

# Puppet for Dumps maintainers

V 0.2

# 50 (+10) minutes of Puppet

- Why should you care?
- What is Puppet?
- Capabilities
- Approach to code
- Language
- Syntax and style
- Default variable values (Hiera)
- Facter
- Classes
- More classes
- Roles, profiles, modules
- Testing with PCC (puppet compiler)
- Testing in deployment-prep
- Merging changes
- Effects of a change
- Disabling puppet
- The private repo
- Q&A

# Why learn puppet?

- Everything about the dumps is in puppet:
- Cron jobs for xml/sql dumps
- Cron jobs for all other dumps
- Exception watcher
- Runtime stats gatherer
- Dump FAQ emailer
- Etc!

# What is Puppet anyways?

- A configuration and orchestration tool
- Manage multiple servers from one host
- PKI for all servers
- All changes to files on a server are detailed with old copies of files preserved



# What Puppet can do

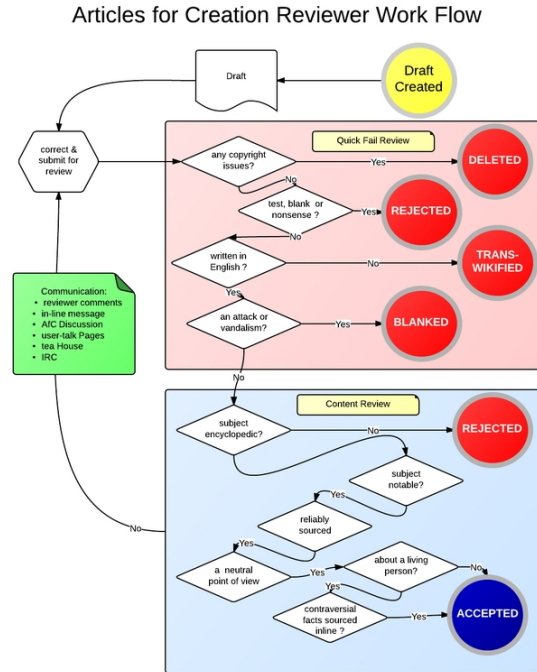
- Create configuration files
- Install packages
- Set up cron jobs and systemd units
- Start and stop services
- Set up firewall rules
- Create users and add their ssh keys
- Make, export or mount NFS shares
- And much much more



Coordinating many systems harmoniously

# Puppet uses a declarative approach

- You declare the state you want your system(s) to be in
- Puppet determines how to get there
- Order of execution of statements in your code is not a given!
- “before”, “after”, “require”
- Example: make sure a package required by some service is installed before the service is started



It's not quite this bad. Status: it's complicated.

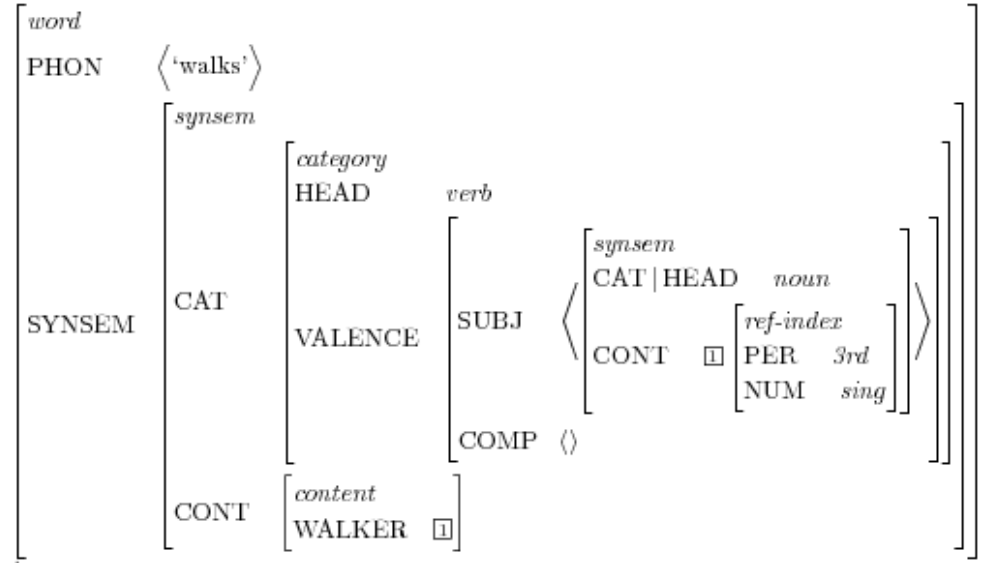
# Puppet language

- Ruby-like for manifests
- Snippets of Ruby in templates
- Full Ruby in library functions
- You don't have to learn Ruby (I didn't!)



