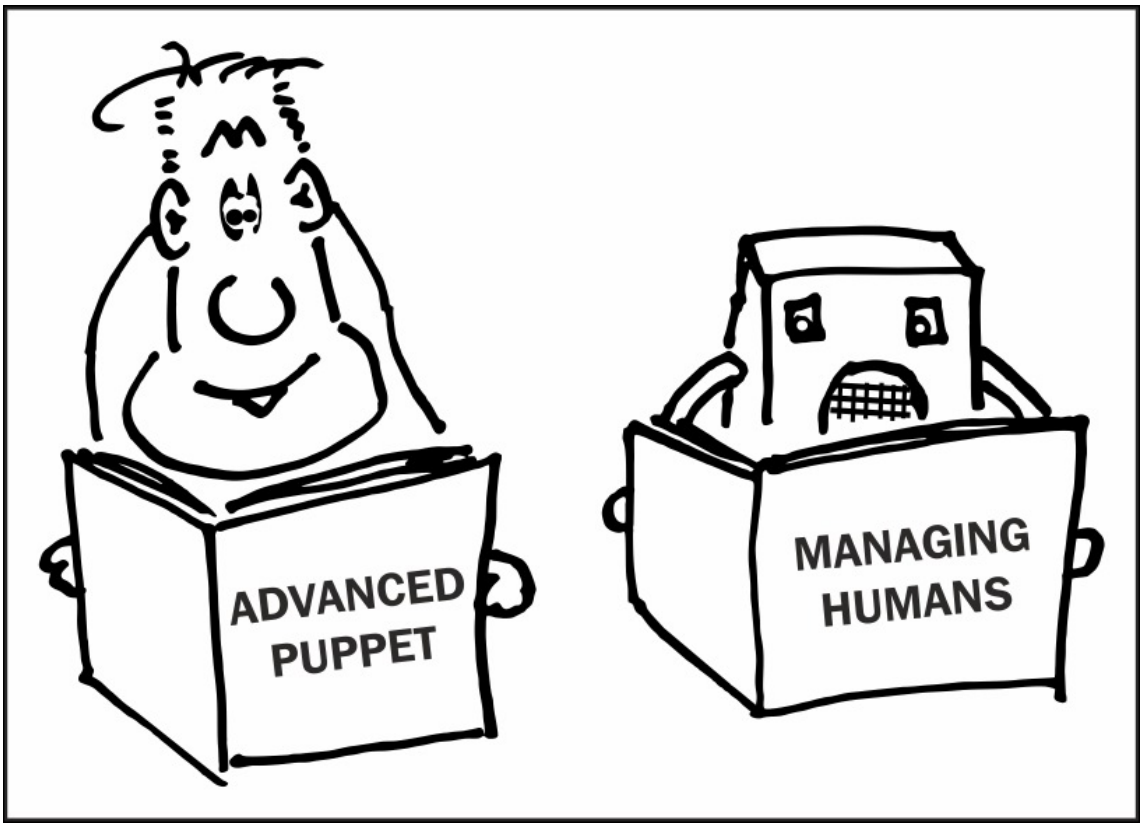
Chapter 7. Mastering modules

In this lab you'll learn about Puppet Forge, the public repository for Puppet modules, and you'll see how to install and use third-party modules from Puppet Forge, using the `r10k` module management tool. You'll see examples of how to use three important Forge modules: `puppetlabs/apache`, `puppetlabs/mysql`, and `puppet/archive`. You'll be introduced to some useful functions provided by `puppetlabs/stdlib`, the Puppet standard library. Finally, working through a complete example, you'll learn how to develop your own Puppet module from scratch, how to add appropriate metadata for your module, and how to upload it to Puppet Forge.



Using Puppet Forge modules

Although you could write your own manifests for everything you want to manage, you can save yourself a lot of time and effort by using public Puppet modules wherever possible. A **module** in Puppet is a self-contained unit of shareable, reusable code, usually designed to manage one particular service or piece of software, such as the Apache web server.

What is the Puppet Forge?

The **Puppet Forge** is a public repository of Puppet modules, many of them officially supported and maintained by Puppet and all of which you can download and use. You can browse the Forge at the following URL:

https://forge.puppet.com/

One of the advantages of using a well-established tool like Puppet is that there are a large number of mature public modules available, which cover the most common software you're likely to need. For example, here is a small selection of the things you can manage with public modules from Puppet Forge:

- MySQL/PostgreSQL/SQL Server
- Apache/Nginx
- Java/Tomcat/PHP/Ruby/Rails
- HAProxy
- Amazon AWS
- Docker
- Jenkins
- Elasticsearch/Redis/Cassandra
- Git repos
- Firewalls (via iptables)

Finding the module you need

The Puppet Forge home page has a search bar at the top. Type what you're looking for into this box, and the website will show you all the modules which match your search keywords. Often, there will be more than one result, so how do you decide which module to use?

The best choice is a **Puppet Supported** module, if one is available. These are officially supported and maintained by Puppet, and you can be confident that supported modules will work with a wide range of operating systems and Puppet versions. Supported modules are indicated by a yellow **SUPPORTED** flag in search results, or you can browse the list of all supported modules at the following URL:

https://forge.puppet.com/modules?endorsements=supported

The next best option is a **Puppet Approved** module. While not officially supported, these modules are recommended by Puppet and have been checked to make sure they follow best practices and meet certain quality standards. Approved modules are indicated by a green **APPROVED** flag in search results, or you can browse the list of all approved modules at the following URL:

https://forge.puppet.com/modules?endorsements=approved

Assuming that a Puppet-Supported or Puppet-Approved module is not available, another useful way to choose modules is by looking at the number of downloads. Selecting the **Most Downloads** tab on the Puppet Forge search results page will sort the results by downloads, with the most popular modules first. The most-downloaded modules are not necessarily the best, of course, but they're usually a good place to start.

