Using Terraform to Deploy Hadoop Infrastructure

Presented By Michael Arnold

CLAIRVOYANT

Agenda

- About the speaker
- What
- How

- Principal Systems Engineer/Consultant
- Automation geek
- 20 years in IT - 8 of those years with Hadoop
- I help people deal with:
    - Servers (physical and virtual)
    - Networks
    - Server operating systems
    - Hadoop distributions
    - Making it all run smoothly

What is Terraform?

- A tool for building infrastructure as code.
- A declarative language.
- Allows for visualizing changes that will be made.
- Automatically determines resource dependencies.
- Can build servers/networks/databases in the cloud.
- Can change DNS records, create git repos, configure monitoring services, or generate TLS certificates.

Resources

- Providers
  - Services which provide an API to which Terraform can communicate.

Providers

Resources

- Provisioners
    - Scripts or commands that can be run to bootstrap a resource.

Provisioners

connection
salt_masterless
remote_exec
null_resource
local_exec
file chef

Commands

- `terraform init    # only needed once`
- `terraform plan`
- `terraform apply`
- `terraform destroy`

Demonstration

DEMO Time!

Other Commands

- `terraform graph`

Code

- https://github.com/teamclairvoyant/terraform-hadoop-talk
- https://github.com/teamclairvoyant/hadoop-deployment-bash

- https://www.terraform.io/

Michael Arnold

@hadoopgeek

www.linkedin.com/in/michaelarnold

www.clairvoyantsoft.com