OpenStack Ironic - Multi-Tenant Baremetal Deployments

Paul Browne (pfb29)
Research Computing Platforms
23 October 2019

Research Computing Services - Background

- Research Computing Services (RCS)

- Operate several large HPC clusters for diverse research computing uses.

- Current compute operated :
  - 1152 Skylake x86 nodes
  - 384 Knights Landing x86 nodes
  - 90 x4 nVidia Tesla P100 GPU nodes
  - Several mini-clusters for departmental users
  - 3 production OpenStack clouds

More hardware coming down the line in Q4 …

# Baremetal Provisioning - Background

Heavy user of xCAT2 for current BM node deployments.

Xtreme Cluster Administration Tool (sic):
- IBM project for deploying clusters of all sizes (10K+)
- Uses inventory DB source-of-truth to key off creation of:
  - Host objects
  - DNS records
  - DHCP leases + PXE records
- PXE boot to stateful/disk or stateless/memory images.
- Use a lightweight ramdisk image for provisioning ("genesis").

But:
- Well-suited for large, relatively static clusters (HPC, render or build farms, etc).
- Not as dynamic as we might like.
- Not as multi-tenant as we might like
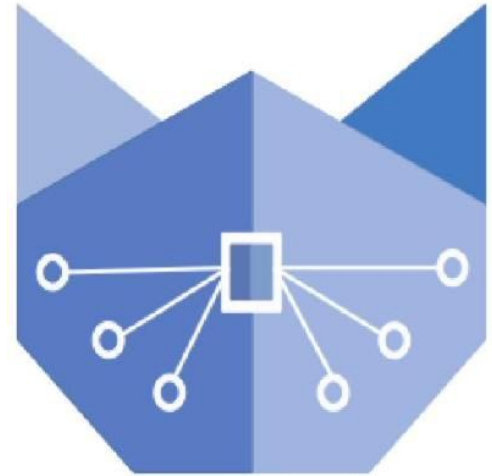
Isn't it Ironic?

- Ironic : The BM provisioning service in OpenStack.

- Similar model to xCAT2:
    - Manage nodes via IPMI/PXE
    - Boot a lightweight image for node deployment tasks

- To end users, BM "instances" can be booted by the same Nova compute service that would (usually) boot their virtual machines.

- Iron is abstracted away; The user just sees a compute instance.
    - User consumes same Nova quotas as VMs

- VMs and BM can share the same networks, if required.

- Ironic and Nova integrate well with the networking service Neutron.

- Result: Boot BM nodes into dedicated networks for
    - Inspection, introspection and validation
    - Provisioning
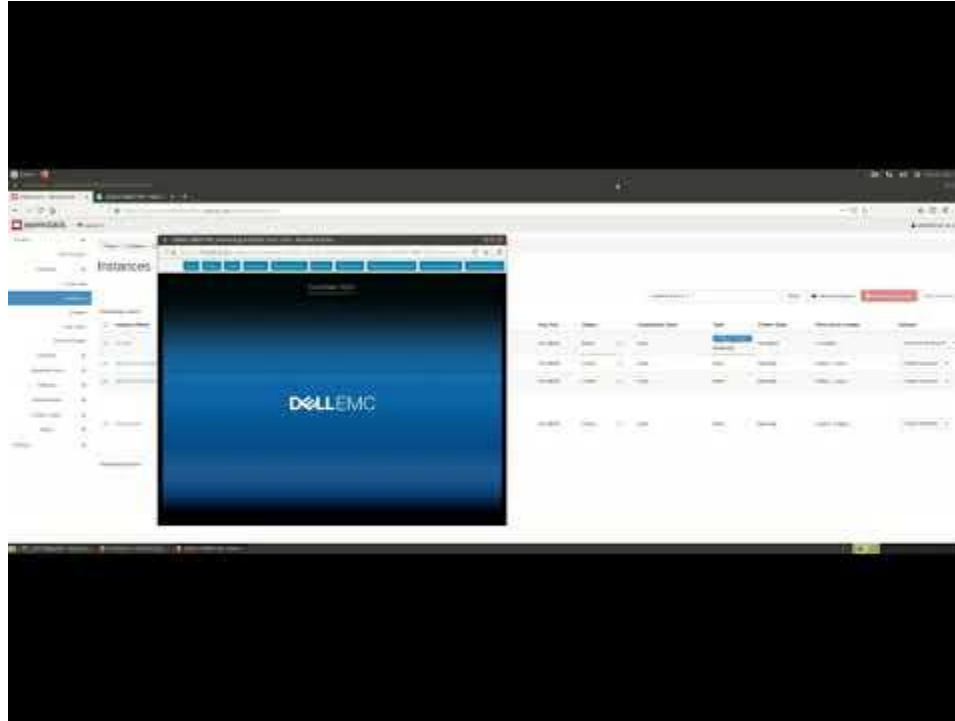   - Node cleaning
    - Finally: User workload network
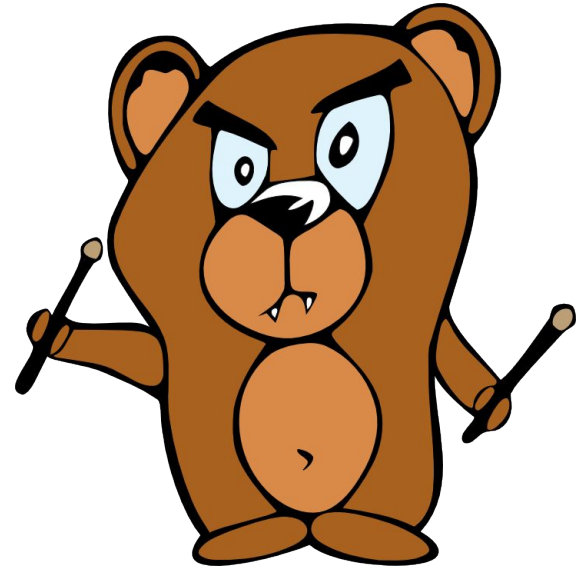
"Pixie Boots" - *Hitting servers with sticks since 2014*

