

Configuration Management Evolution at CERN

Gavin McCance

gavin.mccance@cern.ch

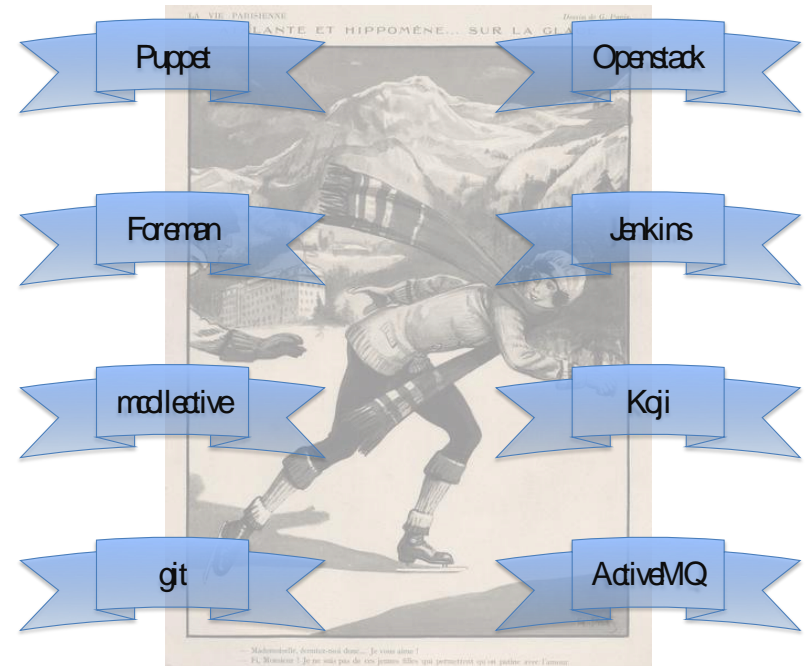
@gmccance

Agile Infrastructure

- Why we changed the stack
- Current status
- Technology challenges
- People challenges
- Community

The Agile Infrastructure

Making IT operations better since 2013



Why?

- Homebrew stack of tools
 - Twice the number machines, no new staff
 - New remote data-centre
 - Adopting more dynamic Cloud model
- “We’re not special”
 - Existence of open source tool chain: OpenStack, Puppet, Foreman, Kibana
- Staff turnover
 - Use standard tools – we can hire for it and people can be hired for it when they leave

Agile Infrastructure “stack”

- Our current stack has been stable for one year now
 - See plenary talk at last CHEP (Tim Bell et al)
- Virtual server provisioning
 - Cloud “operating system”: **OpenStack** -> (Belmiro, next)
- Configuration management
 - **Puppet** + ecosystem as configuration management system
 - **Foreman** as machine inventory tool and dashboard
- Monitoring improvements
 - **Flume + Elasticsearch + Kibana** -> (Pedro, next++)



Puppet

- Puppet manages nodes' configuration via “manifests” written in Puppet DSL
- All nodes check in frequently (~1-2 hours) and ask for configuration
 - Configuration applied frequently to minimise drift
- Using the central puppet master model
 - ..rather than masterless model
 - No shipping of code, central caching and ACLs



Separation of data and code

- Puppet “Hiera” splits configuration “data” from “code”
 - Treat Puppet manifests really as code
 - More reusable manifests
 - Heira is quite new: old manifests are catching up
- Hiera can use multiple sources for lookup
 - Currently we store the data in git
 - Investigating DB for “canned” operations

Modules and Git

- Manifests (code) and hiera (data) are version controlled



- Puppet can use git's easy branching to support parallel environments
 - Later...

