CentOS Infra revealed

(aka “the joy of running on community donated machines”)

Fabian Arrotin

arrfab@centos.org

/whois arrfab

- Belgian guy
- SysAdmin by choice
- CentOS (ab)user for a long time
- CentOS Project member

Agenda

- history
- Infra overview
- Interaction with other distributions' infra teams

A bit of history

- Once upon a time …

  -
    ```
    whois centos.org|grep Creation
    ```

  -
    ```
    Creation Date: 2003-12-04T12:28:30Z
    ```
- A few private nodes used as builders (idle/spare time)
- 5 hosted/donated machines (all roles)

Once upon a time (back in 2004)

- Local scripts
- Centos guys doing that on their free/spare time after daily jobs
- Need for automation (beginning of puppet inside centos.org infra)

Current infra situation

- Core infra :
  - CentOS 5/6/7
    - Slow migration : remotely reinstall some nodes running previous centos supported distro
    - Mix of new and old hw (Pentium 4 anyone ?)
  - Donated machines all around the world => not the classic "on premise" DC, not centralized

Current infra situation

- cfgmgmt:
  - puppet (coming from puppetmasterd 0.23 to 2.7.23 and then to 3.6.2)
  - switched from “plain” puppetmasterd to puppet+Foreman as ENC (trying to put all data/variables at the Foreman level)

Current infra situation

- Monitoring - Zabbix:
  - agent on all nodes
  - external checks
  - Zabbix proxies (because of geo-dispersed topology)

Current infra situation

- DNS for centos.org :
  - BIND, as first choice (3 glue records), but delegation to powerdns (custom pipe backend)
  - more and more msync nodes ({mirror,msync}.centos.org) =>
    - from external mirror => trying to fetch from nearest msync node
    - from end-user => same (mirror.centos.org)
- PowerDNS = 3 nodes , ~400requests/sec (moved multiple times)

CentOS

Current infra situation

- Other end-users facing roles
  - www
  - Forums
  - Bug tracker (mantis)
  - Mailing-list (mailman)
  - Torrent.centos.org (+ seeders)

Msync/mirror role

- msync/mirror : 60 nodes currently
  - pushing up to 4Gb/s during several hours at release time
  - seeding 580+ external public mirrors (and all end users land on those verified mirrors through yum)
  - A lot in the US, a few in Europe, almost nothing in Asia/AP/Africa
  - Sometimes slowly connected (10Mbps)

Verifying mirrors

- mirror-status : check in loop every mirror/release/arch/iso and produce a “per country” list
- used by mirrorlist.centos.org and isoredirect.centos.org (using GeoIP check for correct redirection)
- ex: `curl 'http://mirrorlist.centos.org/?release=7&arch=x86_64&repo=os&infra=stock&cc=be'`
- same for isoredirect : mirror.centos.org/centos/7/isos/x86_64 => http://isoredirect.centos.org/centos/7/isos/x86_64/

“donated machines” - facts

- Almost all infra running on “donated” machines
- lost several msync nodes (sometimes hardware issue, or something else)
- wiki migrated 3 times on different physical nodes (in 2 months window)
- main ns1.centos.org record moved to another node
- Mailing-list, bug tracker, planet, mirmon/mirror-status.centos.org (main mirror management solution)

“donated machines”

- Challenges:
  - not always aware of such needed move
  - sponsoring company disappearing :
    - bankrupcy
    - acquisition and new owner doesn't want to support OSS/CentOS
    - sometimes machine still run (no inventory ?)
    - sometimes it disappears
      - with/without notification

CentOS

“donated” machines

- Almost all infra running on “donated” machines
- what we do :
  - start by reinstalling it remotely :
    - Faster than an audit : ssh keys, different kernel -ovh- with grsec, etc)
  - then puppetize it
  - Start small (non crucial role)

“donated” machines

- all about trust relationship (to be built/proven over time)
- we start with non crucial role (msync comes to mind :
  - all packages are gpg signed
  - can be removed from pdns array/lists)
- "test" their suppor/response time/quality of response
- for more "crucial" roles, prefer using sponsors who showed they can trusted on a long term
  - example : ns1.centos.org is a VM that can be moved between two physical nodes in the same DC (not that often !)

Other resources (DevCloud)

- For developers/tests
- 4 physical nodes in a DC (64Gb RAM each)
- aggregating local sata disks through gluster + infiniband
- opennebula cloud

The Future

- Centralized auth
  - Local users still defined for Infra team through puppet
  - More users for cbs/koji/SIGs
    - X509 certs needed
    - Self-service portal
- Faster updates on mirrors
  - CDN ?
  - Message bus for communications with external mirrors ?

CentOS

Q&A

Questions ?

Thank you !

CentOS