Boot Up, Boot On Up! : The Ironic State Machine

Demo : BM deploy via OpenStack Horizon UI

Demo : BM un-deploy via OpenStack Horizon UI

# Infrastructure-as-Code: HashiCorp Terraform

```
## Cloud provider
provider "openstack" {
  version = "~> 1.7"
}

## Input variables
variable "provider_network" {
  default = "cumulus-internal"
  type = "string"
  description = "Network for BM instances"
}
variable "bm_image" {
  default = "centos7-image"
  type = "string"
  description = "Image for BM instances"
}
variable "bm_flavor" {
  default = "bm.v1.skylake.dirac25y"
  type = "string"
  description = "Flavor for BM instances"
}
variable "bm_keypair" {
  default = "iris-pfb29"
  type = "string"
  description = "Keypair for BM instances"
}
```

```
## Boot BM server
resource
"openstack_compute_instance_v2"
"bm_test" {
  count            = "1"
  name             = "bm_test"
  image_name       = "${var.bm_image}"
  flavor_name      = "${var.bm_flavor}"
  key_pair         = "${var.bm_keypair}"
  config_drive     = true
  network {
    name = "${var.provider_network}"
  }
  metadata {
    Description = "BM server"
    Role        = "bm_test_server"
    Managed     = "true"
  }
  lifecycle { create_before_destroy =
false }
}
```

