



## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

# Cloud computing for research

**Cécile Cavet**

cecile.cavet at apc.univ-paris7.fr

Centre François Arago (FACe), Laboratoire APC, Université Paris Diderot

Labhex **UnivEarthS**



January 11, 2015



# Plan

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

- 1 Cloud computing
- 2 Solutions & providers
- 3 Using the cloud
- 4 Conclusion



# What is cloud computing ?

## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References



➔ Provides IT resources on-demand.

Full definition of cloud computing: NIST report [1].



# What is cloud computing ? A mature project

- 1 **Hardware:** virtualization of all resources of commodity hardware.

## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

# The virtualization revolution...



## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

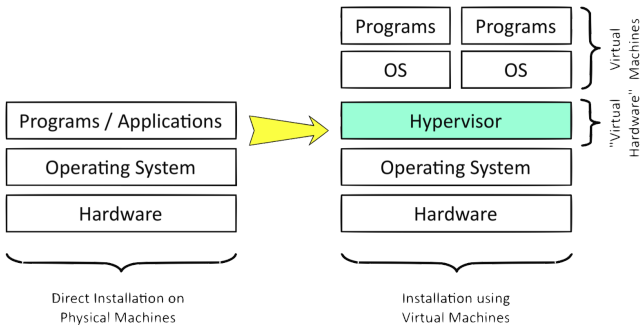


Figure : Principle of virtualization [2].



# ...and the Docker revolution

## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

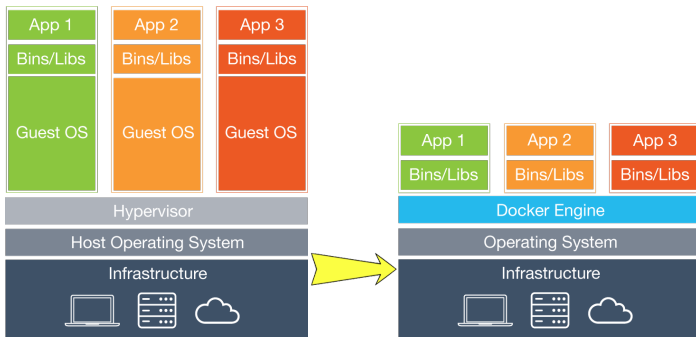


Figure : VM vs container @Docker.



# What is cloud computing ? A mature project

## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

- 1 **Hardware**: virtualization of all resources of commodity hardware.
- 2 **Software**: simplified APIs (software interface) → end-user client, Web interface and HTTP protocols.

# Simplified APIs



## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

The screenshot shows the OpenStack Horizon dashboard. The top right corner indicates the user is logged in as 'cavet' with links for 'Settings', 'Help', and 'Sign Out'. The main content area is titled 'Overview' and displays three progress bars for resource usage: 'Used 4 of 10 Available Instances', 'Used 7 of 10 Available vCPUs', and 'Used 12 800 MB of 15 000 MB Available RAM'. Below these is a 'Select a month to query its usage:' section with a dropdown menu set to 'November' and a 'Submit' button. The summary shows 'Active Instances: 4', 'Active RAM: 12GB', 'This Month's VCPU-Hours: 2259.62', and 'This Month's GB-Hours: 88877.03'. A 'Download CSV Summary' button is also present. The 'Usage Summary' table is as follows:

Instance Name	vCPUs	Disk	RAM	Uptime
cdh4-quickstart-test	2	50	4GB	2 mos
master	2	50	4GB	1 mos
worker1	2	50	4GB	1 mos
ubuntu-12.04	1	0	512MB	2 semaines, 6 jours

At the bottom, a terminal window titled 'xterm' shows the command 'nova list' and its output:

```
apoc271:~$ nova list
+-----+-----+-----+-----+-----+
| ID | Name | Status | Networks |
+-----+-----+-----+-----+
| c91c66f6-ebcb-4103-b955-09045524e68b | cdh4-quickstart-test | ACTIVE | lap=172.17.5.5, 134.158.246.54 |
| 1c88378b-f957-4f38-946c-ec2111ead5b7 | master | ACTIVE | lap=172.17.5.4, 134.158.246.130 |
| f54f8a15-602b-4bef-9463-3f94d83653de | ubuntu-12.04 | ACTIVE | lap=172.17.5.3, 134.158.246.150 |
| a2fee94a-a92b-4400-82bd-0461a0758dee | worker1 | ACTIVE | lap=172.17.5.6, 134.158.246.136 |
+-----+-----+-----+-----+
apoc271:~$
```

Figure : OpenStack API: Horizon dashboard and Nova command-line client.





