Building a Puppet Infrastructure @ DESY

Jan Engels, DESY IT - Systems
HEPiX Fall 2013
University of Michigan

HELMHOLTZ ASSOCIATION

DESY

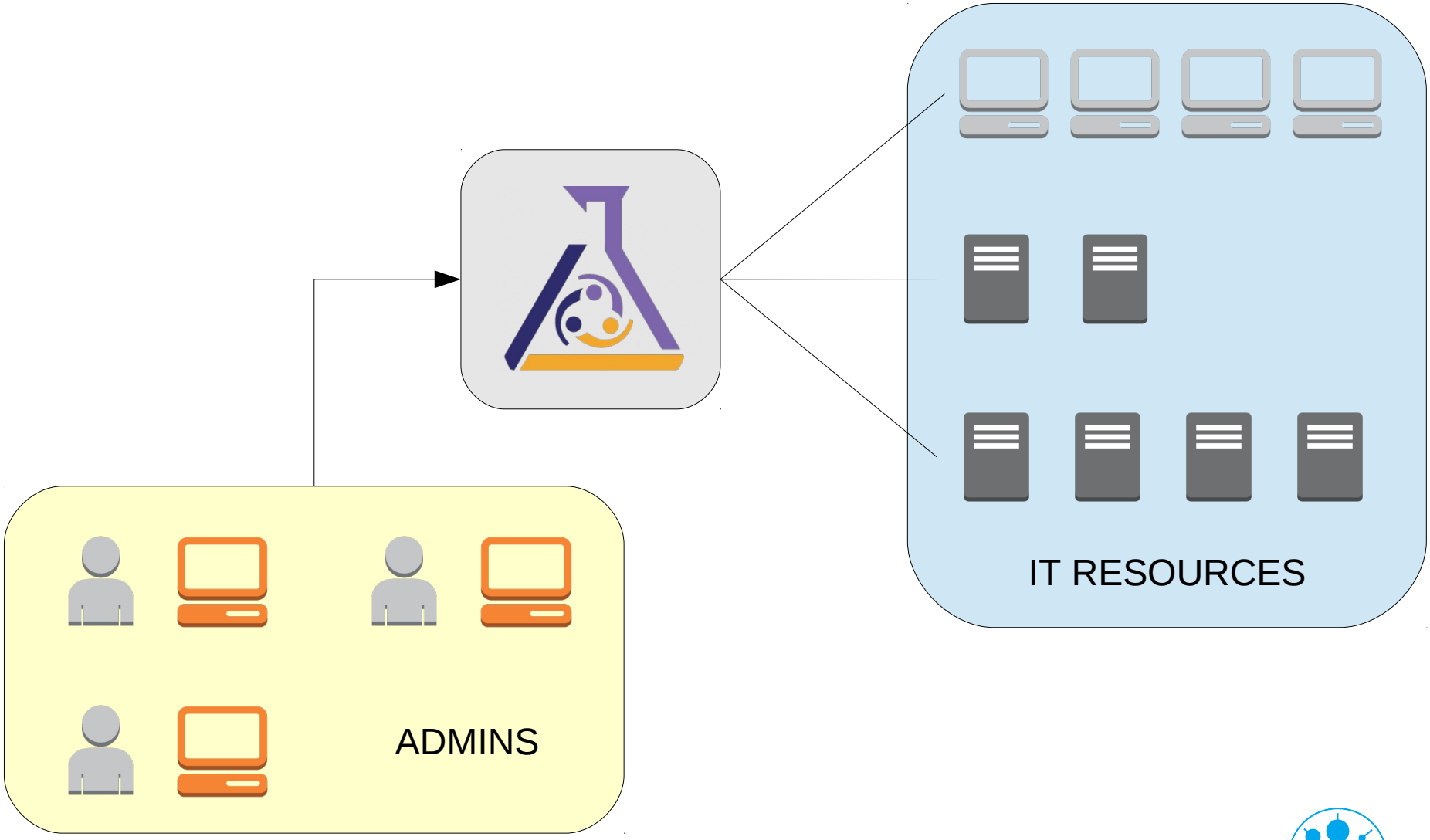
Outline

- Puppet Introduction
- Puppet Infrastructure @ DESY
- Puppet Configuration @ DESY
- Summary & Outlook

Puppet Introduction

> What is Puppet ?

