition to OVMF & UEFI
- Switch to PvGrub by default
 - ▶ avoid maintaining custom kernels
 - ▶ will it still work with PVH2?
- Still need to support all existing VMs and disks
 - ▶ 10 years old VMs still running!

domU kernels

- Used to carry patches, lead to a few issues
- Issue with detaching disks
 - ▶ no automatic unmount
- No daemon inside domU
 - ▶ communication only through system events
 - ▶ e.g. custom udev rules

Live Patching

- Used since upgrading to Xen 4.8 (late 2016)
- Easy to use
 - ▶ `livepatch-build-tools.git` & `xen-livepatch`
- Did not generate any crash/bug

Live Migration

- Especially useful for XSA that cannot be live patched
 - ▶ avoid having to reboot all VMs
- Custom tool to migrate VMs
 - ▶ according to free nodes, their CPUs, ...
- Rolling updates, no VM downtime

Live Migration Issues

- A few, only $< 0.5\%$ of all the VMs crashed
 - ▶ only VMs in bad states, no suspend & resume
- Issues with a custom domU Linux kernel (3.10-i386)

Conclusion

- Overall, really satisfied with Xen
 - ▶ especially Xen 4.8 new features
 - ▶ XSA management, embargoes
- Some downsides:
 - ▶ several XSAs lately
 - ▶ clearer roadmap (e.g. PV/PVHVM, PVH2, OVMF)

Future - Plans

- More testing
 - ▶ push tests on features we rely on
- Staying as close as possible from upstream
 - ▶ avoid as much as possible carrying patches
 - ▶ try/test new releases & features
- Switch to PVH2?

Thanks! Questions?