# What is cloud computing ? A mature project

## Cloud computing

### What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

- 1 **Hardware**: virtualization of all resources of commodity hardware.
- 2 **Software**: simplified APIs (software interface) → end-user client, Web interface and HTTP protocols.
- 3 **Resources**: excess of **commercial** computing resources.
  - In 2006: 50 % of Amazon's resources was not used.
  - Now: growing infrastructures (academic and commercial).



# What is cloud computing ? Essential characteristics

## Cloud computing

What is it ?

### Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

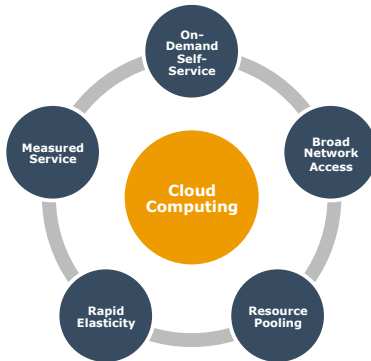


Figure : Attributes of cloud computing.

➔ **Huge flexibility for scientific applications.**



# What is cloud computing ? A clump of clouds...

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

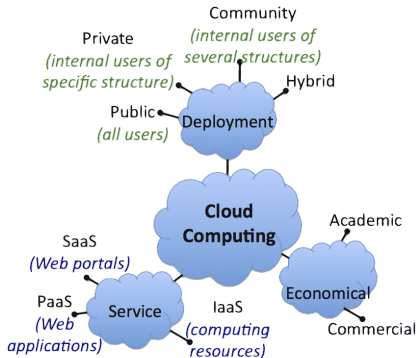


Figure : Different models of cloud infrastructure (from [3]).



# Cloud solutions

## Commercial clouds:

- **AWS**: since 2006.



- **Google**: since 2008.



- **CloudWatt, Numergy**: since 2012.

- ...

### Cloud computing

What is it ?  
Characteristics  
Models

### Solutions & providers

Solutions  
Providers

### Using the cloud

Instanciation  
SlipStream  
Use case

### Conclusion

References



# Cloud solutions: use

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

**Solutions**

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

The screenshot shows the AWS Management Console interface. At the top, there's a navigation bar with 'AWS Services' and filters for 'EC2', 'EMR', and 'S3'. The main content area is titled 'Amazon Web Services' and lists several categories of services:

- Compute:** EC2 (Virtual Servers in the Cloud), EC2 Container Service (Run and Manage Docker Containers), Elastic Beanstalk (Run and Manage Web Apps), Lambda (Run Code in Response to Events).
- Storage & Content Delivery:** S3 (Scalable Storage in the Cloud), CloudFront (Global Content Delivery Network).
- Developer Tools:** CodeCommit (Store Code in Private Git Repositories), CodeDeploy (Automate Code Deployments), CodePipeline (Release Software using Continuous Delivery).
- Management Tools:** CloudWatch (Monitor Resources and Applications), CloudFormation (Create and Manage Resources with Templates), CloudTrail (Track User Activity and API Usage).
- Internet of Things:** AWS IoT (Connect Devices to the Cloud).
- Mobile Services:** Mobile Hub BETA (Build, Test, and Monitor Mobile apps), Cognito (User Identity and App Data Synchronization), Device Farm (Test Android, iOS, and IOS Apps on Real Devices in the Cloud), Mobile Analytics (Collect, View and Export App Analytics), SNS (Send Notifications).