HashiCorp
Terraform

Demo : BM deploy via Terraform and OpenStack APIs
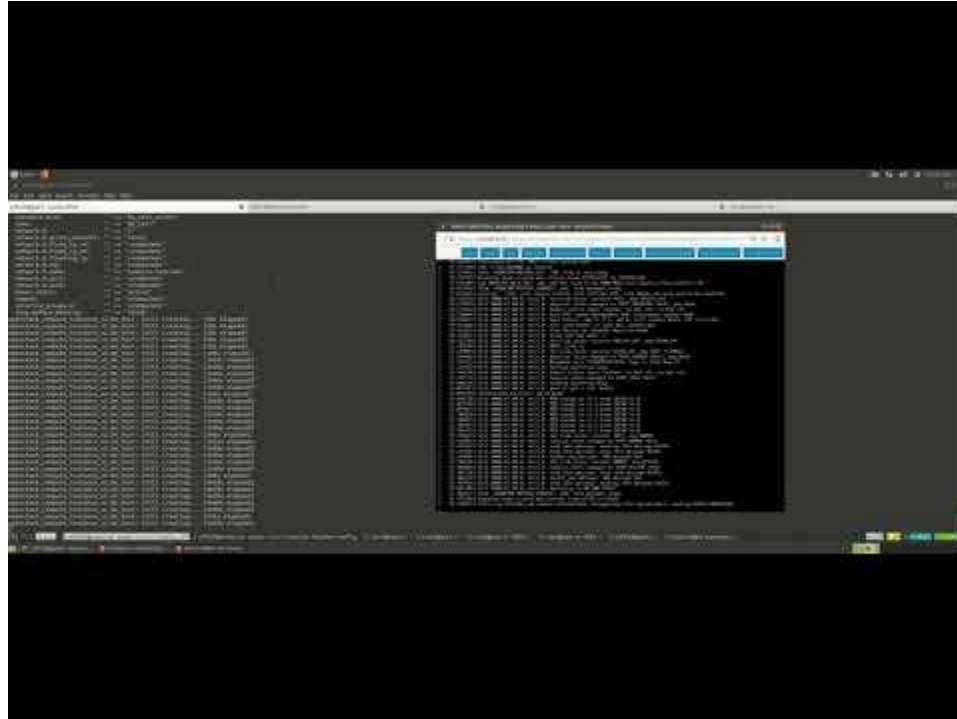
Demo : BM un-deploy via Terraform and OpenStack APIs
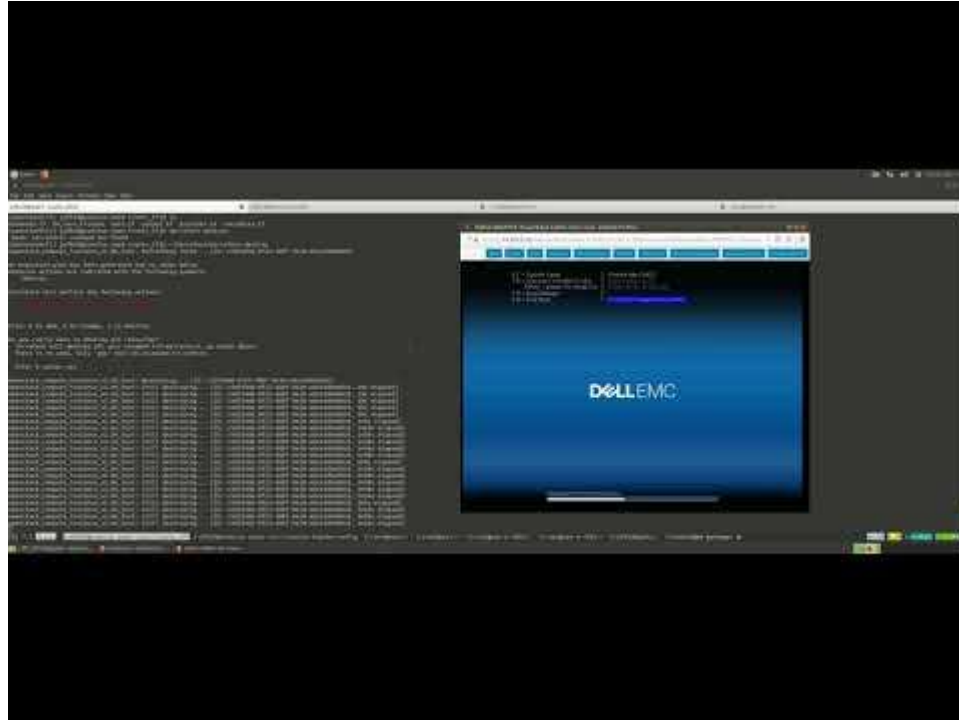
Summary

Why BM provisioning via OpenStack Ironic?

- OpenStack offers us a "single-pane-of-glass" and common APIs for managing:
  - VMs
  - Bare metal machines
  - Container hosts

- Strong network isolation and multi-tenancy built-in
  - Mandatory cleaning of nodes between uses (disk wipes, firmware re-flashes)

- Consolidate inventory accounting and book-keeping
  - Machine re-assignments easier to track, in a more dynamic BM cluster

- Enabling new use cases:
  - Bare-metal container orchestration clusters
  - Serve research uses not met by HPC clusters

MAGNUM
an OpenStack Community Project

Sponsored Content:

- **Unmanaged research VMs/cloud private networks + Ceph block/object storage**
  - Infrastructure available on 1/3/5 year leases via:

    www.hpc.cam.ac.uk/cloud-resources

- Limited test environments available for trying out the environment / learning

- Next OpenStack training dates on how to use the unmanaged cloud service

  - **December 5th 2019**
  - **March 19th 2020**
  - **July 9th 2020**

  Research Computing: Infrastructure as a Service:
  https://training.cam.ac.uk/event/3173728