

Cloud Computing on FreeBSD

Orchestrating jails using pot and nomad

FreeBSD DevSummit – EuroBSD 2019
20190918 – Lillehammer

pizzamig@FreeBSD.org
luca.pizzamiglio@trivago.com

Cloud computing on FreeBSD

- **What type of Cloud Computing?**
- **Nomad and pot**
- **Pot and FreeBSD**
- **Demo (?)**
- **Future work**
- **QA**

Cloud Computing

- **Cloud computing is a generic definition for:**

“on-demand availability of computer system resources”

- **I'll focus only on:**

“container-orchestration system for automating application deployment, scaling and management”

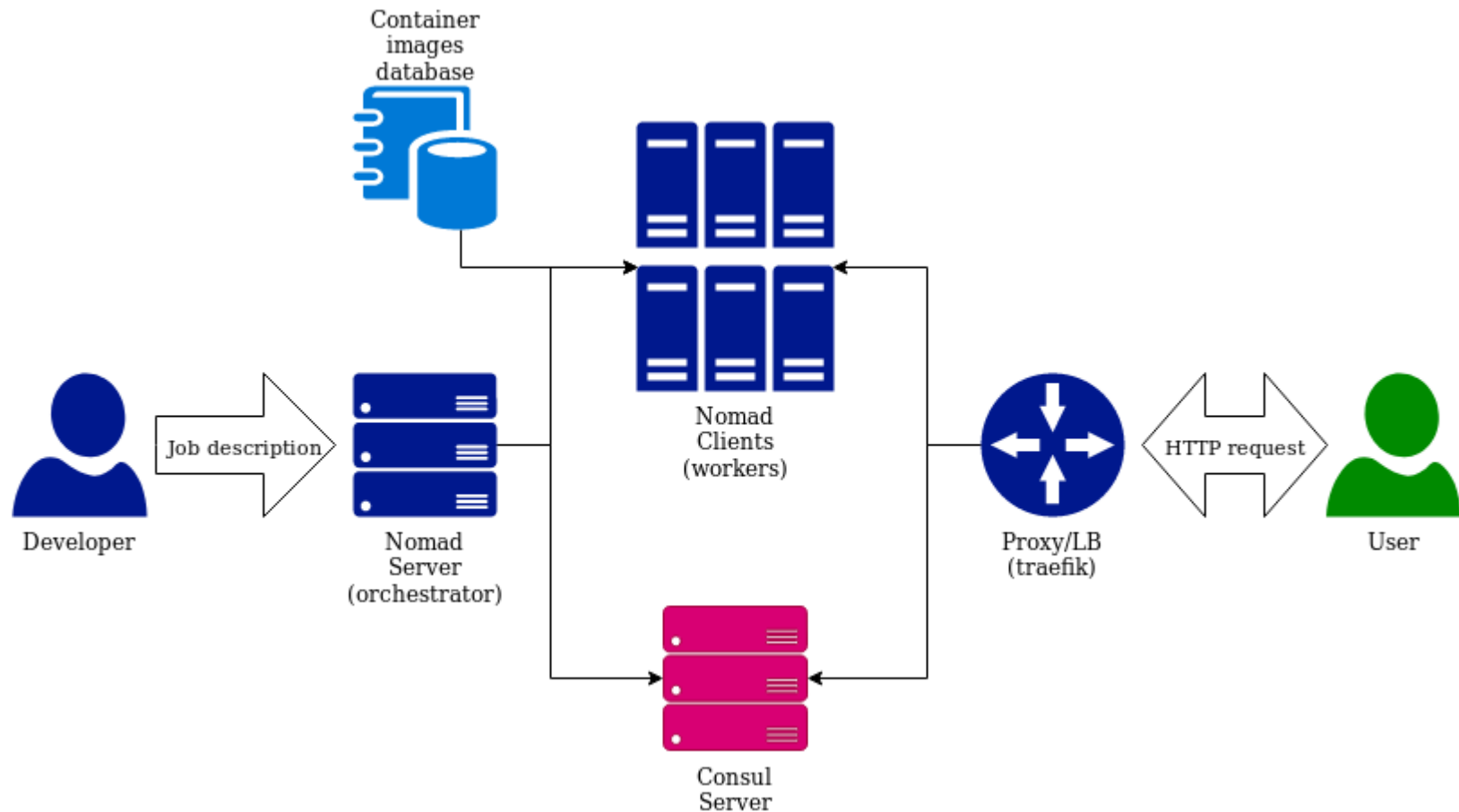
- **Used technologies:**

- pot (jail abstraction framework)
 - FreeBSD only (jails, ZFS, pf)
- Hashicorp stack (nomad, consul, optionally vault)

Container orchestration overview



Container orchestration overview



nomad and pot

- **nomad (server and client) and consul are already working on FreeBSD**
 - nomad client has few drivers:
 - java, raw exec, isolated exec (?)
- **pot already implements several features built on top of jails(8) to provide an easier interface to interact with**
- **a nomad driver has been developed by Esteban Barrios to allow nomad clients to spawn jails**
 - On github: <https://github.com/trivago/nomad-pot-driver>

Job description

```
job "example" {  
  datacenters = ["dc1"]  
  type        = "service"  
  group "example-group" {  
    task "nginx-pot" {  
      driver = "pot"  
      service {  
        tags = ["nginx-jail"]  
        name = "pot-webserver"  
        port = "http"  
        check {  
          type   = "tcp"  
          name   = "tcp"  
          interval = "5s"  
          timeout = "2s"  
        }  
      }  
    }  
  }  
}
```

```
config {  
  image = "https://pot-registry.zapto.org/registry/"  
  pot = "FBSD120-nginx"  
  tag = "1.0"  
  command = "nginx -g 'daemon off;'"  
  port_map = { http = "80" }  
  copy = [  
    "/www/index.html:/usr/local/www/nginx-dist/index.html",  
    "/www/nginx.conf:/usr/local/etc/nginx/nginx.conf" ]  