Foreman

- Lifecycle management tool for VMs and physical servers



- External Node Classifier – tells the puppet master what a node should look like
- Receives reports from Puppet runs and provides dashboard



Hosts

hostgroup_fullname ~ bi/inter/plus/live

<input type="checkbox"/>	Name	Operating system	Environment	Model	Host group	Last report	Owner	
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0353.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0375.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	42 minutes ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0407.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/acron	about 1 hour ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0409.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 1 hour ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0410.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0411.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 1 hour ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0412.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0416.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0417.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 1 hour ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0418.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0419.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 2 hours ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>
<input type="checkbox"/>	<input checked="" type="checkbox"/> lxplus0420.cern.ch	SLC 6.4	production	OpenStack Nova	bi/inter/plus/live/login	about 1 hour ago	it-dep-pes-ps-support	<input type="button" value="Edit"/>

p01001532021656.cern.ch

Reports from the last days - 163 reports found

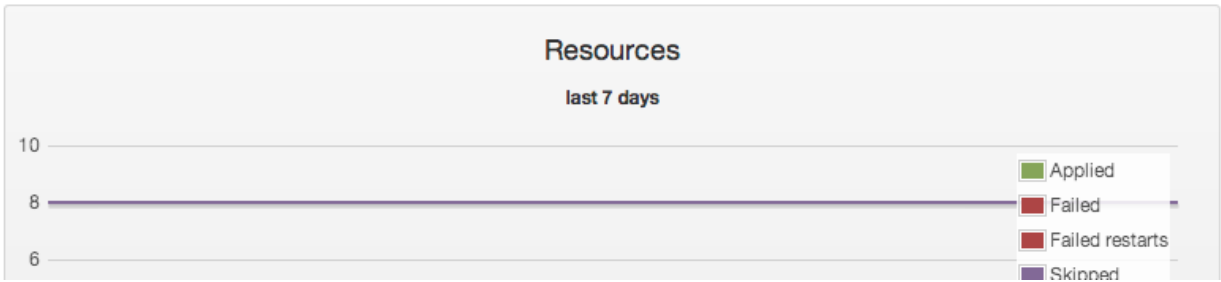
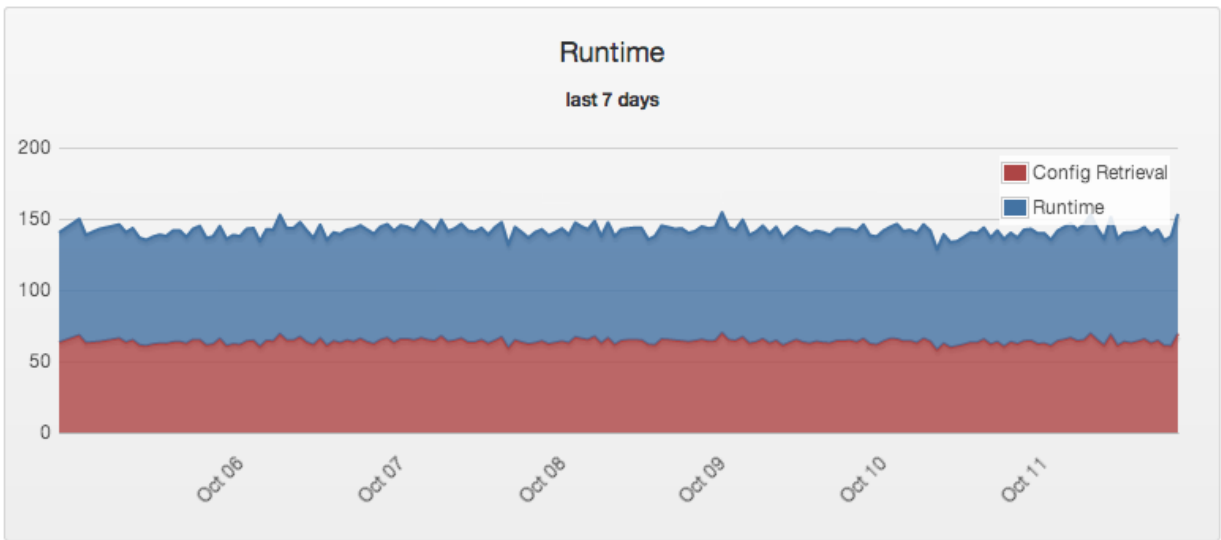
[Edit](#) [Build](#) [Delete](#)

