

Course Code: INFO3606
Course Title: Cloud Computing
Semester Project

Tevin Achong - 816000026, Jonathan Owen -

April 8, 2020

Contents

1	Chef	2
1.1	Components	2
1.2	Architecture	3
1.3	Features	3
1.4	Official Website	4
1.5	Latest Version	4
1.6	Advantages	4
1.7	Disadvantages	5
2	Ansible	6
2.1	Components	6
2.2	Architecture	7
2.3	Features	7
2.4	Official Website	8
2.5	Latest Version	8
2.6	Advantages	8
2.7	Disadvantages	9
3	Fuel	10
3.1	Components	10
3.2	Architecture	10
3.3	Features	10
3.4	Official Website	10
3.5	Latest Version	10
3.6	Advantages	10
3.7	Disadvantages	10
4	Puppet	11
4.1	Components	11
4.2	Architecture	11
4.3	Features	11
4.4	Official Website	11
4.5	Latest Version	11
4.6	Advantages	11
4.7	Disadvantages	11
5	Compass	12
5.1	Components	12
5.2	Architecture	12
5.3	Features	12
5.4	Official Website	12
5.5	Latest Version	12
5.6	Advantages	12
5.7	Disadvantages	12
6	References NEED TO FORMAT PROPERLY	13

1 Chef

Chef is a configuration management tool that is used to streamline the task of configuring and maintaining a company's servers. Additionally, Chef can integrate with various cloud-based platforms to automatically provision and configure new machines. It contains solutions for small and large scale systems, and features and pricing for each of the various ranges. Essentially, Chef ensures that the files and the software that users are expecting to be on a machine are actually present, configured correctly, and working as is expected. Performing these tasks for a single machine is fairly straightforward. However, as an organization's infrastructure scales up (more machines are introduced) it becomes increasingly difficult. This is the reason Chef was developed and is used.

1.1 Components

Chef operates with three core components: **Chef Server**, **Workstations**, and **Nodes**. These three components communicate in a linear fashion, as explained below.

1. **Workstations:** All the configuration code is created, tested, and changed at workstations, which are personal computers or virtual servers. There can exist as many workstations as is necessary. Additionally, cookbooks and policies that will be pushed to the Chef server and pulled by nodes are tested and maintained here. The Chef Workstation provides chef and command line tools, the testing tools Test Kitchen, ChefSpec, Cookstyle, and Foodcritic, and InSpec - a tool that allows you to write automated tests for compliance, security and policy requirements. Cookbooks created on workstations can be used privately or by one organization, or uploaded to the Chef Supermarket for other to use. Workstations can also be used to download cookbooks created by other Chef users and found in the Chef Supermarket.
2. **Nodes:** These are the servers that are managed by Chef i.e. the machines to which changes are being pushed. Nodes are generally a fleet of multiple machines that require the benefits of automation. Chef can manage nodes that are virtual servers, containers, network devices, and storage devices and a Chef client is installed on every node that is under management by Chef.
3. **Chef Server:** The Chef Server is the center of all of Chef's operations. It provides a communication pathway between the workstations (where infrastructure is coded) and the nodes (where the configurations are deployed by the Chef client). Any configuration files, metadata, cookbooks and other information created on workstations are stored on the Chef server. Additionally, the Chef server contains information regarding the state of all nodes at the time of the last chef-client run. Any changes made to infrastructure code must pass through the Chef server in order to be applied to nodes. Before accepting changes, the Chef server authenticates all communication via its REST API using public key encryption. In turn, the Chef Server is also made of several components which aid it in efficiently communicating with workstations and nodes. Each Chef Server uses an NGINX front-end load balancer to route all requests to the Chef Server API, and PostgreSQL to store data. A web interface, known as Chef manage, is used for common Chef server management tasks. An Apache Solr instance, wrapped by chef-solr, is used for indexing and searching. All these components help to make the Chef server capable of

handling requests for several thousands of nodes and make Chef server a resource heavy application.