Puppet Introduction



Puppet Introduction

- Configuration management tool
- Modular
  - Third party modules
    - https://forge.puppetlabs.com
- Customizable
  - Write your own modules
- Secure (Standard SSL, HTTPS)
- Cross-platform (Linux, Unix, Windows)
- Open Source – Apache 2.0 license
- Commercial – Puppet Enterprise

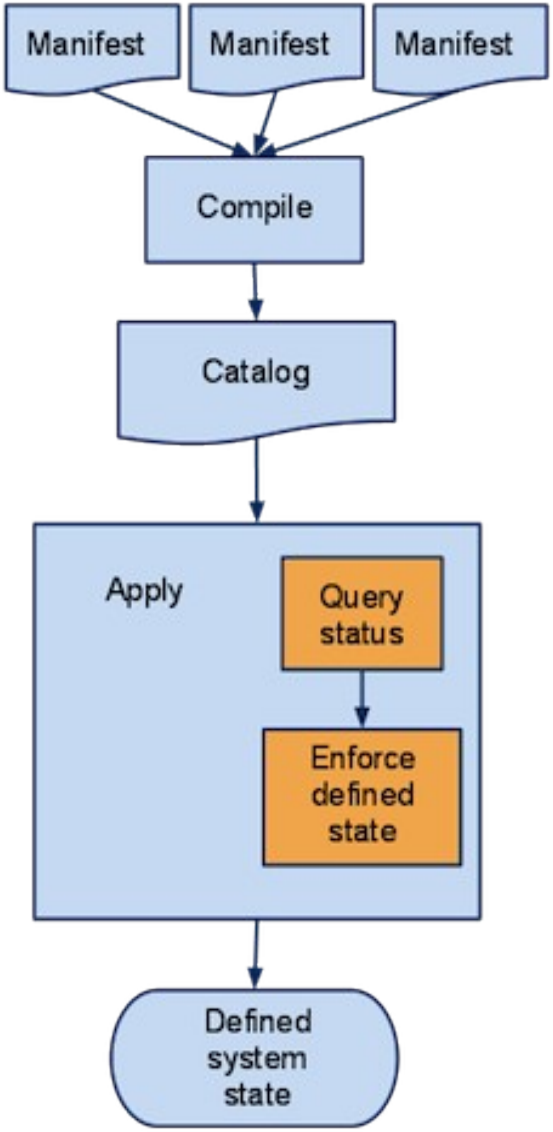
Puppet Introduction

- Manifests
- Declarative syntax
- Define system state
- Resources
- RAL
- Providers
- Facts
- Catalog
- Master/Agent scenario
- Reports

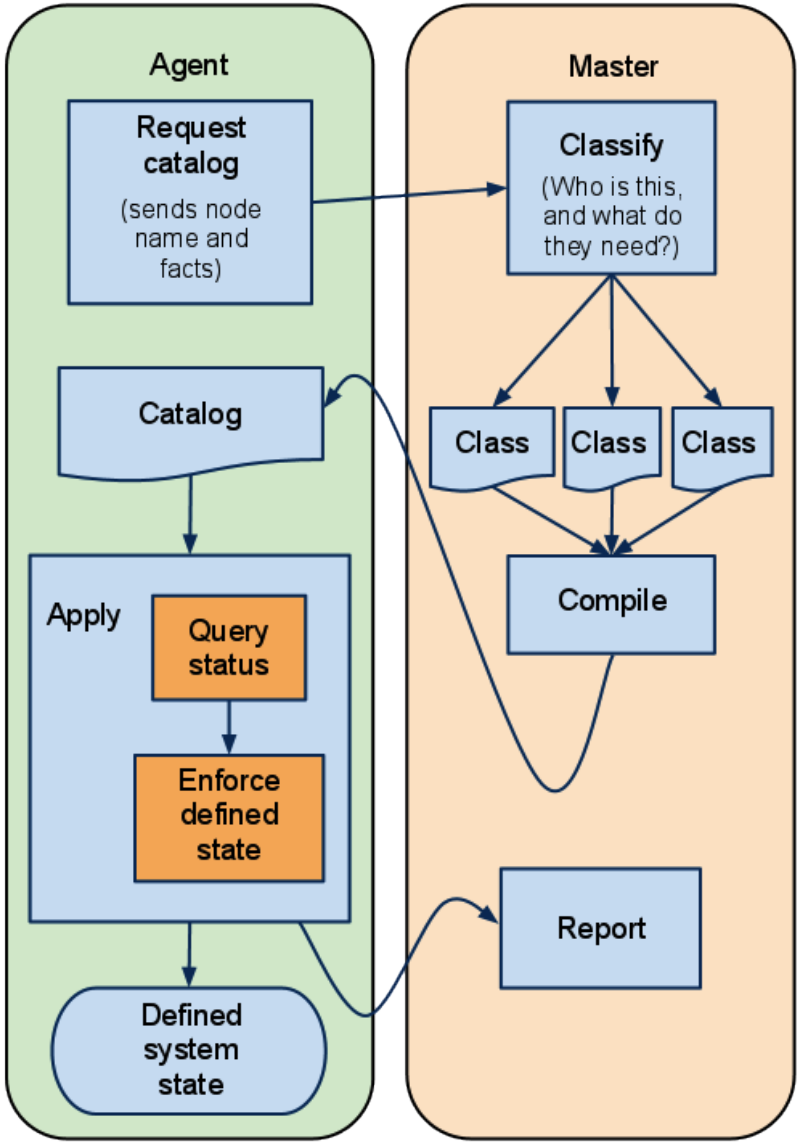
```
# manifest example

package {'htop':
    ensure => present,
}

file { '/tmp/testfile':
    ensure  => file,
    content => "Hi there!",
}

user { 'dave':
    ensure     => present, # absent
    shell      => '/bin/bash',
    home       => '/home/dave',
    managehome => true,
}
```