[Properties](#) [Metrics](#) [Templates](#) [BMC](#)

Details

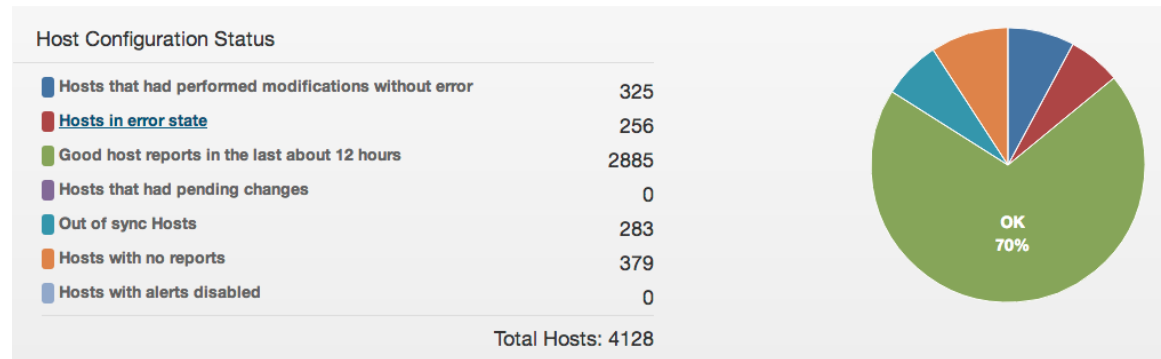
[Audits](#) [Facts](#) [Reports](#) [YAML](#)

Properties	
Remote control	Console
Monitoring	Lemon
Console Username	root@p01001532021656
Console Password	XXXXXXXXXX
Domain	cern.ch
IP Address	128.142.35.220
MAC Address	00:00:00:00:00:00
Puppet Environment	ceph_beesly
Host Architecture	x86_64



Deployment at CERN

- Puppet 3.2
- Foreman 1.2



- Been in real production for 6 months
- Over 4000 hosts currently managed by Puppet
 - SLC5, SLC6, Windows
 - ~100 distinct hostgroups in CERN IT + Expts
 - New EMI Grid service instances puppetised
 - Batch/Lxplus service moving as fast as we can drain it
 - Data services migrating with new capacity
 - AI services (Openstack, Puppet, etc)

Key technical challenges

- Service stability and scaling
- Service monitoring
- Foreman improvements
- Site integration

Scalability experiences

- Most stability issues we had were down to scaling issues
- Puppet masters are easy to load-balance
 - We use standard apache mod_proxy_balancer
 - We currently have 16 masters
 - Fairly high IO and CPU requirements
- Split up services
 - Puppet – critical vs. non critical

12 backend nodes
“Bulk”