  mount = [  
    "/www/images:/usr/local/etc/www/nginx-dist/images" ]  
  }  
  resources {  
    cpu = 200  
    memory = 128  
    network {  
      mbits = 10  
      port "http" {}  
    }  
  }  
} } } } }
```

nomad and pot

- **Currently only 3 network setup supported**

- Host (host network stack)
- Public-bridge (internal virtual network based on VNET)
- Alias (typical jail setup)

- **pot has been extended to provided needed features like:**

- import/export images
- not returning exec.start command

- **one direct contribution to FreeBSD so far**

- R334094 : Improve MAC address uniqueness on if_epair(4)

pot images and the registry

- **a pot image is a compressed ZFS snapshot**
 - Create a pot
 - Provision it (manually or automatically via flavors)
 - Take a snapshot and export it
 - Upload the image to a http server
- **a registry is a http server with pot images**
- **there is no public image registry**
 - I don't want to maintain one
 - Security concerns
 - I'd keep a recipes catalog

Obstacles

- **Not forking command** ✓
- **Not persistent jails** ✓
- **Kernel panic when stopping jails with VNET** ✓
- **Redirection from localhost using external IP** ✓
- **Dynamic insertion/removal of redirection rules** ✓
- **Redefinition of cpusets**

Future development

- **Private bridge support for internal network**
- **CARP support**
- **Image size reduction**
 - Pkg base (?)
- **Multi-dataset support**
- **Move pot from sh to Rust**
- **Contribute to FreeBSD in key areas**
 - racct
 - bridge & epair
 - libpf (?)

Adoption plan in trivago

- **trivago has an internal cloud solution**
 - It would be nice to have more services based on FreeBSD
- **Plan (?) to adopt nomad-jail orchestration for:**
 - Managed services
 - Infrastructure services (high network traffic services)
- **Moonshots**
 - poudriere-nomad

Questions?

- **Questions?**
- **More questions?**

pizzamig@FreeBSD.org

<https://github.com/pizzamig/pot>

<https://github.com/trivago/nomad-pot-driver>

<https://www.nomadproject.io/>