1.2 Architecture

In order to achieve its goals, Chef treats infrastructure as code. Instead of manually changing anything, the machine setup is described in a Chef *recipe*. *Cookbooks* store collections of recipes. Ideally, one cookbook relates to a single task, but it can have numerous server configurations involved. The chef server stores each of the cookbooks and as a new chef client node checks in with the server, recipes are sent to tell the node how to configure itself. Afterwards, the client will occasionally check in with the server to see if anything needs to be changed. If something does need to be changed, then the client deals with it. Patches and updates can be applied to the entire infrastructure by changing the recipe; there is no need to interact with each machine individually. *Bookshelf* is used to store cookbooks and related files and templates. The Chef Server uses a Bookshelf that operates as a versioned repository. Full root access is required. Whenever a cookbook is uploaded to the Chef server, the new version of the cookbook is compared to the one already stored on the server. If any changes exist, a new version is stored. If resources are shared between cookbooks and cookbook versions, they will not be stored more than once. This is because the Chef Server stores one copy of a file or template. As such, Chef utilizes a three-tier client-server architecture. The working units (like cookbooks) are developed on the Chef workstation. From the command line utilities like Knife, they are uploaded to the Chef server and all the nodes which are present in the architecture are registered with the Chef server. Figure 1 below depicts this architecture.

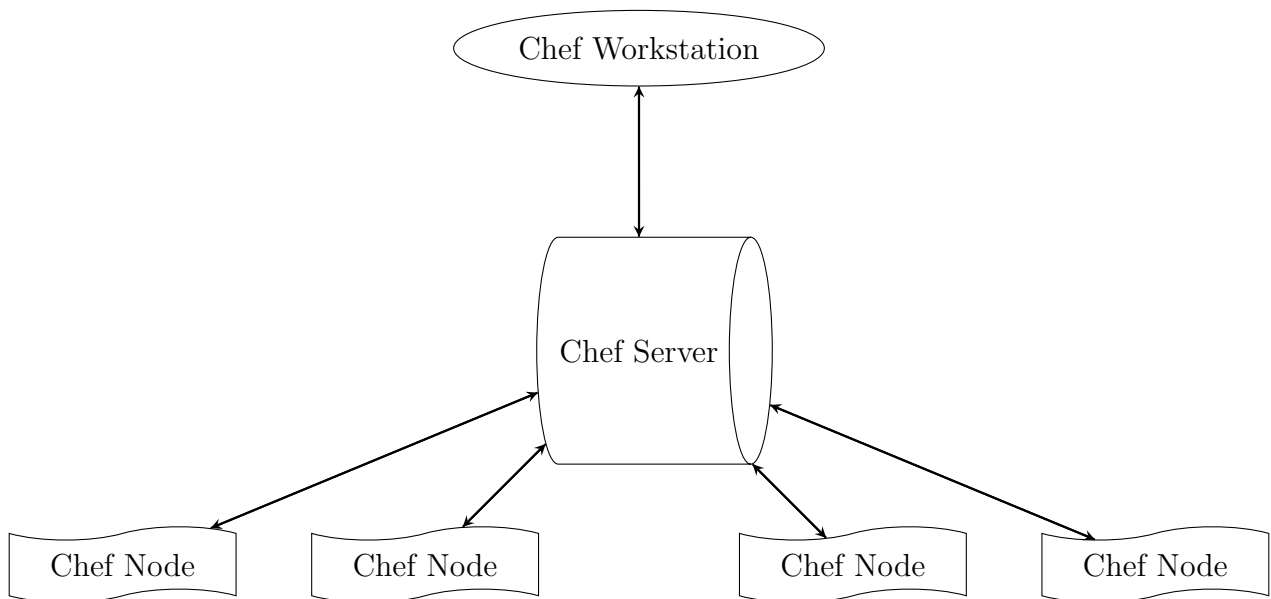


Figure 1: Architecture of Chef

1.3 Features

Chef provides the following features:

- Automatic Backup
- Automatic Notifications
- Compliance Management
- Configuration Management
- Data Recovery
- Data Visualization
- Real Time Analytics
- Real Time Data
- Real Time Monitoring
- Real Time Notifications
- Real Time Reporting
- Real Time Updates
- Server Monitoring

1.4 Official Website

The official website for Chef is located at <https://www.chef.io>.

1.5 Latest Version

As of January 28, 2019, the latest stable release of the Chef client is v14.10.9.

As of October 15, 2018, the latest stable release of the Chef server is v12.18.14.