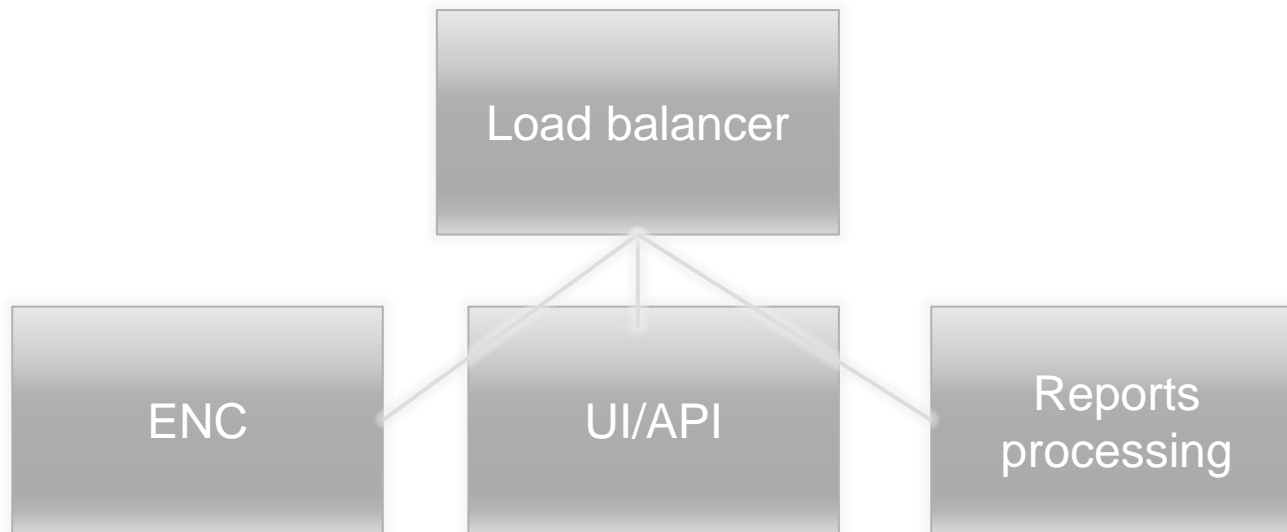
4 backend nodes
“Interactive”

**SERVING LARGE FILES OVER
PUPPET**



Scalability experiences

- Foreman is easy to load-balance
- Also split into different services
 - That way Puppet and Foreman UI don't get affected by e.g. massive installation bursts



PuppetDB

- All puppet data sent to PuppetDB
 - Querying at compile time for Puppet manifests
 - e.g. configure load-balancer for all workers
- Scaling is still a challenge
 - Single instance – manual failover for now
 - Postgres scaling
 - Heavily IO bound (we moved to SSDs)
 - Get the book



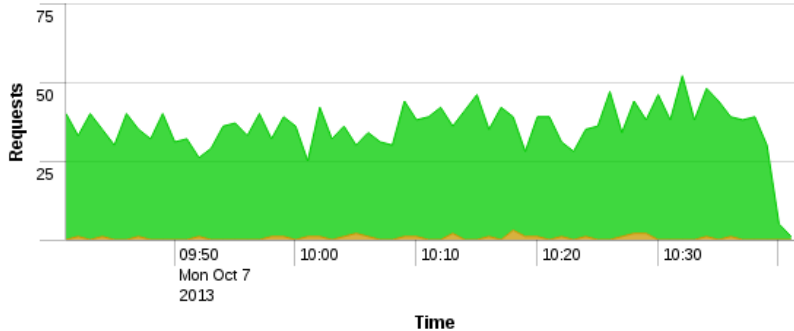
Monitor response times

- Monitor response, errors and identify bottlenecks

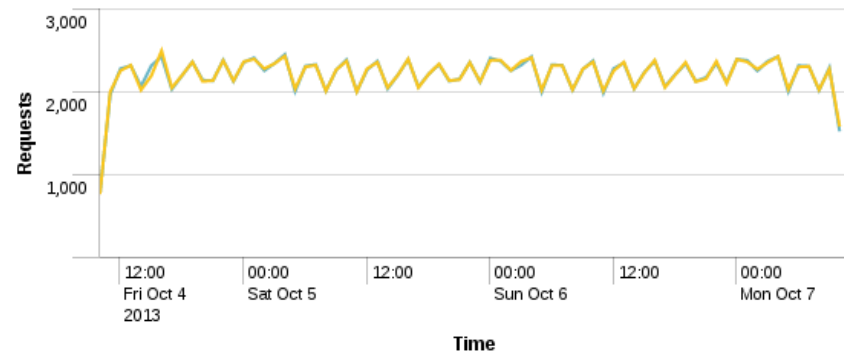
LB average compilation time (inter)



Total catalogs by response code group (batch)