DESY

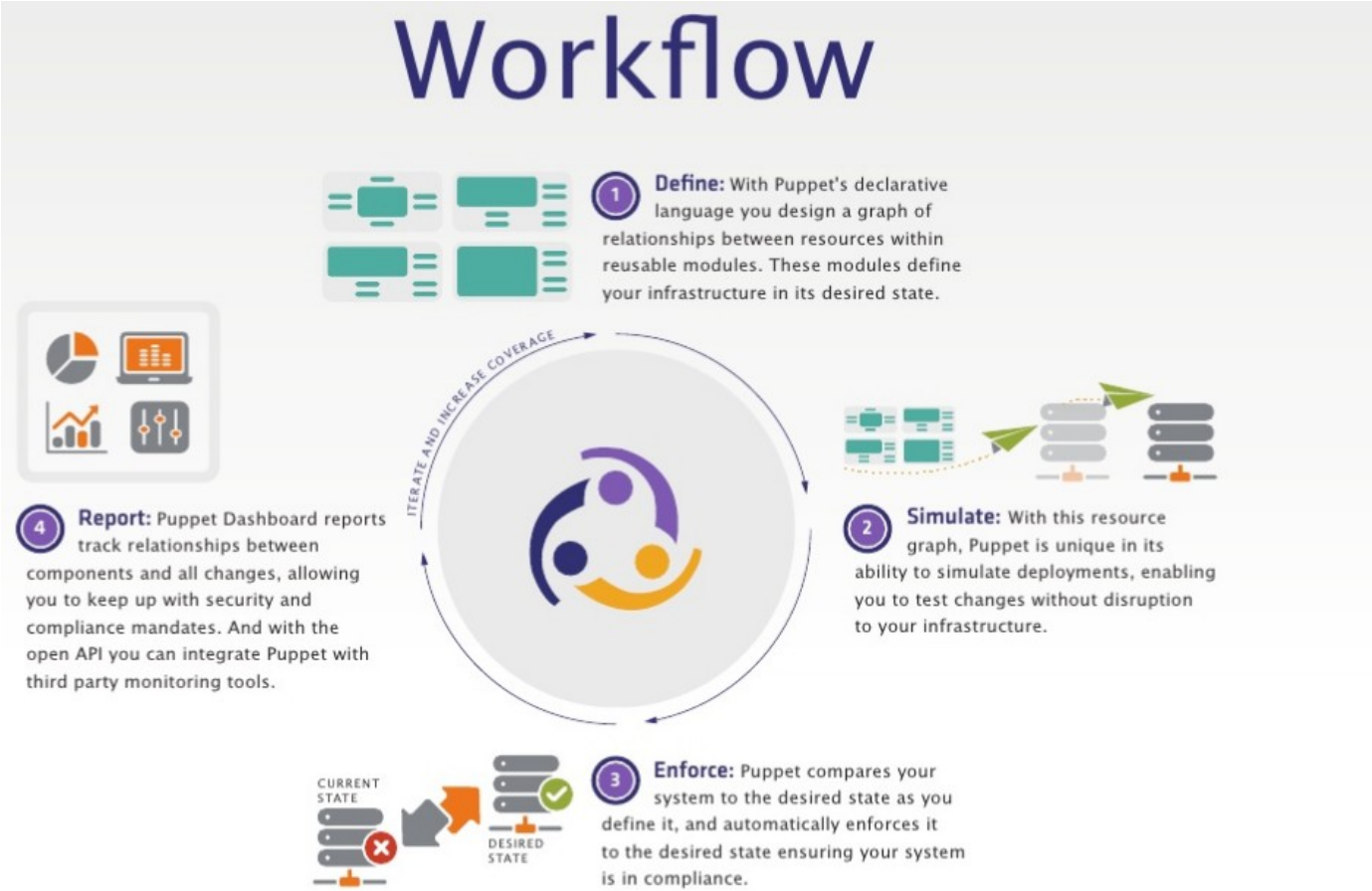
Puppet Introduction



Puppet Introduction



Puppet Introduction



puppet labs

DESY

Puppet Infrastructure @ DESY

Why Puppet?