# Syntax, linting

- puppet parser validate
- puppet-lint
- quotes and variables in strings
- arrow spacing
- 4-space indentation
- classes, profiles, roles... later.





# Arrows

- [http://puppet-lint.com/checks/arrow\\_alignment/](http://puppet-lint.com/checks/arrow_alignment/)

```
file { $tempdir:  
  ensure => 'directory',  
  mode   => '0755',  
  owner  => $user,  
  group => $group,  
}
```

**NOPE**

```
file { $tempdir:  
  ensure => 'directory',  
  mode   => '0755',  
  owner  => $user,  
  group => $group,  
}
```

**YEP**

# Single quotes with vars

- [http://puppet-lint.com/checks/single\\_quote\\_string\\_with\\_variables/](http://puppet-lint.com/checks/single_quote_string_with_variables/)

**NOPE**

```
$cirrussearchdir = '${basedir}/cirrussearch'
```

**YEP**

```
$cirrussearchdir = "${basedir}/cirrussearch"
```

# Quoting lone vars

- [http://puppet-lint.com/checks/only\\_variable\\_string/](http://puppet-lint.com/checks/only_variable_string/)

```
file { $tempdir:  
  ensure => 'directory',  
  mode   => '0755',  
  owner  => "${user}",  
  group  => "${group}",  
}
```

**NOPE**

```
file { $tempdir:  
  ensure => 'directory',  
  mode   => '0755',  
  owner  => $user,  
  group  => $group,  
}
```

**YEP**

# Hiera

- Set values for vars for one or a group of servers
- Puppet/hieradata
- Yaml syntax
- Scope explicit in name

`puppet / hieradata / eqiad / profile / ceph.yaml`

8 lines (7 sloc) | 296 Bytes

```
1  profile::ceph::fsid: '5917e6d9-06a0-4928-827a-f489384975b1'
2  profile::ceph::cluster_network: 192.168.4.0/22
3  profile::ceph::public_network: 10.64.20.0/24
4
5  profile::ceph::openstack_controllers:
6    - cloudcontrol1003.wikimedia.org
7    - cloudcontrol1004.wikimedia.org
8    - cloudcontrol1005.wikimedia.org
```

Restrictions on use, info about lookup scope:

[https://wikitech.wikimedia.org/wiki/Puppet\\_coding#Hiera](https://wikitech.wikimedia.org/wiki/Puppet_coding#Hiera)

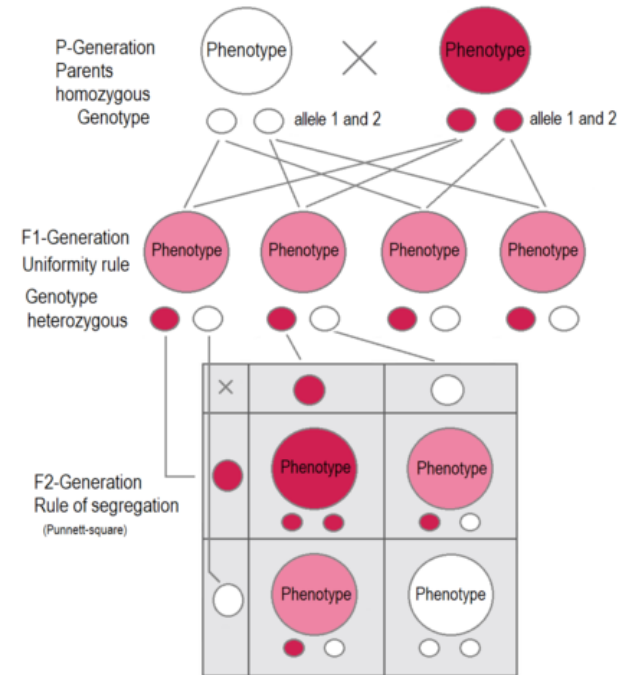
# Facter

- For “facts” retrievable via OS
- OS distro/version
- Cores, memory
- Mounted filesystems
- Networking info
- Custom facts!

```
disks => {  
  sda => {  
    model => "TOSHIBA-TR150",  
    size => "894.25 GiB",  
    size_bytes => 960197124096,  
    vendor => "ATA"  
  },  
  sr0 => {  
    model => "DVDRAM GUC0N",  
    size => "1.00 GiB",  
    size_bytes => 1073741312,  
    vendor => "HL-DT-ST"  
  }  
}  
dmi => {  
  bios => {  
    release_date => "01/16/2015",  
    vendor => "American Megatrends Inc.",  
    version => "N551JX.202"  
  },  
  board => {  
    asset_tag => "ATN12345678901234567",  
    manufacturer => "ASUSTeK COMPUTER INC.",  
    product => "N551JX"  
  },  
}
```

# Classes and inheritance

- Short version:  
don't.
- Long version:  
still don't.