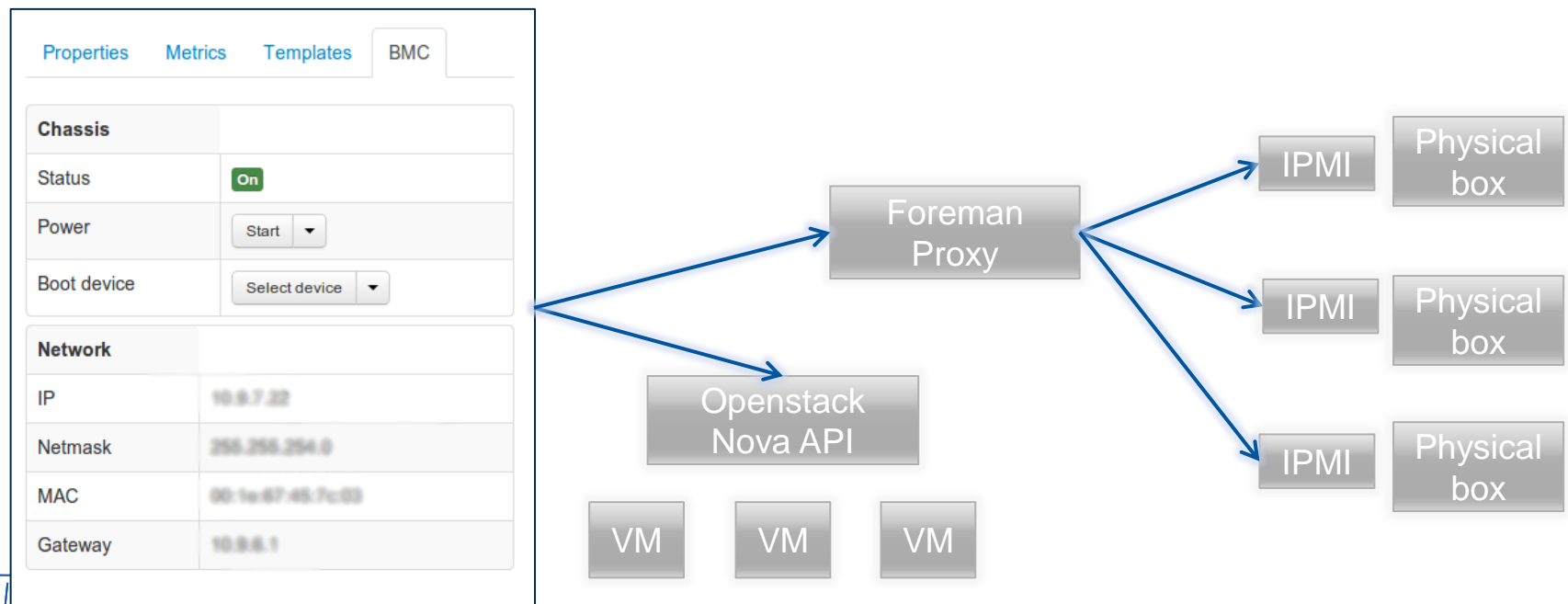
Catalogs and reports (batch)



- Currently using Splunk – will likely migrate to Elasticsearch and Kibana

Upstream improvements

- CERN strategy is to run the main-line upstream code
 - Any developments we do gets pushed upstream
 - e.g Foreman power operations, CVE reported

