



Cloud Computing

Summer Term 2017

{Anton.Gulenko|Marcel.Wallschlaeger}@tu-berlin.de

Project Assignment No. 1

~~Due: 11.06.2017 23:59~~

Update: New Due date: 18.06.2017 23:59

The primary goal of this assignment is to gain insight into performance characteristics of virtualization using different cloud providers and to compare these values to a setup without virtualization. The secondary goal is to familiarize yourself with Amazon AWS, OpenStack, and the respective tools as an example of today's Infrastructure-as-a-Service clouds.

1. Prerequisites

1.1. Student groups

- Sign up for a group in the course website in ISIS (link under „Tutorials“). Each group must consist of 4-5 students.

1.2. Amazon AWS Educate

- Go to <https://aws.amazon.com/de/education/awseducate/> and apply.
- In order to create an Amazon AWS account and redeem a grant code, a credit card is required. Accounts can be shared using AWS Identity and Access Management (IAM) via the [AWS Management Console](#). How to use IAM is described [here](#).
- One grant is more than sufficient to successfully complete this assignment, if you **shut down** your virtual machines when you are not using them.
- AWS Documentation: Using the [Amazon EC2 online documentation](#) you can find out how to setup and use the API and AMI tools, how to work with instances etc.

IMPORTANT: Remember to shut down your virtual machines every day when you are done with them!

- The Amazon Educate grants are worth \$100 each.
- Unless you shut down your machines, the \$100 will be used up at some point and your personal credit card will be charged by Amazon!
 - **For each AWS region** check in the web-based AWS Console for running machines.
 - The AWS console only displays machines within the current region. You can change the region in the upper right corner of the screen.
- It is not possible to redeem a second grant code within the same Amazon account.
- **Define a yearly Budget** that will notify you via Email when you use more than 70% of your Amazon Educate credits. This will be graded as part of this assignment, see below.

1.3. OpenStack

The CIT group maintains an OpenStack cloud for teaching, project and research purposes. Each group will get one account with limited resource quota within this OpenStack installation. You can access the OpenStack dashboard via the following link: <https://cloud.cit.tu-berlin.de>

The connection is possible when logged into the TU Berlin Eduroam Wifi, the University LAN network or by using the [Tubit VPN](#). You will receive your credentials through ISIS.

Refer to the [OpenStack documentation](#) for details about the [web-based dashboard](#) and the [command-line tools](#).

2. Preparing Cloud Environments and Images

2.1. Prepare an AWS EC2 Instance

First, create an EC2 instance that has all the necessary benchmarking tools and packages installed to perform the assignments below. The benchmarking tools can be found in the additional materials on ISIS. Use the Amazon CLI tools to create the instance and all required resources (keypair, security group). Make sure to keep a listing of the AWS shell commands you executed, as they are part of your submission (see section „Submission Deliverables”). Remember not to include any private information such as your access key or secret key in your submission (replace them with dummy strings). Your resulting instance must...

- ...be a 64bit machine
- ...be paravirtualized (PV, not HVM)
- ...be Instance Store backed (not EBS-backed)
- ...run in the „eu-central-1” region
- ...use the instance type „m3.medium”. This is the smallest instance type available for virtual machines following the requirements defined above (see [Instance Type Matrix](#))
- ...be accessible by you via SSH (see [here](#))
- ...contain the benchmark scripts provided on ISIS (assignment1_benchmarking_tools.tar.gz) and required tools (gcc, dd, fio)

Hints:

- Shut down the virtual machine until you start working on the assignments below
- We suggest you use one of Amazon's official images: **Amazon Linux AMI 2017.03.0** minimal PV S3 (~~ami-f603d399~~ **Update: ami-f52bfa9a**). However if you want to use a different one you may do so. Amazon Linux is based on RedHat Linux and the following additional packages need to be installed for the benchmarking tools:
 - gcc, dd, fio
 - You have to install additional tools such as the EC2 API/AMI Tools and their dependencies.

2.2. Prepare a VM in OpenStack

Launch and prepare a virtual machine in OpenStack, that has all the necessary tools and packages installed to perform the assignments below. You can optionally create a snapshot of your VM if everything is set up. Keep a listing of all OpenStack commands that you execute to launch the VM (including the commands for configuring security groups, key pairs and floating IPs). You can start out by playing with the web-based dashboard, but your submission must include shell commands.