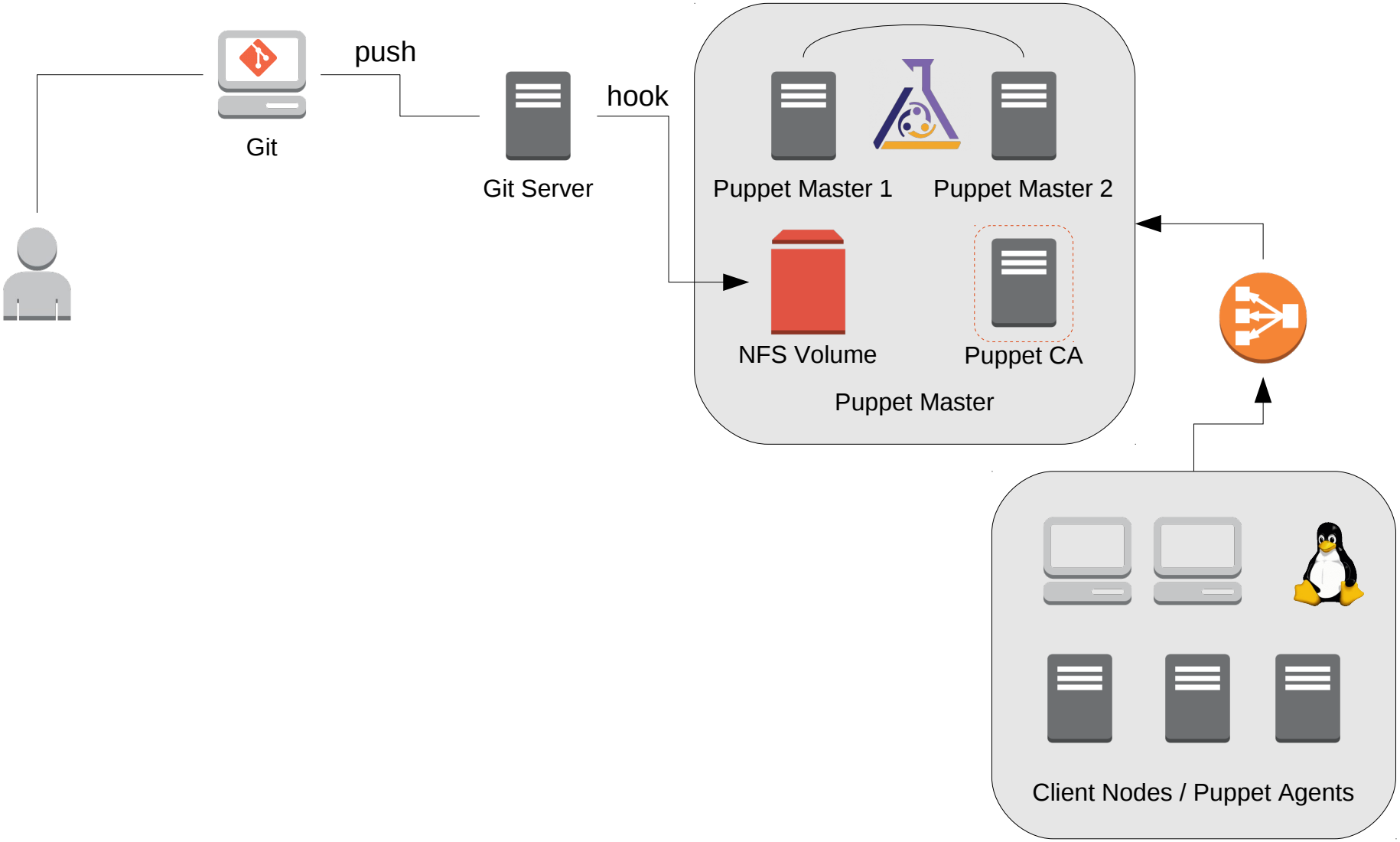
- Quattor
  - Only RedHat support
- FAI
  - Poor RedHat support
- Salad
  - Homegrown, to be retired
- Puppet
  - Cross-platform support
  - Third party modules
  - Defined system state
  - Queriable infrastructure
- Other tools like Chef also considered...
- Puppet seems to be the right choice!