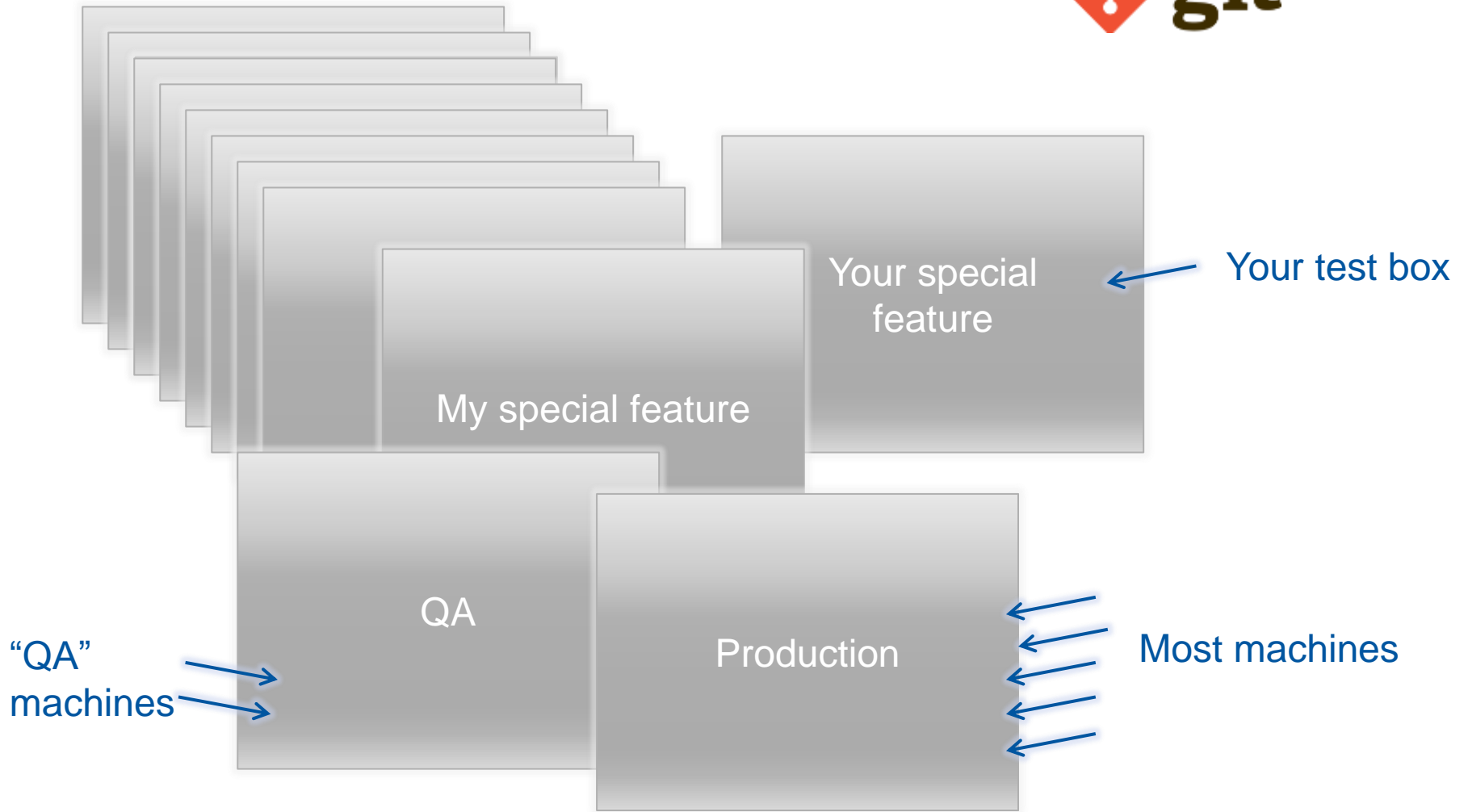


Site integration

- Using Opensource doesn't get *completely* get you away from coding your own stuff
- We've found every time Puppet touches our existing site infrastructure a new "service" or "plugin" is born
 - Implementing our CA audit policy
 - Integrating with our existing PXE setup and burn-in/hardware allocation process - possible convergence on tools in the future – Razor?
 - Implementing Lemon monitoring "masking" use-cases – nothing upstream, yet..