It's also worth checking the latest release date for modules. If the module you're looking at hasn't been updated in over a year, it may be better to go with a more actively-maintained module, if one is available. Clicking on the **Latest release** tab will sort search results by the most recently updated.

You can also filter search results by operating system support and Puppet version compatibility; this can be very useful for finding a module that works with your system.

Having chosen the module you want, it's time to add it to your Puppet infrastructure.

Using r10k

In the past, many people used to download Puppet Forge modules directly and check a copy of them into their codebase, effectively forking the module repo (and some still do this). There are many drawbacks to this approach. One is that your codebase becomes cluttered with code that is not yours, and this can make it difficult to search for

the code you want. Another is that it's difficult to test your code with different versions of public modules, without creating your own Git branches, redownloading the modules, and so on. You also won't get future bug fixes and improvements from the Puppet Forge modules unless you manually update your copies. In many cases, you will need to make small changes or fixes to the modules to use them in your environment, and your version of the module will then diverge from the upstream version, storing up maintenance problems for the future.

A much better approach to module management, therefore, is to use the `r10k` tool, which eliminates these problems. Instead of downloading the modules you need directly and adding them to your codebase, `r10k` installs your required modules on each Puppet-managed node, using a special text file called a **Puppetfile**. `r10k` will manage the contents of your `modules/` directory, based entirely on the Puppetfile metadata. The module code is never checked into your codebase, but always downloaded from the Puppet Forge when requested. So you can stay up to date with the latest releases if you want, or pin each module to a specified version which you know works with your manifest.

`r10k` is the de facto standard module manager for Puppet deployments, and we'll be using it to manage modules throughout the rest of this course.

In this example, we'll use `r10k` to install the `puppetlabs/stdlib` module. The Puppetfile in the example repo contains a list of all the modules we'll use in this course. Here it is (we'll look more closely at the syntax in a moment):

```
forge 'http://forge.puppetlabs.com'

mod 'garethr/docker', '5.3.0'
mod 'puppet/archive', '1.3.0'
mod 'puppet/staging', '2.2.0'
mod 'puppetlabs/apache', '2.0.0'
mod 'puppetlabs/apt', '3.0.0'
mod 'puppetlabs/aws', '2.0.0'
mod 'puppetlabs/concat', '4.0.1'
mod 'puppetlabs/docker_platform', '2.2.1'
mod 'puppetlabs/mysql', '3.11.0'
mod 'puppetlabs/stdlib', '4.17.1'
mod 'stahnma/epel', '1.2.2'

mod 'pbg_ntp',
  :git => 'https://github.com/bitfield/pbg_ntp.git',
  :tag => '0.1.4'
```

Follow these steps:

1. Run the following commands to clear out your `modules/` directory, if there's anything in it (make sure you have backed up anything here you want to keep):

   ```
   cd /etc/puppetlabs/code/environments/pbg
   sudo rm -rf modules/
   ```

2. Run the following command to have `r10k` process the example Puppetfile here and install your requested modules:

   ```
   sudo r10k puppetfile install --verbose
   ```

`r10k` downloads all the modules listed in the Puppetfile into the `modules/` directory. All modules in this directory will be automatically loaded by Puppet and available for use in your manifests. To test that the `stdlib` module is correctly installed, run the following command:

```
sudo puppet apply --environment pbg -e "notice(upcase('hello'))"
Notice: Scope(Class[main]): HELLO
```

The `upcase` function, which converts its string argument to uppercase, is part of the `stdlib` module. If this doesn't work, then `stdlib` has not been properly installed. As in previous examples, we're using the `--environment pbg` switch to tell Puppet to look for code, modules, and data in the `/etc/puppetlabs/code/environments/pbg` directory.

Understanding the Puppetfile

The example Puppetfile begins with the following:

```
forge 'http://forge.puppetlabs.com'
```

The `forge` statement specifies the repository where modules should be retrieved from.

There follows a group of lines beginning with `mod`:

```
mod 'garethr/docker', '5.3.0'
mod 'puppet/archive', '1.3.0'
mod 'puppet/staging', '2.2.0'
...
```

The `mod` statement specifies the name of the module (`puppetlabs/stdlib`) and the specific version of the module to install (`4.17.0`).

Managing dependencies with generate-puppetfile

`r10k` does not automatically manage dependencies between modules. For example, the `puppetlabs/apache` module depends on having both `puppetlabs/stdlib` and `puppetlabs/concat` installed. `r10k` will not automatically install these for you unless you specify them, so you also need to include them in your Puppetfile.

However, you can use the `generate-puppetfile` tool to find out what dependencies you need so that you can add them to your Puppetfile.