DESY

Initial Puppet Setup @ DESY

- Puppet project started in 2012
- 2 Puppet Master Server (2-way 32-core / 64Gb RAM)
- 1 Git Server (2-way 16-core / 12Gb RAM)
- 1 Puppet Certificate Authority (CA)
- NFS 3 Volume used as shared storage (NetApp)
- F5 Load balancer
  - Single endpoint for Puppet Agents
  - Redirects agents with a Certificate Signing Request (CSR) to the CA

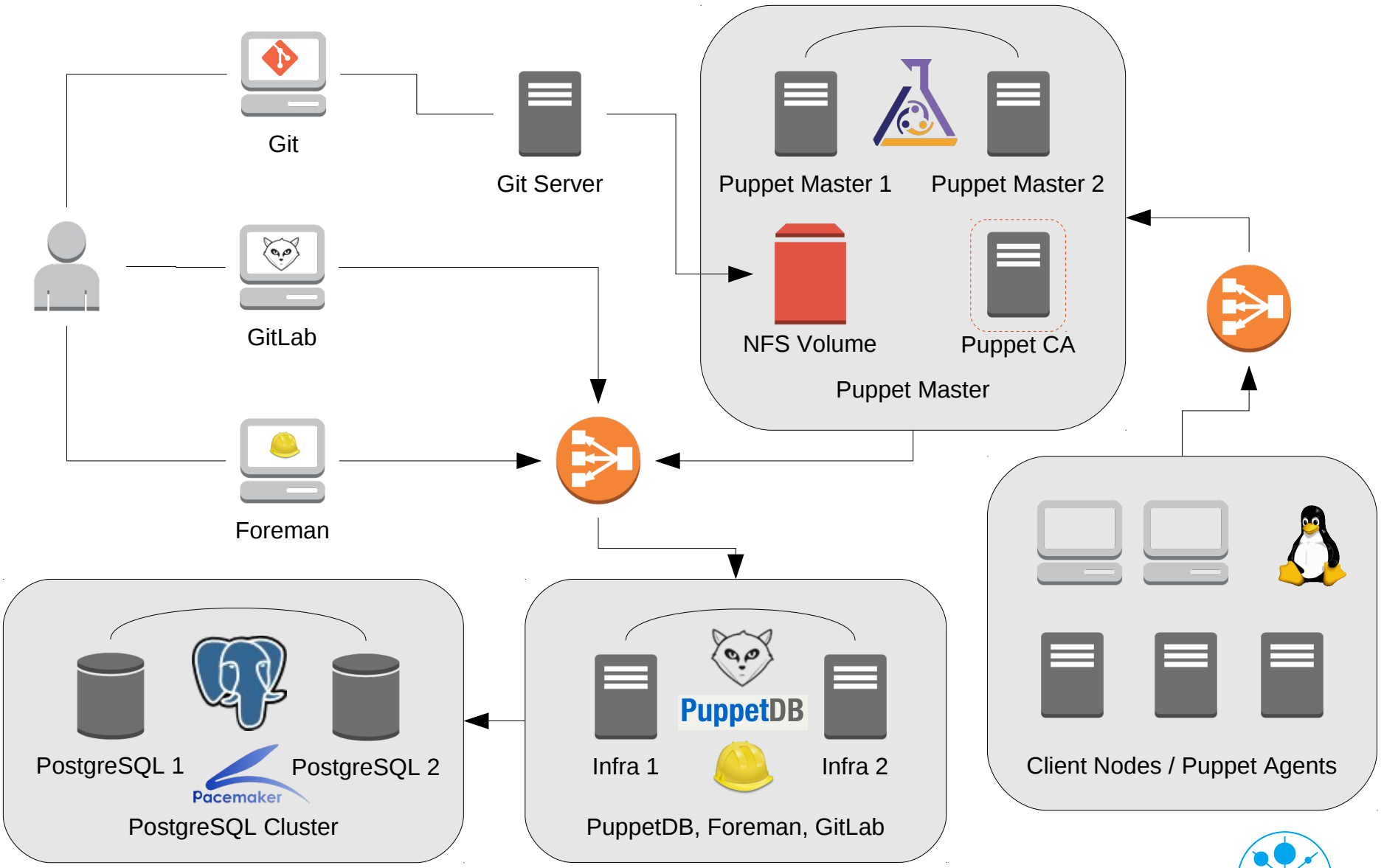
Initial Puppet Setup @ DESY



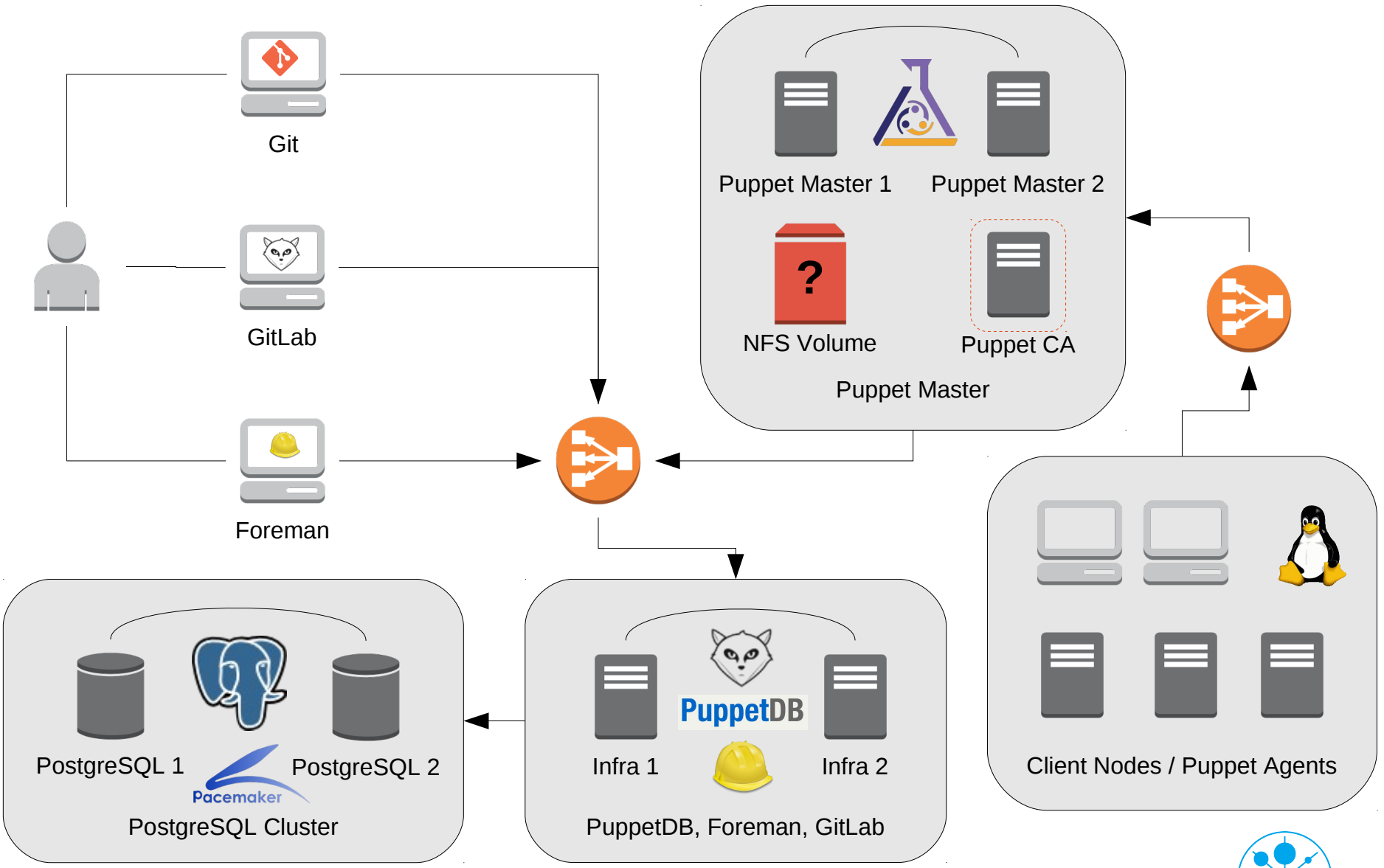
Current Puppet Infrastructure @ DESY

- 2 Puppet Master Server (2-way 32-core / 64Gb RAM)
- 2 PostgreSQL DB Server (2-way 24-core / 64Gb RAM)
- 2 PuppetDB/Gitlab/Foreman Server (2-way 24-core / 64Gb RAM)
- 1 Git Server (2-way 16-core / 12Gb RAM)
- 1 Puppet Certificate Authority (CA)
- NFS 3 Volume used as shared storage (NetApp)
- F5 Load balancer
  - Single endpoint for Puppet Agents
  - Redirects agents with a Certificate Signing Request (CSR) to the CA

Current Puppet Infrastructure @ DESY



Future Puppet Infrastructure @ DESY



Puppet Configuration @ DESY

- Puppet Development Platform
- Development vs Production
- Bootstrapping
- Host grouping
- Repository and Package Management
- Node Management
- Secrets Management

Puppet Development Platform

- Revision Control with Git
  - Best suited for our Puppet environments scenario
- Documentation
  - Wiki style
- Issue Tracking
- Code Reviewing
  - What goes into the “Master” ?
- Collaboration platform