People challenges

- Debugging tools and docu needed!
 - PuppetDB helpful here
- Can we have X', Y' and Z' ?
 - Just because the old thing did it like that, doesn't mean it was the only way to do it
 - Real requirements are interesting to others too
 - Re-understanding requirements and **documentation and training**
- Great tools – how do 150 people use them without stepping on each other?
 - Workflow and process



Use git branches to define isolated puppet environments



```
o [devel] configure elasticsearch endpoint for kibana
o remove lb configuration and flume from teststack
o Adding virtual host for ermis
o condor faster discovery
o Do some clean-up in my hostgroup and create a new on
o condor hg, preemption disabled
o Adding template folder to my hostgroup
o Creating hostgroup for ermis.
o hg_vobox new variable
o extract all fields from nova api requests
o add regex nova api for flume
o kibana configuration data
o add standalone kibana configuration
o Added landb set integration
o Fixed error with defined/undef
o More complete configuration for myproxy
o Updated hiera following changes in myproxy hg/module
o Reorganized myproxy module
o use sssd_filter_users so that values defined in diff
o Add new its6 koji repository
o hg_bi rsyslog hiera server
o remove useless notify statment
o AI-2281 - Open ssh ports so aiadm can access.
o remove setting of rules from puppet on request of se
o add teststak - the openstack test instance
```

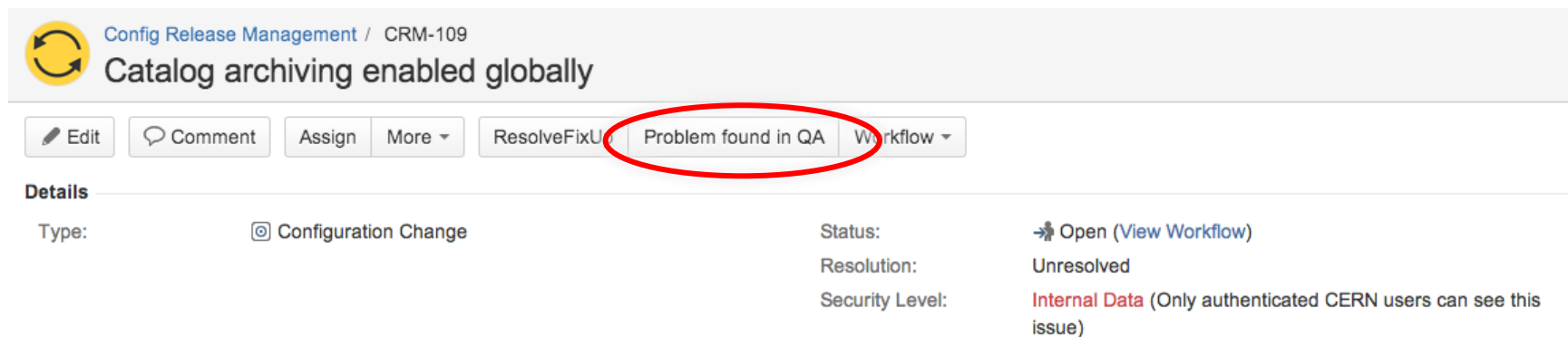
Easy git cherry pick

Git model and flexible environments

- For simplicity we made it more complex
 - Each Puppet module / hostgroup now has its own git repo (~200 in all)
 - Simple git-merge process within module
 - Delegated ACLs to enhance security
 - Standard “QA” and “production” branches that machines can subscribe to
 - Flexible tool (Jens, to be open-sourced by CERN) for defining “feature” developments
 - Everything from “production” except for the change I’m testing on my module

Strong QA process

- Mandatory QA process for “shared” modules
 - Recommended for non-shared modules
 - Everyone is expected to have some nodes from their service in the QA environment
 - Normally changes are QA'd for at least 1 week. Hit the button if it breaks your box!



Config Release Management / CRM-109
Catalog archiving enabled globally

Edit Comment Assign More ResolveFixU **Problem found in QA** Workflow

Details

Type: Configuration Change

Status: Open (View Workflow)

Resolution: Unresolved

Security Level: Internal Data (Only authenticated CERN users can see this issue)

- Still iterating on the process
 - Not bound by technology
 - Is one week enough? Can people “freeze”?