1. Run the following command to install the `generate-puppetfile` gem:

   ```
   sudo gem install generate-puppetfile
   ```

2. Run the following command to generate the Puppetfile for a list of specified modules (list all the modules you need on the command line, separated by spaces):

   ```
   generate-puppetfile puppetlabs/docker_platform
   Installing modules. This may take a few minutes.
   Your Puppetfile has been generated. Copy and paste between the markers:
   ===========================================
   forge 'http://forge.puppetlabs.com'

   # Modules discovered by generate-puppetfile
   ```

```
mod 'garethr/docker', '5.3.0'
mod 'puppetlabs/apt', '3.0.0'
mod 'puppetlabs/docker_platform', '2.2.1'
mod 'puppetlabs/stdlib', '4.17.1'
mod 'stahnma/epel', '1.2.2'
============================================
```

3. Run the following command to generate a list of updated versions and dependencies for an existing Puppetfile:

```
generate-puppetfile -p /etc/puppetlabs/code/environments/pbg/Puppetfile
```

This is an extremely useful tool both for finding dependencies you need to specify in your Puppetfile and for keeping your Puppetfile up to date with the latest versions of all the modules you use.

Using modules in your manifests

Now that we know how to find and install public Puppet modules, let's see how to use them. We'll work through a few examples, using the `puppetlabs/mysql` module to set up a MySQL server and database, using the `puppetlabs/apache` module to set up an Apache website, and using `puppet/archive` to download and unpack a compressed archive. After you've tried out these examples, you should feel quite confident in your ability to find an appropriate Puppet module, add it to your `Puppetfile`, and deploy it with `r10k`.

Using puppetlabs/mysql

Follow these steps to run the `puppetlabs/mysql` example:

1. If you've previously followed the steps in the [*Using r10k*] section, the required module will already be installed. If not, run the following commands to install it:

```
cd /etc/puppetlabs/code/environments/pbg
sudo r10k puppetfile install
```

2. Run the following command to apply the manifest:

```
sudo puppet apply --environment=pbg /examples/module_mysql.pp
Notice: Compiled catalog for ubuntu-xenial in environment pbg in 0.89 seconds
Notice: /Stage[main]/Mysql::Server::Config/File[/etc/mysql]/ensure: created
Notice: /Stage[main]/Mysql::Server::Config/File[/etc/mysql/conf.d]/ensure:
created
Notice: /Stage[main]/Mysql::Server::Config/File[mysql-config-file]/ensure:
defined content as '{md5}44e7aa974ab98260d7d013a2087f1c77'
Notice: /Stage[main]/Mysql::Server::Install/Package[mysql-server]/ensure:
created
Notice:
/Stage[main]/Mysql::Server::Root_password/Mysql_user[root@localhost]/password_has
 password_hash changed '' to '*F4AF2E5D85456A908E0F552F0366375B06267295'
Notice: /Stage[main]/Mysql::Server::Root_password/File[/root/.my.cnf]/ensure:
defined content as '{md5}4d59f37fc8a385c9c50f8bb3286b7c85'
Notice: /Stage[main]/Mysql::Client::Install/Package[mysql_client]/ensure:
created
Notice:
```

```
/Stage[main]/Main/Mysql::Db[cat_pictures]/Mysql_database[cat_pictures]/ensure:
created
Notice:
/Stage[main]/Main/Mysql::Db[cat_pictures]/Mysql_user[greebo@localhost]/ensure:
created
Notice:
/Stage[main]/Main/Mysql::Db[cat_pictures]/Mysql_grant[greebo@localhost/cat_pictur
 created
Notice: Applied catalog in 79.85 seconds
```

Let's take a look at the example manifest (`module_mysql.pp`). The first part installs the MySQL server itself, by including the class `mysql::server` :

```
# Install MySQL and set up an example database
include mysql::server
```

The `mysql::server` class accepts a number of parameters, most of which we needn't worry about for now, but we would like to set a couple of them for this example. Although you can set the values for class parameters directly in your Puppet manifest code, just as you would for resource attributes, I'll show you a better way to do it: using Hiera's automatic parameter lookup mechanism.

Note

We mentioned briefly in [Lab 6], [*Managing data with Hiera*], that Hiera can supply values for class and module parameters, but how does it work, exactly? When you include a class `x` which takes a parameter `y` , Puppet automatically searches Hiera for any keys matching the name `x::y` . If it finds one, it uses that value for the parameter. Just as with any other Hiera data, you can use the hierarchy to set different values for different nodes, roles, or operating systems.

In this example, our parameters are set in the example Hiera data file (`/etc/puppetlabs/code/environments/pbg/data/common.yaml`):

```
mysql::server::root_password: 'hairline-quotient-inside-tableful'
mysql::server::remove_default_accounts: true
```

The `root_password` parameter, as you'd expect, sets the password for the MySQL `root` user. We also enable `remove_default_accounts` , which is a security feature. MySQL ships with various default user accounts for testing purposes, which should be turned off in production. This parameter disables these default accounts.

Note

Note that although we've specified the password in plain text for the purposes of clarity, in your production manifests, this should be encrypted, just as with any other credentials or secret data (see [Lab 6], [*Managing data with Hiera*]).

Next comes a resource declaration:

```
mysql::db { 'cat_pictures':
  user     => 'greebo',
  password => 'tabby',
  host     => 'localhost',
  grant    => ['SELECT', 'UPDATE'],
}
```

As you can see, this looks just like the built-in resources we've used before, such as the `file` and `package` resources. In effect, the `mysql` module has added a new resource type to Puppet: `mysql::db`. This resource models a specific MySQL database: `cat_pictures` in our example.

The title of the resource is the name of the database, in this case, `cat_pictures`. There follows a list of attributes. The `user`, `password`, and `host` attributes specify that the user `greebo` should be allowed to connect to the database from `localhost` using the password `tabby`. The `grant` attribute specifies the MySQL privileges that the user should have: `SELECT` and `UPDATE` on the database.