On the right side, there's a 'Resource Groups' section with a 'Learn more' link. Below it, there's a 'Create a Group' button and a 'Tag Editor' button. At the bottom right, there's an 'Additional Resources' section with a 'Getting Started' link and text: 'Read our documentation or view our training to learn more about AWS.'

Figure : AWS dashboard @AWS.

- **Pay-as-you-go:** Hadoop cluster of 6 standard workers is \$50 per month.
- **Free Tier instances: tests.**
- **Grant: performance benchmarks.**
- **Limitation: Patriot Act.**



# Cloud solutions

## Cloud computing

What is it ?  
Characteristics  
Models

## Solutions & providers

Solutions  
Providers

## Using the cloud

Instantiation  
SlipStream  
Use case

## Conclusion

References

## Academic clouds:

- **OpenNebula**: since 2008.



- **OpenStack**: since 2010, stable project since 2013.



- **StratusLab**: since 2012.





# Cloud solutions: use

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

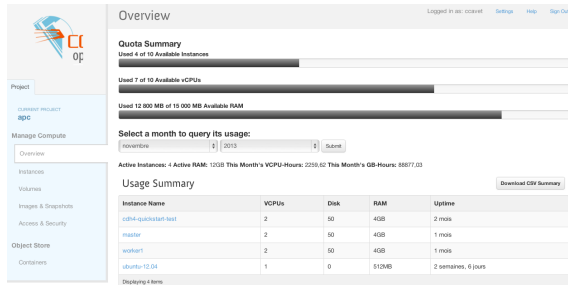


Figure : Horizon dashboard @OpenStack.

- Free.
- Performance benchmarks [4] and tests.
- Porting of scientific applications.
- Limitation: resources and services.



# French academic cloud providers

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

**Providers**

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

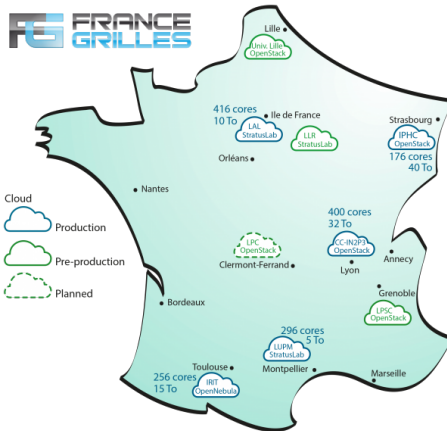


Figure : More information in [5] @France Grilles.



# How to use the cloud



## Cloud computing

What is it ?  
Characteristics  
Models

## Solutions & providers

Solutions  
Providers

## Using the cloud

Instantiation  
SlipStream  
Use case

## Conclusion

References

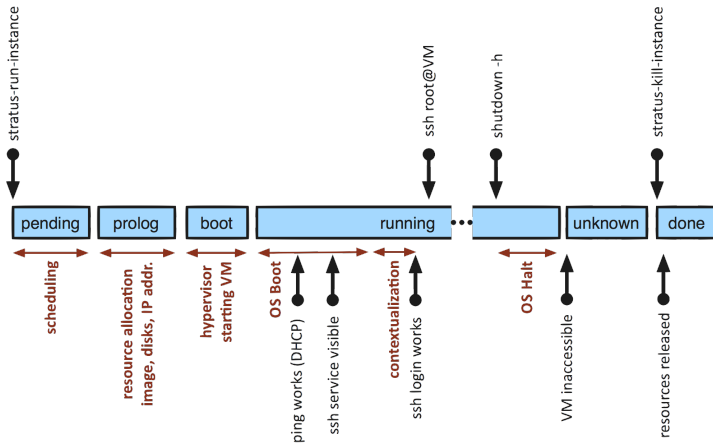


Figure : Virtual machine life-cycle [2].