DESY

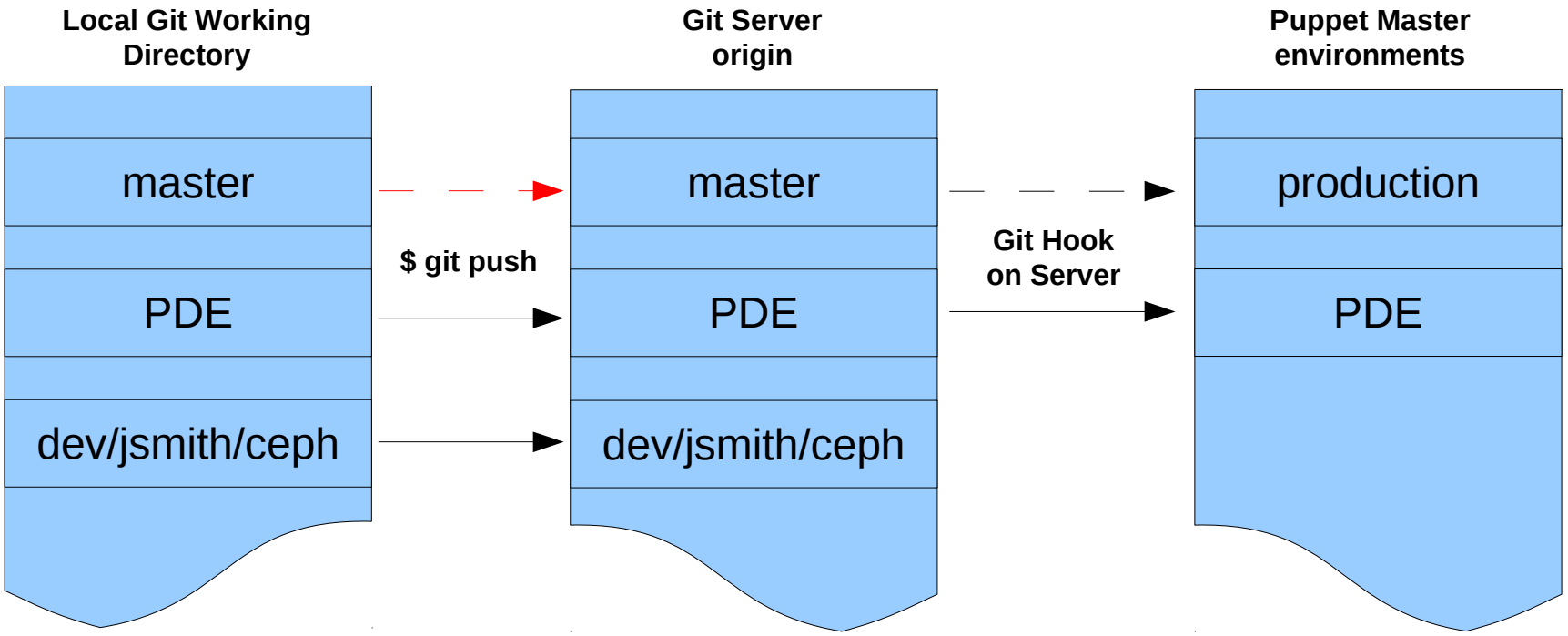
Puppet Development Platform

Features	Trac	Redmine	GitLab	Gerrit	ReviewBoard
Wiki	X	X	**X**	-	-
Git Read Access	X	X	**X**	X	X
Git Write Access	-	-	**X**	X	-
Git ACLs	-	-	**0,5**	X	-
Issue Tracker	X	X	**X**	-	-
Forum	-	X	**-**	-	-
Comments	(?)	X	**X**	X	X
Workflow Mgmt.	-	-	**X**	X	0,5
Code Review	(?)	0,5	**X**	X	X
Usability	(?)	-	**X**	(?)	X
Maintenance	X	0,5	**-**	0,8	(?)
DESY Integration	X	0,5	**X**	X	(?)
Result	**5**	**6,5**	**9,5**	**7,8**	**4,5**

DESY

Development vs Production

- Git hook dynamically updates puppet environments
- Git branch matching the user id used as “Personal Development Environment”
- Deploy step for additional layer of “security” on production environment(s)



Bootstrapping

> WBOOM (Workstation and BOOt Management – T. Finnern)

- Host registration
- Anaconda kickstart / FAI
- Network registration (PXE, DNS, DHCP, IP)
- Generate puppet node definitions