When this manifest is applied, Puppet will create the `cat_pictures` database and set up the `greebo` user account to access it. This is a very common pattern for Puppet manifests which manage an application: usually, the application needs some sort of database to store its state, and user credentials to access it. The `mysql` module lets you configure this very easily.

So we can now see the general principles of using a Puppet Forge module:

- We add the module and its dependencies to our `Puppetfile` and deploy it using `r10k`
- We `include` the class in our manifest, supplying any required parameters as Hiera data
- Optionally, we add one or more resource declarations of a custom resource type defined by the module (in this case, a MySQL database)

Almost all Puppet modules work in a similar way. In the next section, we'll look at some key modules which you're likely to need in the course of managing servers with Puppet.

Using puppetlabs/apache

Most applications have a web interface of some kind, which usually requires a web server, and the venerable Apache remains a popular choice. The `puppetlabs/apache` module not only installs and configures Apache, but also allows you to manage virtual hosts (individual websites, such as the frontend for your application).

Here's an example manifest which uses the `apache` module to create a simple virtual host serving an image file (`module_apache.pp`):

```
include apache

apache::vhost { 'cat-pictures.com':
  port          => '80',
  docroot       => '/var/www/cat-pictures',
  docroot_owner => 'www-data',
  docroot_group => 'www-data',
}

file { '/var/www/cat-pictures/index.html':
  content => "<img
    src='http://bitfieldconsulting.com/files/happycat.jpg'>",
  owner   => 'www-data',
  group   => 'www-data',
}
```

Follow these steps to apply the manifest:

1. If you've previously followed the steps in the [*Using r10k*] section, the required module will already be installed. If not, run the following commands to install it:

```
cd /etc/puppetlabs/code/environments/pbg
sudo r10k puppetfile install
```

2. Run the following command to apply the manifest:

```
sudo puppet apply --environment=pbg /examples/module_apache.pp
```

3. To test the new website, point your browser to (for Vagrant users; if you're not using the Vagrant box, browse to port `80` on the server you're managing with Puppet) `http://localhost:8080/`

You should see a picture of a happy cat:

Let's go through the manifest and see how it works in detail.

1. It starts with the `include` declaration which actually installs Apache on the server (`module_apache.pp`):

```
include apache
```

2. There are many parameters you could set for the `apache` class, but in this example, we only need to set one, and as with the other examples, we set it using Hiera data in the example Hiera file:

```
apache::default_vhost: false
```

```
This disables the default **Apache 2 Test Page** virtual
host.
```

3. Next comes a resource declaration for an `apache::vhost` resource, which creates an Apache virtual host or website.

```
apache::vhost { 'cat-pictures.com':
  port          => '80',
  docroot       => '/var/www/cat-pictures',
  docroot_owner => 'www-data',
  docroot_group => 'www-data',
}
```

```
The title of the resource is the domain name which the virtual host
will respond to (`cat-pictures.com`). The `port`
tells Apache which port to listen on for requests. The
`docroot` identifies the pathname of the directory where
Apache will find the website files on the server. Finally, the
`docroot_owner` and `docroot_group` attributes
specify the user and group which should own the `docroot/`
directory.
```

4. Finally, we create an `index.html` file to add some content to the website, in this case, an image of a happy cat.

```
file { '/var/www/cat-pictures/index.html':
  content => "<img
    src='http://bitfieldconsulting.com/files/happycat.jpg'>",
  owner   => 'www-data',
  group   => 'www-data',
}
```

Note

Note that port `80` on the Vagrant box is mapped to port `8080` on your local machine, so browsing to `http://localhost:8080` is the equivalent of browsing directly to port `80` on the Vagrant box. If for some reason you need to change this port mapping, edit your `Vagrantfile` (in the Puppet Beginner's Guide repo) and look for the following line:

```
config.vm.network "forwarded_port", guest: 80, host: 8080
```

Change these settings as required and run the following command on your local machine in the PBG repo directory:

```
vagrant reload
```

Using puppet/archive

While installing software from packages is a common task, you'll also occasionally need to install software from archive files, such as a tarball (a `.tar.gz` file) or ZIP file. The `puppet/archive` module is a great help for this, as it provides an easy way to download archive files from the Internet, and it can also unpack them for you.

In the following example, we'll use the `puppet/archive` module to download and unpack the latest version of the popular WordPress blogging software. Follow these steps to apply the manifest:

1. If you've previously followed the steps in the [*Using r10k*] section, the required module will already be installed. If not, run the following commands to install it:

```
cd /etc/puppetlabs/code/environments/pbg
sudo r10k puppetfile install
```

2. Run the following command to apply the manifest:

```
sudo puppet apply --environment=pbg /examples/module_archive.pp
Notice: Compiled catalog for ubuntu-xenial in environment production in 2.50
seconds
Notice: /Stage[main]/Main/Archive[/tmp/wordpress.tar.gz]/ensure: download
archive from https://wordpress.org/latest.tar.gz to /tmp/wordpress.tar.gz and
extracted in /var/www with cleanup
```

Unlike the previous modules in this lab, there's nothing to install with `archive`, so we don't need to include the class itself. All you need to do is declare an `archive` resource. Let's look at the example in detail to see how it works (`module_archive.pp`):

```
archive { '/tmp/wordpress.tar.gz':
  ensure       => present,
  extract      => true,
  extract_path => '/var/www',
  source       => 'https://wordpress.org/latest.tar.gz',
  creates      => '/var/www/wordpress',
  cleanup      => true,
}
```