Community collaboration

- Traditionally one of HEPs strong points
- There's a large existing Puppet community with a good model - we can join it and open-source our modules
- New HEPiX working group being formed now
 - Engage with existing Puppet community
 - Advice on best practices
 - Common modules for HEP/Grid-specific software
 - <https://twiki.cern.ch/twiki/bin/view/HEPIX/ConfigManagement>
 - <https://lists.desy.de/sympa/info/hepix-config-wg>

The screenshot shows the GitHub profile for 'cernops'. The profile includes the CERN logo, the name 'CERN Operations cernops', location 'CERN, Meyrin, Switzerland', website 'https://agileinf.its.cern.ch/jira', and join date 'Joined on May 21, 2012'. Statistics show 34 public repos, 0 private repos, and 20 members. A list of repositories is displayed, including:

- puppet-fetchcrl**: Installs and Configures fetch-crl along with IGTf certificates. Last updated a day ago. 1 star, 2 forks.
- voms-admin-server**: forked from italiangrid/voms-admin-server. The VOMS Administration service. Last updated a day ago. 0 stars, 2 forks.
- puppet-fts**: Puppet module for File Transfer Service (FTS). Last updated 2 days ago. 1 star, 0 forks.
- puppet-glexecwn**: EMI glExec and worker node module. Last updated 5 days ago. 0 stars, 2 forks.
- puppet-bdii**: EMI BDIi. Last updated 5 days ago. 0 stars, 1 fork.
- puppet-argus**: EMI Argus module. Last updated 5 days ago. 0 stars, 2 forks.
- puppet-creamce**: puppet module to install and configure a cream CE (EMI3). 0 stars, 0 forks.

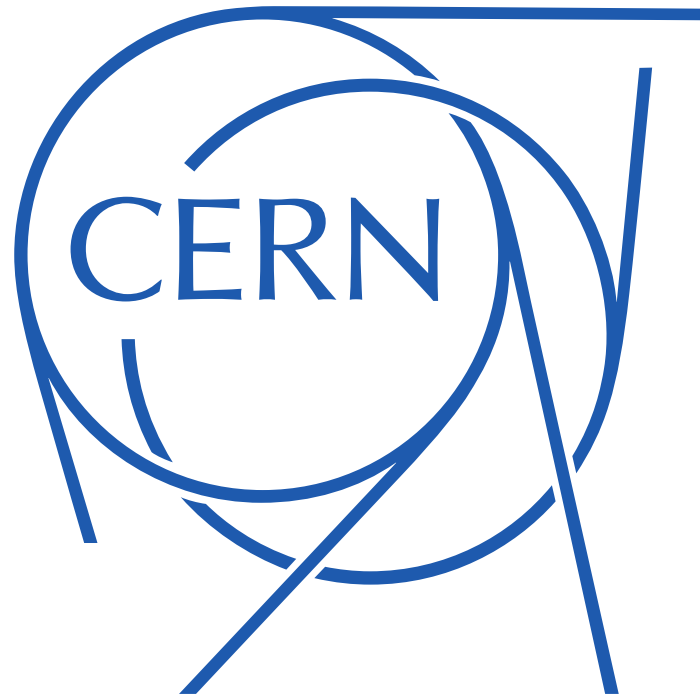
<http://github.com/cernops>
 for the modules we share
Pull requests welcome!



Summary

- The Puppet / Foreman / Git / Openstack model is working well for us
 - 4000 hosts in production, migration ongoing
- Key technical challenges are scaling and integration which are under control
- Main challenge now is people and process
 - How to maximise the utility of the tools
- The HEP and Puppet communities are both strong and we can benefit if we join them together

<https://twiki.cern.ch/twiki/bin/view/HEPIX/ConfigManagement>
<http://github.com/cernops>



Backup slides