Even Puppet says don't:

[https://puppet.com/docs/puppet/5.5/lang\\_classes.html#inheritance](https://puppet.com/docs/puppet/5.5/lang_classes.html#inheritance)

# Classes as building blocks

```
1 class dumps::generation::server::dirs(
2     $datadir          = undef,
3     $xmldumpsdir     = undef,
4     $tempdir         = undef,
5     $miscdatasetsdir = undef,
6     $user            = undef,
7     $group           = undef,
8 ) {
9     class { 'dumps::server_dirs':
10         datadir          => $datadir,
11         xmldumpsdir     => $xmldumpsdir,
12         miscdatasetsdir => $miscdatasetsdir,
13         user            => $user,
14         group           => $group,
15     }
16
17     # Directories where dumps of any type are generated
18     # This list is not for one-off directories, nor for
19     # directories with incoming rsyncs of datasets
20     $cirrussearchdir    = "${miscdatasetsdir}/cirrussearch"
21     $xlationdir        = "${miscdatasetsdir}/contenttranslation"
22     $categoriesrdfdir  = "${miscdatasetsdir}/categoriesrdf"
23     $categoriesrdfdailydir = "${miscdatasetsdir}/categoriesrdf/daily"
```

```
35     # top level directories for various dumps/datasets, on generation hosts only
36     file { $tempdir:
37         ensure => 'directory',
38         mode   => '0755',
39         owner  => $user,
40         group  => $group,
41     }
42
43     # subdirs for various generated dumps
44     file { [ $cirrussearchdir, $xlationdir, $categoriesrdfdir,
45             $categoriesrdfdailydir, $globalblocksdir, $medialistsdir, $incrsdir,
46             $mediatitlesdir, $pagetitlesdir, $shorturlsdir, $machinevisiondir ]:
47
48         ensure => 'directory',
49         mode   => '0755',
50         owner  => $user,
51         group  => $group,
52     }
53
54     # needed for wikidata weekly crons
55     file { [ $otherwikibasedir, $otherwikibasewikidatadir, $otherwikidatadir ]:
56         ensure => 'directory',
57         mode   => '0755',
58         owner  => $user,
59         group  => $group,
60     }
61 }
```

<https://github.com/wikimedia/puppet/blob/production/modules/dumps/manifests/generation/server/dirs.pp>

# Classes are built from other classes

- ...but not from profiles or roles
- ...and only from classes in the same module
- Profiles are built from classes in multiple modules
- Roles are built from multiple profiles





# Building with classes

- Include: no params, as often as you want
- Require: no params, as often as you want
- Declare: params, only once



Yes, this is built entirely from Legos.

# Resources

- Files, cron jobs, services, packages
- Declare any given resource only once
- `require_package()` to save headaches

```
44     host.evaluate_code(cls)
45   rescue
46     nil
47   end
48
49   # Create package resource
50   begin
51     host_scope = compiler.topscope.class_scope(host)
52     host_scope.call_function(:create_resources,
53                             ['package', { package_name => { :ensure => 'present' } }])
54   rescue Puppet::Resource::Catalog::DuplicateResourceError
55     nil
56   end
57
58   # Declare dependency
59   call_function :require, [class_name]
60 end
```

[https://github.com/wikimedia/puppet/blob/production/modules/wmflib/lib/puppet/parser/functions/require\\_package.rb](https://github.com/wikimedia/puppet/blob/production/modules/wmflib/lib/puppet/parser/functions/require_package.rb)

# Cron jobs as resources


```
1 class snapshot::cron::pagetitles(  
2     $user      = undef,  
3     $filesonly = false,  
4 ) {  
5     $cronsdire = $snapshot::dumps::dirs::cronsdire  
6     $reporidire = $snapshot::dumps::dirs::reporidire  
7     $confsdire = $snapshot::dumps::dirs::confsdire  
8  
9     if !$filesonly {  
10        cron { 'pagetitles-ns0':  
11            ensure      => 'present',  
12            environment => 'MAILTO=ops-dumps@wikimedia.org',  
13            user        => $user,  
14            command     => "cd ${reporidire}; python3 onallwikis.py --configfile ${confsdire}/wikidump.conf.dumps:monitor --filena  
15            minute      => '10',  
16            hour        => '8',  
17        }  
18  
19        cron { 'pagetitles-ns6':  
20            ensure      => 'present',  
21            environment => 'MAILTO=ops-dumps@wikimedia.org',
```