1. The title gives the path to where you want the archive file to be downloaded (`/tmp/wordpress.tar.gz`). Assuming you don't need to keep the archive file after it's been unpacked, it's usually a good idea to put it in `/tmp` .
2. The `extract` attribute determines whether or not Puppet should unpack the archive; this should usually be set to `true` .
3. The `extract_path` attribute specifies where to unpack the contents of the archive. In this case, it makes sense to extract it to a subdirectory of `/var/www/` , but this will vary depending on the nature of the archive. If the archive file contains software which will be compiled and installed, for example, it may be a good idea to unpack it in `/tmp/` , so that the files will be automatically cleaned up after the next reboot.
4. The `source` attribute tells Puppet where to download the archive from, usually (as in this example) a web URL.
5. The `creates` attribute works exactly the same way as `creates` on an `exec` resource, which we looked at in [Lab 4], [*Understanding Puppet resources*]. It specifies a file which unpacking the archive will create. If this file exists, Puppet knows the archive has already been unpacked, so it does not need to unpack it again.
6. The `cleanup` attribute tells Puppet whether or not to delete the archive file once it has been unpacked. Usually, this will be set to `true` , unless you need to keep the archive around or unless you don't need to unpack it in the first place.

Note

Once the file has been deleted by `cleanup`, Puppet won't redownload the archive file `/tmp/wordpress.tar.gz` the next time you apply the manifest, even though it has `ensure => present`. The `creates` clause tells Puppet that the archive has already been downloaded and extracted.

Exploring the standard library

One of the oldest-established Puppet Forge modules is `puppetlabs/stdlib`, the official Puppet standard library. We looked at this briefly earlier in the lab when we used it as an example of installing a module with `r10k`, but let's look more closely now and see what the standard library provides and where you might use it.

Rather than managing some specific software or file format, the standard library aims to provide a set of functions and resources which could be useful in any piece of Puppet code. Consequently, well-written Forge modules use the facilities of the standard library rather than implementing their own utility functions which do the same thing.

You should do the same in your own Puppet code: when you need a particular piece of functionality, check the standard library first to see if it solves your problem rather than implementing it yourself.

Before trying the examples in this section, make sure the `stdlib` module is installed by following these steps: If you've previously followed the steps in the [*Using r10k*] section, the required module will already be installed. If not, run the following commands to install it:

```
cd /etc/puppetlabs/code/environments/pbg
sudo r10k puppetfile install
```

Safely installing packages with ensure_packages

As you know, you can install a package using the `package` resource, like this (`package.pp`):

```
package { 'cowsay':
  ensure => installed,
}
```

But what happens if you also install the same package in another class in a different part of your manifest? Puppet will refuse to run, with an error like this:

```
Error: Evaluation Error: Error while evaluating a Resource Statement, Duplicate
declaration: Package[cowsay] is already declared in file /examples/package.pp:1;
cannot redeclare at /examples/package.pp:4 at /examples/package.pp:4:1 on node ubuntu-
xenial
```

If both of your classes really require the package, then you have a problem. You could create a class which simply declares the package, and then include that in both classes, but that is a lot of overhead for a single package. Worse, if the duplicate declaration is in a third-party module, it may not be possible, or advisable, to change that code.

What we need is a way to declare a package which will not cause a conflict if that package is also declared somewhere else. The standard library provides this facility in the `ensure_packages()` function. Call `ensure_packages()` with an array of package names, and they will be installed if they are not already declared elsewhere (`package_ensure.pp`):

```
ensure_packages(['cowsay'])
```

To apply this example, run the following command:

```
sudo puppet apply --environment=pbg /examples/package_ensure.pp
```

You can try all the remaining examples in this lab in the same way. Make sure you supply the `--environment=pbg` switch to `puppet apply`, as the necessary modules are only installed in the `pbg` environment.

If you need to pass additional attributes to the `package` resource, you can supply them in a hash as the second argument to `ensure_packages()`, like this (`package_ensure_params.pp`):

```
ensure_packages(['cowsay'],
  {
    'ensure' => 'latest',
  }
)
```

Why is this better than using the `package` resource directly? When you declare the same `package` resource in more than one place, Puppet will give an error message and refuse to run. If the package is declared by `ensure_packages()`, however, Puppet will run successfully.

Since it provides a safe way to install packages without resource conflicts, you should always use `ensure_packages()` instead of the built-in `package` resource. It is certainly essential if you're writing modules for public release, but I recommend you use it in all your code. We'll use it to manage packages throughout the rest of this course.

Modifying files in place with file_line

Often, when managing configuration with Puppet, we would like to change or add a particular line to a file, without incurring the overhead of managing the whole file with Puppet. Sometimes it may not be possible to manage the whole file in any case, as another Puppet class or another application may be managing it. We could write an `exec` resource to modify the file for us, but the standard library provides a resource type for exactly this purpose: `file_line`.

Here's an example of using the `file_line` resource to add a single line to a system config file (`file_line.pp`):

```
file_line { 'set ulimits':
  path => '/etc/security/limits.conf',
  line => 'www-data         -     nofile          32768',
}
```

If there is a possibility that some other Puppet class or application may need to modify the target file, use `file_line` instead of managing the file directly. This ensures that your class won't conflict with any other attempts to control the file.

You can also use `file_line` to find and modify an existing line, using the `match` attribute (`file_line_match.pp`):

```
file_line { 'adjust ulimits':
  path  => '/etc/security/limits.conf',
  line  => 'www-data         -     nofile          9999',
  match => '^www-data .* nofile',
}
```

The value of `match` is a regular expression, and if Puppet finds a line in the file which matches this expression, it will replace it with the value of `line`. (If you need to potentially change multiple lines, set the `multiple` attribute to `true` or Puppet will complain when more than one line matches the expression.)