> After bootstrapping control is handed over to puppet

Host Grouping

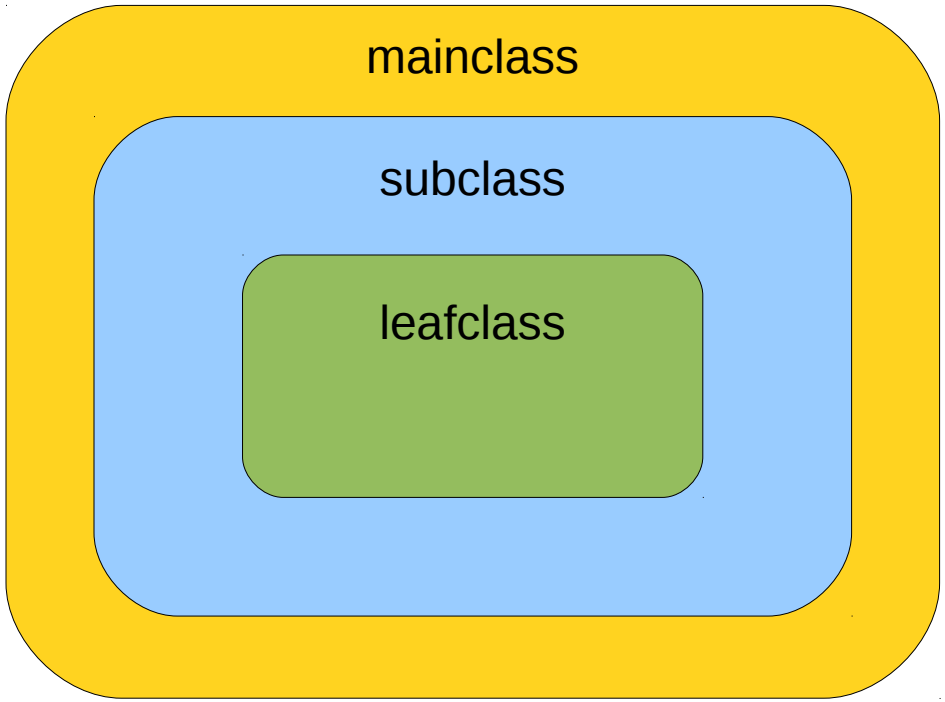
> Hierarchical approach

- Generic classes used for classifying hosts
- Some examples:

> mainclass = grid
> subclass = wn

> mainclass = grid
> subclass = server

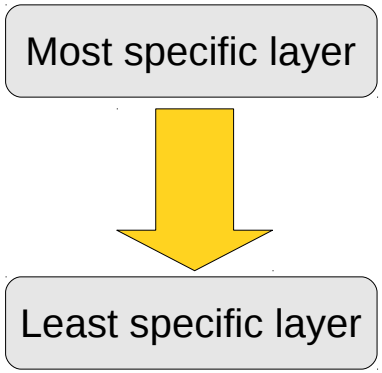
> mainclass = it
> subclass = desktop



Host Grouping

> Hiera Configuration

```
---
:backends:
    - yaml
:yaml:
    :datadir: '/srv/puppet/environments/%{environment}/hieradata'
:hierarchy:
    - "%{::clientcert}"
    - "%{::mainclass}_%{::subclass}_%{::leafclass}"
    - "%{::mainclass}_%{::subclass}"
    - "%{::mainclass}"
    - "%{::operatingsystem}_%{::lsbdistrelease}"
    - "%{::operatingsystem}_%{::lsbmajdistrelease}"
    - "%{::operatingsystem}"
    - common
```



Repository and Package Management

- Mirror external repositories
- Yum
  - Yum exclude to freeze package versions
  - Priorities/Protected
  - Added repositories stage in Puppet
  - Considering using Pulp for user managed repositories

Node Management

- Node selection based on Puppet facts
- Trigger package updates
- Reboot machines
- Evaluating MCollective
  - Middleware configuration
  - Parallel-ssh still needed for now

Secrets Management

- DESY Tool: HKDC
  - Stores/retrieves secrets on a “per-node” basis
  - OK for host certificates
  - Not enough as a general purpose “Secrets Management Tool”
- Hiera encryption backends
  - Not suitable for our scenario
- Work in progress...

Summary & Outlook

> Summary

- Puppet running in production since summer 2013
- Managing ~500 Linux clients (SL6, Ubuntu 12.04)
- 1st milestone reached: Grid!
- Continuously moving existing services into puppet
- Working on MCollective and GitLab

> Outlook

- MCollective in production
- Foreman ENC
- Secrets Management Tool
- Preparing first release on Puppet Forge!

Thank you! Questions or comments?

DESY

References

- Facter
  - http://puppetlabs.com/facter
- FAI
  - http://fai-project.org/
- Foreman
  - http://theforeman.org/
- Git
  - http://git-scm.com/
- GitLab
  - http://gitlab.org/
- Grid
  - http://grid.desy.de/

References

- Hiera
  - http://projects.puppetlabs.com/projects/hiera/
- MCollective
  - http://puppetlabs.com/mcollective
- Pacemaker
  - http://clusterlabs.org/wiki/Main_Page
- Parallel-ssh
  - http://code.google.com/p/parallel-ssh/
- PostgreSQL
  - http://www.postgresql.org/
- Pulp
  - http://www.pulpproject.org/

References

- Puppet Labs
  - http://puppetlabs.com/
- Puppet Forge
  - https://forge.puppetlabs.com/
- PuppetDB
  - http://docs.puppetlabs.com/puppetdb/index.html
- Quattor
  - http://quattor.org/
- WBOOM / Salad
  - http://it.desy.de/services/administration/wboom_device_management_system/index_eng.htr