Use the following values for your instance:

- Availability Zone: Cloud Computing 2017
- Source Image: ubuntu-16.04
- Flavor: Cloud Computing
- Networks: cc17-net
- Key Pair: a key uploaded by you to make the instance reachable through ssh
- Security Groups: either default, or a custom security group. Make sure to open port 22 and ICMP traffic to make the VM pingable and reachable.

Before you can connect to the OpenStack API from the command-line, create an `openrc` file like the this:

```
export OS_USERNAME="GROUP_NAME"
export OS_TENANT_NAME="GROUP_NAME"
export OS_PASSWORD="PROVIDED_PASSWORD"
export OS_AUTH_URL="http://cloud.cit.tu-berlin.de:5000/v3"
```

Run `source openrc` to load these parameters into your environment. Afterwards you can start using the OpenStack command-line tools to create a key-pair, configure your security groups, launch an instance, and assign a floating IP to the instance. This should allow you to log into the instance via ssh (use the username `'ubuntu'` for Ubuntu-based cloud images).

3. Performance Benchmarks

Execute the benchmarks outlined below on the following three platforms:

1. Amazon AWS, using the EC2 instance you prepared
2. OpenStack, using the virtual machine you prepared
3. Your personal laptop or workstation

Tasks: Design and implement a meaningful benchmarking procedure. Describe your approach, including any written source code or scripts. Plot your results using a technology of your choice (e.g. MS Excel, [Python matplotlib](#), [Gnuplot](#)).

The benchmarks should give a good assessment of the cloud performance, so think about how external influences affect your measurements. For example, since you share resources with other users, the benchmarks should at least cover multiple times of day (e.g. day, evening, night). It might also be useful to execute the benchmarks multiple times to avoid measurement errors.

Hint: When executing long-running experiments in Amazon AWS, remember to keep an eye on your credits. You might have to shut down your VM between repeated measurements.

3.1. Disk benchmark

Benchmark sequential (Bytes/s) and random (Operations/s) read and write access to the virtualized disk. For sequential read/write use the [dd](#) tool, for random access use the [fio](#) tool. No shell script is provided for this benchmark, so you have to develop your own.

Answer the following questions:

1. Look at the disk measurements. Are they consistent with your expectations. If not, what could be the reason?
2. Based on the comparison with the measurements on your local hard drive, what kind of storage solutions do you think the two clouds use?

3.2. CPU benchmark (`linpack.sh`)

[LINPACK](#) is a popular CPU benchmark in the supercomputing community. Use the `linpack.sh` script from the `assignment1_benchmarking_tools.tar.gz` archive provided on ISIS to benchmark the CPU of the three platforms.

Answer the following questions:

1. Look at `linpack.sh` and `linpack.c` and shortly describe how the benchmark works.
2. Find out what the LINPACK benchmark measures (try Google). Would you expect paravirtualization to affect the LINPACK benchmark? Why?
3. Look at your LINPACK measurements. Are they consistent with your expectations? If not, what could be the reason?

3.3. Memory benchmark (`memsweep.sh`)

Use the `memsweep.sh` script from the `assignment1_benchmarking_tools.tar.gz` archive provided on ISIS to benchmark the memory of the three platforms.

Answer the following questions:

1. Find out how the memsweep benchmark works by looking at the shell script and the C code. Would you expect virtualization to affect the memsweep benchmark? Why?
2. Look at your memsweep measurements. Are they consistent with your expectations. If not, what could be the reason?

4. Submission Deliverables

Your submission on ISIS must be a single .pdf file with the following contents:

- A screenshot showing the budget you created in Amazon AWS that notifies you when you used 70% of your yearly budget (2 points)
- A *commented* command-line listing used to prepare the Amazon EC2 instance (5 points)
- A *commented* command-line listing used to prepare and launch the virtual machine in OpenStack (5 points)
- For every benchmark mentioned above:
 - A description of your benchmarking methodology, including any written source code or scripts (6 points per benchmark)
 - The benchmarking results for the three platforms, including descriptions and plots (6 points per benchmark)
 - Answers to the questions (2 points per question)

Total points: 62

Final Warning: Please make sure your virtual machines are shut down! :-)