You can also use `file_line` to delete a line in a file if it is present (`file_line_absent.pp`):

```
file_line { 'remove dash from valid shells':
  ensure            => absent,
  path              => '/etc/shells',
  match             => '^/bin/dash',
  match_for_absence => true,
}
```

Note that when using `ensure => absent`, you also need to set the `match_for_absence` attribute to `true` if you want Puppet to actually delete matching lines.

Introducing some other useful functions

The `grep()` function will search an array for a regular expression and return all matching elements (`grep.pp`):

```
$values = ['foo', 'bar', 'baz']
notice(grep($values, 'ba.*'))

# Result: ['bar', 'baz']
```

The `member()` and `has_key()` functions return `true` if a given value is in the specified array or hash, respectively (`member_has_key.pp`):

```
$values = [
  'foo',
  'bar',
  'baz',
]
notice(member($values, 'foo'))

# Result: true

$valuehash = {
  'a' => 1,
  'b' => 2,
  'c' => 3,
}
notice(has_key($valuehash, 'b'))

# Result: true
```

The `empty()` function returns `true` if its argument is an empty string, array, or hash (`empty.pp`):

```
notice(empty(''))

# Result: true

notice(empty([]))
```

```
# Result: true

notice(empty({}))

# Result: true
```

The `join()` function joins together the elements of a supplied array into a string, using a given separator character or string (`join.pp`):

```
$values = ['1', '2', '3']
notice(join($values, '... '))

# Result: '1... 2... 3'
```

The `pick()` function is a neat way to provide a default value when a variable happens to be empty. It takes any number of arguments and returns the first argument which is not undefined or empty (`pick.pp`):

```
$remote_host = ''
notice(pick($remote_host, 'localhost'))

# Result: 'localhost'
```

Sometimes you need to parse structured data in your Puppet code which comes from an outside source. If that data is in YAML format, you can use the `loadyaml()` function to read and parse it into a native Puppet data structure (`loadyaml.pp`):

```
$db_config = loadyaml('/examples/files/database.yml')
notice($db_config['development']['database'])

# Result: 'dev_db'
```

The `dirname()` function is very useful if you have a string path to a file or directory and you want to reference its parent directory, for example to declare it as a Puppet resource (`dirname.pp`):

```
$file = '/var/www/vhosts/mysite'
notice(dirname($file))

# Result: '/var/www/vhosts'
```

The pry debugger

When a Puppet manifest doesn't do quite what you expect, troubleshooting the problem can be difficult. Printing out the values of variables and data structures with `notice()` can help as can running `puppet apply -d` to see detailed debug output, but if all else fails, you can use the standard library's `pry()` method to enter an interactive debugger session (`pry.pp`):

```
pry()
```

With the `pry` gem installed in Puppet's context, you can call `pry()` at any point in your code. When you apply the manifest, Puppet will start an interactive Pry shell at the point where the `pry()` function is called. You can then

run the `catalog` command to inspect Puppet's catalog, which contains all the resources currently declared in your manifest:

```
sudo puppet apply --environment=pbg /examples/pry_install.pp
sudo puppet apply --environment=pbg /examples/pry.pp
...
[1] pry(#<Puppet::Parser::Scope>)> catalog
=> #<Puppet::Resource::Catalog:0x00000001bbcf78
...
 @resource_table={["Stage", "main"]=>Stage[main]{}, ["Class",
"Settings"]=>Class[Settings]{}, ["Class", "main"]=>Class[main]{}},
 @resources=[["Stage", "main"], ["Class", "Settings"], ["Class", "main"]],
...
```

Once you've finished inspecting the catalog, type `exit` to quit the debugger and continue applying your Puppet manifest.

Writing your own modules

As we've seen, a Puppet module is a way of grouping together a set of related code and resources that performs some particular task, like managing the Apache web server or dealing with archive files. But how do you actually create a module? In this section, we'll develop a module of our own to manage the NTP service, familiar to most system administrators as the easiest way to keep server clocks synchronized with the Internet time standard. (Of course, it's not necessary to write your own module for this because a perfectly good one exists on Puppet Forge. But we'll do so anyway, for learning purposes.)

Creating a repo for your module

If we're going to use our new module alongside others that we've installed from Puppet Forge, then we should create a new Git repo just for our module. Then we can add its details to our Puppetfile and have `r10k` install it for us.

If you've already worked through [Lab 3], [*Managing your Puppet code with Git], you'll have created a GitHub account. If not, go to that lab and follow the instructions in the [Creating a GitHub account and project*] section before continuing:

1. Log in to your GitHub account and click on the **Start a project** button.
2. On the **Create a new repository** screen, enter a suitable name for your repo (I'm using `pbg_ntp` for the Puppet Beginner's Guide's NTP module).
3. Check the **Initialize this repository with a README** box.
4. Click on **Create repository**.
5. GitHub will take you to the project page for the new repository. Click on the **Clone or download** button. If you're using GitHub with an SSH key, as we discussed in [Lab 3], [*Managing your Puppet code with Git], copy the **Clone with SSH** link. Otherwise, click on **Use HTTP**S and copy the **Clone with HTTPS* link.
6. On your own computer, or wherever you develop Puppet code, run the following command to clone the new repo (use the GitHub URL you copied in the previous step instead of this one):

   ```
   git clone https://github.com/bitfield/pbg_ntp.git
   ```

When the clone operation completes successfully, you're ready to get started with creating your new module.