1.6 Advantages

Advantages of using Chef include:

1. There is a low barrier for entry. Chef uses native Ruby language for configuration. This means that it can be easily picked up by anyone who has some development experience.
2. It integrates excellently with cloud. Using the knife utility, Chef can be easily integrated with most existing cloud technologies. It is a suitable tool for organizations that wish to distribute their infrastructure on multi-cloud environment.
3. Chef can help organizations to accelerate software delivery, i.e. how quickly the software is able to be changed in response to new requirements or conditions.
4. Helps organizations to catch bugs and issues before they occur. Infrastructure automation increases a system's resiliency just as much as it accelerates delivery speed.
5. Improving risk management - the *quality* of changes made to the software.
6. When using Chef, you can deliver all your infrastructure everywhere continuously.

1.7 Disadvantages

Some disadvantages include:

1. Cookbooks require constant monitoring so that the people who are working do not cause any issues with them.
2. Only Chef solo is available.
3. Currently, Chef is only a good fit for Amazon Web Services Cloud.
4. Someone who is not familiar with Ruby may have a hard time learning it.
5. There is currently a lack of proper documentation.

2 Ansible

Ansible is an open-source automation tool, or platform, used for IT tasks such as configuration management, application deployment, intraservice orchestration, and provisioning. Thus, simplifying complex tasks, making the jobs of developers more manageable and allowing them to focus attention on other tasks that add value to their organization.

2.1 Components

- **Plays**

Playbooks contain plays. Plays are essentially groups of tasks that are performed on defined hosts to enforce your defined functions. Each play must specify a host or group of hosts.

- **Tasks**

Tasks are actions carried out by playbooks. A task definition can contain modules such as yum, git, service, and copy.

- **Roles**

A role is the way Ansible bundles automation content to make it reusable. Roles are organizational components that can be assigned to a set of hosts to organize tasks. Therefore, instead of creating a monolithic playbook, we can create multiple roles, with each role assigned to complete a unit of work. For example: a webserver role can be defined to install Apache and Varnish on a specified group of servers.

- **Handlers**

Handlers are very similar to tasks. The major difference is that a handler will be executed only when it is called by an event. The handler is called by the [notify] directive. The name of the notify directive and the handler must be the same.

- **Templates**

Templates files are based on Python's Jinja2 template engine and have a .j2 extension. If necessary, you can place contents of your index.html file into a template file. But the real power of templates is comes when variables are used. Ansible's [facts] can be used in template files, and users can even call custom variables.

- **Variables**

You can include custom-made variables in your playbooks. Variables can be defined in five (5) different ways:

1. Variables defined under the vars_file attribute.
2. Variables defined in *role*/vars/apache-install.yml
3. Variables passed through the command line.
4. Variables defined in the play under vars.
5. Variables defined in group_vars/ directory.

2.2 Architecture

The architecture that makes up Ansible is as follows:

- **Modules**

Ansible works by connecting to your nodes and pushing out small programs, called "Ansible Modules" to them. These programs are written to be resource of the desired state of the system. Ansible then executes these modules (over SSH by default), and removes them when finished. Your library of modules can reside on any machine, and there are no servers, daemons, or databases required. Typically you'll work with your favorite terminal program, a text editor, and probably a version control system to keep track of changes to your content.

- **Plugins**

Plugins are pieces of code that augment Ansible's core functionality.

- **Inventory**

The inventory is a configuration file where you define the host information.

- **Playbooks**

A playbook is where users define how to apply policies, declare configurations, orchestrate steps and launch tasks either synchronously or asynchronously on their servers. Each playbook is composed of one or more plays. Playbooks are normally maintained and managed in a version control system like Git. They are expressed in YAML (Yet Another Markup Language). In most cases - especially in enterprise environments - using Ansible playbooks is recommended.

2.3 Features

- **Configuration Management**

Ansible is designed to be very simple, reliable, and consistent for configuration management. Users who are familiar with the field of Information Technology can usually get up and running fairly quickly with Ansible. Ansible configurations are simple data descriptions of infrastructure and are both readable by humans and parsable by machines. A password or an SSH key is all that is needed to start managing systems. For example, if a user wants to install an updated version of a specific type of software on all the machines in their enterprise, all they need to do is write out all the IP addresses of the nodes (also called remote hosts) and write an Ansible playbook to install it on all the nodes, then run the playbook from their control machine.