<https://github.com/wikimedia/puppet/blob/production/modules/snapshot/manifests/cron/pagetitles.pp>


# Roles, profiles, classes

- Classes build other classes and profiles
- Profiles build other profiles and roles
- Roles don't build anything
- ONE ROLE PER HOST

 beta\_testbed.pp

 dumper.pp

 dumper\_misc\_crons\_only.pp

 dumper\_monitor.pp

 testbed.pp

# Role for xml/sql dumps workers

12 lines (10 sloc) | 410 Bytes

```
1  class role::dumps::generation::worker::dumper {
2      include ::profile::standard
3      include ::profile::base::firewall
4
5      include profile::dumps::generation::worker::common
6      include profile::dumps::generation::worker::dumper
7      include profile::dumps::generation::worker::crontester
8
9      system::role { 'dumps::generation::worker::dumper':
10         description => 'dumper of XML/SQL wiki content',
11     }
12 }
```

<https://github.com/wikimedia/puppet/blob/production/modules/role/manifests/dumps/generation/worker/dumper.pp>

# Profile common to dumps workers

```
10 # mw packages and dependencies
11 require profile::mediawiki::scap_proxy
12 require profile::mediawiki::common
13 require profile::mediawiki::nutcracker
14 class { 'profile::mediawiki::mcrouter_wancache':
15     prometheus_exporter => false
16 }
17 require profile::services_proxy::envoy
18
19 $xmldumpsmount = '/mnt/dumpsdata'
20
21 class { '::dumpsuser': }
22
23 if ($dumps_misc_cronrunner) {
24     $nfs_server = $cron_nfs_server
25 }
26 else {
27     $nfs_server = $dumps_nfs_server
28 }
```

```
41 class { '::snapshot::dumps::dirs':
42     user          => 'dumpsgen',
43     xmldumpsmount => $xmldumpsmount,
44     xmldumpspublicdir => "${xmldumpsmount}/xmldatadumps/public",
45     xmldumpsprivatedir => "${xmldumpsmount}/xmldatadumps/private",
46     dumpstempdir    => "${xmldumpsmount}/xmldatadumps/temp",
47     cronsdir        => "${xmldumpsmount}/otherdumps",
48     apachedir       => '/srv/mediawiki',
49 }
50 class { '::snapshot::dumps': php => $php}
51
52 # scap3 deployment of dump scripts
53 scap::target { 'dumps/dumps':
54     deploy_user => 'dumpsgen',
55     manage_user => false,
56     key_name    => 'dumpsdeploy',
57 }
58 ssh::userkey { 'dumpsgen':
59     content => secret('keyholder/dumpsdeploy.pub'),
60 }
```

<https://github.com/wikimedia/puppet/blob/production/modules/profile/manifests/dumps/generation/worker/common.pp>

# Puppet Compiler

- Build puppet catalog for gerrit change
- Target specific hosts
- Shows the diff, errors
- Not foolproof but pretty awesome
- Use Wikitech creds to log in

Compilation results for deploy1001.eqiad.wmnet: **changes detected**

## Catalog differences

### Summary

Total Resources: 3887  
Resources added: 0  
Resources removed: 0  
Resources modified: 4  
Change percentage: 0.10%

### Resources modified

- Class[Profile::Kubernetes::Deployment\_server::Global\_config]  
Parameters differences:  
--- Class[Profile::Kubernetes::Deployment\_server::Global\_config].orig  
+++ Class[Profile::Kubernetes::Deployment\_server::Global\_config]  
  
@@  
- general\_values => {'default': {'tls': {'image\_version': '1.15.1-2'}}}  
+ general\_values => {'default': {'tls': {'image\_version': '1.15.1-2'}, 'monitoring': {'image\_version': '0.0.7'}}}
- File[/etc/helmfile-defaults/general-codfw.yaml]  
Content differences:  
--- /etc/helmfile-defaults/general-codfw.yaml.orig  
+++ /etc/helmfile-defaults/general-codfw.yaml  
@@ -5,6 +5,8 @@  
prometheus\_nodes:  
- 10.192.0.145  
- 10.192.16.189  
+monitoring:  
+ image\_version: 0.0.7  
puppet\_ca\_cert: |  
-----BEGIN CERTIFICATE-----  
MIIFDYCCA0IgwIBAqIBATANBgkqhkiG9w0BAQsFADAcMR0wGAYDVQQDDBFQdXBw
- File[/etc/helmfile-defaults/general-eniad.yaml]