Writing the module code

As you'll see if you look inside the Puppet Forge modules you've already installed, modules have a standard directory structure. This is so that Puppet can automatically find the manifest files, templates, and other components within the module. Although complex modules have many subdirectories, the only ones we will be concerned with in this example are manifests and files. In this section, we'll create the necessary subdirectories, write the code to manage NTP, and add a config file which the code will install.

Note

All the code and files for this module are available in the GitHub repo at the following URL:

https://github.com/bitfield/pbg_ntp

1. Run the following commands to create the `manifests` and `files` subdirectories:

```
cd pbg_ntp
mkdir manifests
mkdir files
```

2. Create the file `manifests/init.pp` with the following contents:

```
# Manage NTP
class pbg_ntp {
  ensure_packages(['ntp'])

  file { '/etc/ntp.conf':
    source  => 'puppet:///modules/pbg_ntp/ntp.conf',
    notify  => Service['ntp'],
    require => Package['ntp'],
  }

  service { 'ntp':
    ensure => running,
    enable => true,
  }
}
```

3. Create the file `files/ntp.conf` with the following contents:

```
driftfile /var/lib/ntp/ntp.drift

pool 0.ubuntu.pool.ntp.org iburst
pool 1.ubuntu.pool.ntp.org iburst
pool 2.ubuntu.pool.ntp.org iburst
pool 3.ubuntu.pool.ntp.org iburst
pool ntp.ubuntu.com

restrict -4 default kod notrap nomodify nopeer noquery limited
restrict -6 default kod notrap nomodify nopeer noquery limited
restrict 127.0.0.1
restrict ::1
```

4. Run the following commands to add, commit, and push your changes to GitHub (you'll need to enter your GitHub username and password if you're not using an SSH key):

```
git add manifests/ files/
git commit -m 'Add module manifest and config file'
[master f45dc50] Add module manifest and config file
 2 files changed, 29 insertions(+)
 create mode 100644 files/ntp.conf
 create mode 100644 manifests/init.pp
git push origin master
```

Notice that the `source` attribute for the `ntp.conf` file looks like the following:

```
puppet:///modules/pbg_ntp/ntp.conf
```

We haven't seen this kind of file source before, and it's generally only used within module code. The `puppet://` prefix indicates that the file comes from within the Puppet repo, and the path `/modules/pbg_ntp/` tells Puppet to look within the `pbg_ntp` module for it. Although the `ntp.conf` file is actually in the directory `modules/pbg_ntp/files/`, we don't need to specify the `files` part: that's assumed, because this is a `file` resource. (It's not just you: this confuses everybody).

Rather than installing the `ntp` package via a `package` resource, we use `ensure_packages()` from the standard library, as described earlier in this lab.

Creating and validating the module metadata

Every Puppet module should have a file in its top-level directory named `metadata.json`, which contains helpful information about the module that can be used by module management tools, including Puppet Forge.

Create the file `metadata.json` with the following contents (use your own name and GitHub URLs):

```
{
  "name": "pbg_ntp",
  "version": "0.1.1",
  "author": "John Arundel",
  "summary": "Example module to manage NTP",
  "license": "Apache-2.0",
  "source": "https://github.com/bitfield/pbg_ntp.git",
  "project_page": "https://github.com/bitfield/pbg_ntp",
  "tags": ["ntp"],
  "dependencies": [
    {"name":"puppetlabs/stdlib",
      "version_requirement":">= 4.17.0 < 5.0.0"}
  ],
  "operatingsystem_support": [
    {
      "operatingsystem": "Ubuntu",
      "operatingsystemrelease": [ "16.04" ]
    }
  ]
}
```

Most of these are fairly self-explanatory. `tags` is an array of strings which will help people find your module if it is listed on Puppet Forge, and it's usual to tag your module with the name of the software or service it manages (in this case, `ntp`).

If your module relies on other Puppet modules, which is very likely (for example, this module relies on `puppetlabs/stdlib` for the `ensure_packages()` function) you use the `dependencies` metadata to record this. You should list each module used by your module along with the earliest and latest versions of that module which will work with your module. (If the currently-released version works, specify the next major release as the latest version. For example, if your module works with `stdlib` version 4.17.0 and that's the latest version available, specify 5.0.0 as the highest compatible version.)

Finally, the `operatingsystem_support` metadata lets you specify which operating systems and versions your module works with. This is very helpful for people searching for a Puppet module which will work with their operating system. If you know your module works with Ubuntu 16.04, as the example module does, you can list that in the `operatingsystem_support` section. The more operating systems your module can support, the better, so if possible, test your module on other operating systems and list them in the metadata once you know they work.

Note

For full details on module metadata and how to use it, see the Puppet documentation:

https://docs.puppet.com/puppet/latest/reference/modules_metadata.html

It's important to get the metadata for your module right, and there's a little tool that can help you with this, called `metadata-json-lint`.

1. Run the following commands to install `metadata-json-lint` and check your metadata:

   ```
   sudo gem install metadata-json-lint
   metadata-json-lint metadata.json
   ```

2. If `metadata-json-lint` produces no output, your metadata is valid and you can go on to the next steps. If you see error messages, fix the problem before continuing.

3. Run the following commands to add, commit, and push your metadata file to GitHub:

   ```
   git add metadata.json
   git commit -m 'Add metadata.json'
   git push origin master
   ```

Tagging your module

Just like when you use third-party Puppet Forge modules, it's important to be able to specify in your Puppetfile the exact version of your module to be installed. You can do this by using Git tags to attach a version tag to a specific commit in your module repo. As you develop the module further and make new releases, you can add a new tag for each release.