- **Application Deployment**

Ansible also allows users to quickly and easily deploy multitier applications. Users do not need to write custom code to automate their systems; they simply need to list the tasks required to be done in a playbook, and Ansible will figure out how to get their systems to the state they want them to be in. In other words, users need not configure the applications on every machine manually. When a playbook is run from the user's control machine, Ansible uses SSH to communicate with the remote hosts and run all the commands (tasks).

- **Orchestration** Orchestration involves bringing different elements into an efficiently-running, whole operation. For example, with application deployment, users need to manage the front-end and backend services, databases, networks, storage, and so on. They also need to ensure that all tasks are handled in the proper order. Ansible uses automated workflows, provisioning, and more to simplify the process of orchestration. Once the infrastructure has been defined using Ansible playbooks, that same orchestration can be used wherever it is needed. This is due largely to the portability of Ansible playbooks.
- **Security and Compliance**
As with application deployment, sitewide security policies (such as firewall rules) can be implemented along with other automated processes. If the security details on the control machine are configured and the associated playbook is run, all the remote hosts will automatically be updated with those details. Therefore, users will not need to monitor each machine for security compliance continually manually. For extra security, an admin's user ID and password are not retrievable in plain text on Ansible.
- **Cloud Provisioning**
The first step in automating your applications' life cycle is automating the provisioning of your infrastructure. With Ansible, you can provision cloud platforms, virtualized hosts, network devices, and bare-metal servers.

2.4 Official Website

The official website for Ansible is <https://www.ansible.com/>.

2.5 Latest Version

As of June 14, 2018, the latest version of Ansible is 2.5.5.

2.6 Advantages

Ansible offers the following advantages:

1. *Free*: Ansible is an open-source tool.
2. *Very simple to set up and use*: No special coding skills are necessary to use playbooks in Ansible.
3. *Powerful*: Ansible lets you model highly complex IT workflows.
4. *Flexible*: The entire application environment can be orchestrated no matter where it is deployed. It can also be customized to suit user's needs.
5. *Agentless*: No other software or firewall ports need to be installed on the client systems users wish to automate. Setting up a separate management structure is also unnecessary.
6. *Efficient*: Because there is no need to install any extra software, there is more room for application resources on the server.

2.7 Disadvantages

Disadvantages of Ansible include:

1. **No Notion of State:** Unlike comparable automation tools, Ansible has no notion of state. Since it does not keep track of dependencies, the tool simply executes a sequential series of tasks, stopping when it finishes, fails or encounters an error. Many users tend to prefer that their automation tool maintain an extensive catalog for ordering, allowing them to reach a defined state regardless of any variance in environmental conditions.
2. **Nascent Windows Support:** Ansible supports both Unix/Linux and Windows nodes. For Windows nodes it uses native powershell remoting (as opposed to SSH), and a Linux control machine is still required for managing Windows hosts. Ansible is still early in its efforts to extend support for Windows.
3. **Minimal Enterprise Support Experience:** Ansible has had considerably less experience working with large enterprises than competitors like Chef and Puppet.
4. **Relatively New:** Ansible has not been around as long as competitors such as Chef or Puppet. As a result, it has the smallest developer/user community and has the least easily accessible material on the internet for self-help and troubleshooting. Less time on the market means more bugs, edge scenarios, and software issues have perhaps not yet appeared.

3 Fuel

3.1 Components

3.2 Architecture

3.3 Features

3.4 Official Website

3.5 Latest Version

3.6 Advantages

3.7 Disadvantages

4 Puppet

4.1 Components

4.2 Architecture

4.3 Features

4.4 Official Website

4.5 Latest Version

4.6 Advantages

4.7 Disadvantages

5 Compass

5.1 Components

5.2 Architecture

5.3 Features

5.4 Official Website

5.5 Latest Version

5.6 Advantages

5.7 Disadvantages

6 References **NEED TO FORMAT PROPERLY**

<https://shadow-soft.com/chef-benefits/>

https://www.tutorialspoint.com/chef/chef_overview.htm

<https://www.linode.com/docs/applications/configuration-management/beginners-guide-chef/>

<https://www.simplilearn.com/what-is-ansible-article>

<https://cloudacademy.com/blog/ansible-aws/>