<https://integration.wikimedia.org/ci/job/operations-puppet-catalog-compiler/build?delay=0sec>

# Puppet in deployment-prep

- Automatic sync every ten minutes
- Sometimes breaks, check that your change arrived
- Host: deployment-puppetmaster04.deployment-prep.eqiad.wmflabs
- Repo: `/var/lib/git/operations/puppet` and `git log`
- Sync errors in: `/var/log/git-sync-upstream.log`



# Hiera in deployment-prep

- Use Horizon interface
- Check your instance, your prefix-puppet and your project-puppet
- Wikitech creds to log in

## deployment-snapshot01

Overview Log Action Log **Puppet Configuration**

This instance is also affected by the following puppet configs: [project config](#), [deployment-snapshot](#)

### Puppet Classes

```
role::beta::mediawiki
role::dumps::generation::worker::beta_testbed
```

Role revision history can be viewed on  [Gerrit](#).

Edit

### Hiera Config

```
profile::dumps::generation::worker::common::dumps_misc_cronrunner: false
profile::dumps::generation::worker::common::nfs_extra_mountopts: actimeo=0
profile::dumps::generation::worker::common::php: /usr/bin/php7.2
profile::dumps::generation_worker_cron_php: /usr/bin/php7.2
```

[https://horizon.wikimedia.org/project/instances/898fe007-ec78-465e-b298-d9c57a-ee7aad/?prev\\_marker=a82c5f27-a92b-42a6-9f3a-530f9feae621](https://horizon.wikimedia.org/project/instances/898fe007-ec78-465e-b298-d9c57a-ee7aad/?prev_marker=a82c5f27-a92b-42a6-9f3a-530f9feae621)

# Puppet-merge in production

- Last chance after +2 and submit to abort
- Check the diff to be sure it's what you added
- “Multiple merge?” = someone else trying at the same time, GET APPROVAL



# Testing in production (heh)

- Do a live run on a target
- (root) puppet agent --test
- Output:  
/var/log/puppet.log
- Trivial changes aren't
- No-ops aren't either



# Disabling puppet

- On one host: puppet agent --disable "MY NAME AND MESSAGE HERE OR ELSE"
- On multiple hosts: via cumin server as root AND YOU STILL BETTER PROVIDE A MESSAGE
- Reasons to do this: testing on one host, don't want puppet overwriting tests, or
- Staged rollout of a change across the cluster

# Private puppet repo?

- Yes there is one
- On puppetmaster in production
- NO COPIES ANYWHERE, yes this means you
- WMCS has a “fake” private repo with passwords visible to the world, for testing
- Dumps doesn't use these but gtk anyways

# Further reading

- <https://wikitech.wikimedia.org/wiki/Puppet>
- [https://wikitech.wikimedia.org/wiki/Puppet\\_coding](https://wikitech.wikimedia.org/wiki/Puppet_coding)
- <https://wikitech.wikimedia.org/wiki/Help:Puppet-compiler>
- [https://wikitech.wikimedia.org/wiki/Puppet\\_Hiera](https://wikitech.wikimedia.org/wiki/Puppet_Hiera)
- <https://puppet.com/docs/puppet/6.19/facter.html>
- [https://puppet.com/docs/puppet/6.19/puppet\\_language.html](https://puppet.com/docs/puppet/6.19/puppet_language.html)

Thanks!  
Questions, comments, gripes?  
You don't know where to find me;  
I'll be hiding.

“puppet isn't so much of a language as it is an incantation phrase book powered by souls devoured thousands of years ago and given form by the heartache of opsen everywhere”

# Credits

All images from commons.wikimedia.org or derived from them, or my own screenshots.

- File:Dof\_blocks\_f22.jpg
- File:Flow\_chart\_for\_flow\_in\_AfC\_on\_english\_wikipedia\_(2).png
- File:Hajrullah\_SYLA,\_Composer\_and\_Conductor.jpeg
- File:Heinold%u2019s\_First\_and\_Last\_Chance\_2007.jpg
- File:Intermediate\_inheritance\_P\_-\_F1\_-\_F2.png
- File:LEGO\_Notre\_Dame\_de\_Paris\_1.jpg
- File:Puppet\_Logo.svg
- File:Ruby\_logo.svg
- File:Tux\_Paint\_t-shirt\_01.svg
- File:USMC-17547\_cropped.jpg
- File:Walks-avm.png
- [https://bash.toolforge.org/quip/AU\\_3cMwj1oXzWjit5obk](https://bash.toolforge.org/quip/AU_3cMwj1oXzWjit5obk)