For the first release of your module, which according to the metadata is version 0.1.1, run the following commands to create and push the release tag:

```
git tag -a 0.1.1 -m 'Release 0.1.1'
git push origin 0.1.1
```

Installing your module

We can use `r10k` to install our new module, just as we did with the Puppet Forge modules, with one small difference. Since our module isn't on the Puppet Forge (yet), just specifying the name of the module in our Puppetfile isn't enough; we need to supply the Git URL so that `r10k` can clone the module from GitHub.

1. Add the following `mod` statement to your Puppetfile (using your GitHub URL instead of mine):

   ```
   mod 'pbg_ntp',
     :git => 'https://github.com/bitfield/pbg_ntp.git',
     :tag => '0.1.1'
   ```

2. Because the module also requires `puppetlabs/stdlib`, add this `mod` statement too:

   ```
   mod 'puppetlabs/stdlib', '4.17.0'
   ```

3. Now install the module in the normal way with `r10k`:

   ```
   sudo r10k puppetfile install --verbose
   ```

`r10k` can install a module from any Git repo you have access to; all you have to do is add the `:git` and `:tag` parameters to the `mod` statement in your Puppetfile.

Applying your module

Now that you've created, uploaded, and installed your module, we can use it in a manifest:

```
sudo puppet apply --environment=pbg -e 'include pbg_ntp'
```

If you're using the Vagrant box or a recent version of Ubuntu, your server will most likely be running NTP already, so the only change you'll see Puppet apply will be the `ntp.conf` file. Nonetheless, it confirms that your module works.

More complex modules

Of course, the module we've developed is a very trivial example. However, it demonstrates the essential requirements of a Puppet module. As you become a more advanced Puppet coder, you will be creating and maintaining much more complicated modules, similar to those you download and use from Puppet Forge.

Real-world modules often feature one or more of the following components:

- Multiple manifest files and subdirectories
- Parameters (which may be supplied directly or looked up from Hiera data)
- Custom facts and custom resource types and providers
- Example code showing how to use the module
- Specs and tests which developers can use to validate their changes
- Dependencies on other modules (which must be declared in the module metadata)
- Support for multiple operating systems

Note

You can find more detailed information about modules and advanced features of modules in the Puppet documentation:

https://docs.puppet.com/puppet/latest/reference/modules_fundamentals.html

Uploading modules to Puppet Forge

It's very easy to upload a module to the Puppet Forge: all you need to do is sign up for an account, use the `puppet module build` command to create an archive file of your module, and upload it via the Puppet Forge website.

Before deciding to write a module in the first place, though, you should check whether there is already a module on the Puppet Forge which does what you need. There are over 4,500 modules available at the time of writing, so it's quite likely that you'll be able to use an existing Puppet Forge module instead of writing your own. Contributing a new module when there is already one available just makes it more difficult for users to choose which module they should use. For example, there are currently 150 modules which manage the Nginx web server. Surely this is at least 149 too many, so only submit a new module if you've made sure that there are no similar modules already on the Puppet Forge.

If there is a module which covers the software you want to manage, but it doesn't support your operating system or version, consider improving this module instead of starting a new one. Contact the module author to see whether and how you can help improve their module and extend support to your operating system. Similarly, if you find bugs in a module or want to make improvements to it, open an issue (if there is an issue tracker associated with the module), fork the GitHub repo (if it's versioned on GitHub), or contact the author to find out how you can help. The vast majority of Puppet Forge modules are written and maintained by volunteers, so your support and contributions benefit the entire Puppet community.

If you don't want to fork or contribute to an existing module, consider writing a small wrapper module which extends or overrides the existing module, rather than creating a new module from scratch.

If you do decide to write and publish your own module, use facilities from the standard library wherever possible, such as `ensure_packages()`. This will give your module the best chance of being compatible with other Forge modules.

Note

If you want to contribute more to the Puppet module community, consider joining the Vox Pupuli group, which maintains over a hundred open source Puppet modules:

https://voxpupuli.org/

Summary

In this lab, we've gained an understanding of Puppet modules, including an introduction to the Puppet Forge module repository. We've seen how to search for the modules we need and how to evaluate the results, including **Puppet Approved** and **Puppet Supported** modules, operating system support, and download count.

We've looked at using the `r10k` tool to download and manage Puppet modules in your infrastructure and how to specify the modules and versions you need in your Puppetfile. We've worked through detailed examples of using three important Forge modules: `puppetlabs/apache`, `puppetlabs/mysql`, and `puppet/archive`.

Introducing the standard library for Puppet, we've covered the use of `ensure_packages()` to avoid package conflicts between modules, the `file_line` resource, which provides line-level editing for config files, and a host of useful functions for manipulating data, as well as looking at the Pry debugger.

To fully understand how modules work, we've developed a simple module from scratch to manage the NTP service, hosted in its own Git repository and managed via a Puppetfile and `r10k`. We've seen what metadata modules

require and how to create it and validate it using `metadata-json-lint`.

Finally, we've looked at some of the features of more sophisticated modules, discussed uploading modules to the Puppet Forge, and outlined some considerations to bear in mind when you're deciding whether to start a new module or extend and improve an existing one.

In the next lab, we'll look at how to organize your Puppet code into classes, how to pass parameters to your classes, how to create defined resource types, and how to structure your manifests using roles, profiles, and how to include classes on a node using Hiera data.