



HPC Cloud

Workshop 2017-07-03

Ander Astudillo, Markus van Dijk

Agenda

- **Introduction** *presentation Markus van Dijk*
- **Hands-on: parts A and B**
- **Lunch**
- **Parallelism & API** *presentation Ander Astudillo*
- **Hands-on: extras**
- **Assignment** *optional*



HPC Cloud

Introduction

Cloud?

Essential cloud characteristics

- On-demand self-service
- Broad network access
- Resource pooling
- Rapid elasticity
- Measured service

What as a Service?

Cloud service models

- SaaS: Software as a Service
- PaaS: Platform as a Service
- IaaS: Infrastructure as a Service

Why *any* Cloud?

Benefits

- No hardware to buy and maintain
- No software to buy and maintain (SaaS, PaaS)
- No maintenance downtime: live migration
- Dynamic scalability – add when needed

Drawbacks

- Control over data – privacy, business secrets, legal obligations (patient data, Patriot Act)
- Control over computing – availability, processing power (SLA, overcommitting)
- Different environment – virtualization layer, VM management

SURFsara's users

SURFsara wants to offer services to **researchers**

- Scientists \neq computer gurus
- Software \neq commercial application
- Messy software dependencies
- Uptime not so important

Who uses SURFsara HPC Cloud?

Software

- Galaxy
- RStudio
- Matlab
- CFD: MPI on virtual cluster, multicore VMs
- De novo genome assembly
- Stock exchange ticker tape database

Hardware

- Multicore VMs
- Multiple VMs with private network
- GPUs
- High memory VMs
- Large data volumes

Research fields:

- Biology
- Genetics
- Informatics
- Chemistry
- Ecology
- Linguistics
- Robotics
- Business studies
- Social sciences
- Engineering
- Humanities

Why SURFsara HPC Cloud?

The SURFsara HPC Cloud is IaaS, Infrastructure as a Service,
so you build and maintain your virtual machines.

General SURFsara benefits

- Data and computing in Amsterdam
- No ties to US and its Homeland Security, Patriot Act
- Others cannot access data inside your VM (including SURFsara)
- Unrestricted Internet access, free data transfer

HPC Cloud benefits

- No overcommitting, you alone use 100% of your cores
- Tailor VM to your needs: cores, RAM, disks
- Root access to your VM
- Free choice of OS, packages, versions
- Fast private network for all VMs in your project

Why *not* SURFsara HPC Cloud?

The SURFsara HPC Cloud is IaaS, Infrastructure as a Service,
so you build and maintain your virtual machines.

Drawbacks

- You maintain everything in your VM
- You are responsible for all of your VM's behavior
- You must protect yourself against threats from the Internet (DDOS, virus)
- Pay for VM uptime, not just compute time (like gas, light)
- No automatic backups
- Your laptop is faster than a 1 core VM
- Interface to construct/start/stop VMs is not user friendly
- Service and incident response only during office hours

Alternatives

SaaS/PaaS at SURFsara

- Gina (Grid Computing)
- Hadoop, Spark
- Lisa
- Cartesius

Well known public clouds

Note: "US-EU Safe Harbor" is worthless

- Amazon Elastic Compute Cloud (EC2)
- Google Compute Engine
- Microsoft Azure
- Rackspace and other hosting providers

HPC Cloud project layout

Network

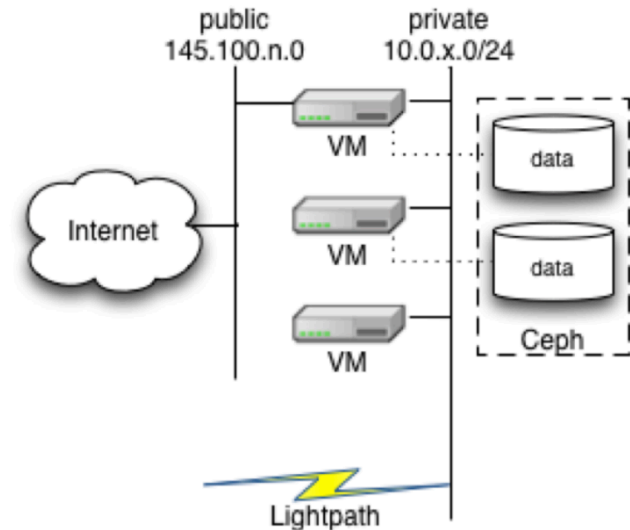
- Direct Internet access, no external firewall
- One private virtual network per project
- Fast interconnect between VMs
- IP and MAC addresses change every launch

Storage

- Host-local SSDs for OS disks (limited size)
- On-line Ceph storage for big data
- No backups, but redundant storage

External data

- Beehub, NFS on VM, iRODS, ...
- Lightpath to your institute
- Community repository on Internet



Hands-on

Support portal

<https://doc.hpccloud.surfsara.nl/>

→ Tutorials → UvA HPC Course 2017-07-03

Self-service portal

<https://ui.hpccloud.surfsara.nl/>

login: **uvaws2017-NN**