# Stack of services

## Cloud computing

What is it ?  
Characteristics  
Models

## Solutions & providers

Solutions  
Providers

## Using the cloud

Instantiation  
SlipStream  
Use case

## Conclusion

References

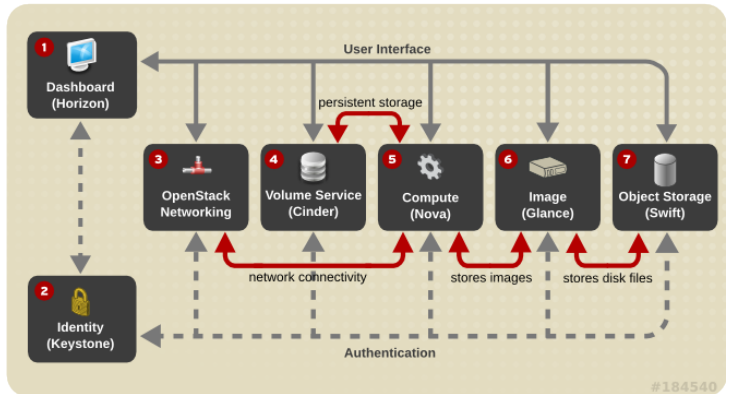


Figure : [OpenStack](#) services @OpenStack.



# Tool for the cloud: **SlipStream**



- PaaS.
- Sixsq private company  
(<http://sixsq.com/products/slipstream/>).
- Commercial connector for commercial cloud (i.e. AWS...).

➔ **Automatic deployment + Cloud infrastructure interoperability.**

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

**SlipStream**

Use case

## Conclusion

References



# Use case of APC on the cloud

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

## Scientific projects:

- **LISAPathfinder** (2015) / **eLISA** (2034) / **PTA** (2006): collaborative Virtual Machines, Beowulf cluster on-demand ➔ see next presentation.
  - **Euclid/LSST** (~2020): Hadoop cluster on-demand (Big Data prospective, see [6]).
  - **Planck** (2009 - 2013): algorithm preservation.
- ➔ **Need for Infrastructure-as-a-Service (IaaS) and Platform-as-a-Service (PaaS).**



# Conclusion

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

### Cloud is very adaptative to scientific problems:

- Easy and quick access.
- Elasticity: resources on-demand.
- Huge flexibility: adapt the environment to the code (root user, OS and libraries on-demand, similar working framework).
- Based on virtualization.
- Stack of services: [MarketPlace/Glance](#) image catalog, [Persistent disk/Cinder](#) disk manager, [Swift](#) long tem storage...



# Conclusion

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instantiation

SlipStream

Use case

## Conclusion

References

## Future:

- Growing infrastructures and community.
- Federated clouds.
- Multi-cloud solution.
- Container (Docker...) replacing VM ?





# Thank you for your attention.

## Cloud computing

- What is it ?
- Characteristics
- Models

## Solutions & providers

- Solutions
- Providers

## Using the cloud

- Instanciation
- SlipStream
- Use case

## Conclusion

- References





# References

## Cloud computing

What is it ?

Characteristics

Models

## Solutions & providers

Solutions

Providers

## Using the cloud

Instanciation

SlipStream

Use case

## Conclusion

References

End-user client installation and full command description: [FACe Wiki](#)

- [1] NIST report: <http://csrc.nist.gov/publications/nistpubs/800-145/SP800-145.pdf>
- [2] Loomis, Présentation aux Journées Cloud de Paris Sud (2012) : <http://indico2.lal.in2p3.fr/indico/getFile.py/access?contribId=0&resId=0&materialId=slides&confId=1897>
- [3] Cloud report: <http://cordis.europa.eu/fp7/ict/ssai/docs/cloud-report-final.pdf>
- [4] Cavet et al. (2012), hal-00766067: <http://hal.archives-ouvertes.fr/hal-00766067>
- [5] Airaj et al., hal-00927506 (2013): <http://hal.archives-ouvertes.fr/hal-00927506>
- [6] Cavet, JDeV (2015): [http://devlog.cnrs.fr/\\_media/jdev2015/poster\\_jdev15\\_hadooponcloud\\_cecile\\_cavet.pdf?id=jdev2015%3Aposters&cache=cache](http://devlog.cnrs.fr/_media/jdev2015/poster_jdev15_hadooponcloud_cecile_cavet.pdf?id=jdev2015%3Aposters&cache=cache)