Why to build a Cloud

Antonio Messina <antonio.messina@uzh.ch>
Cloud Architect
University of Zurich, S³IT

Who are we

S^{3}IT is a central service of the University of Zurich.

We provide solutions for researcher's data analysis usecase:

– Usecase analysis

– Solution enginneering and implementaion

– tools to run large scale data alaysis and to automate the infrastructure provisioning:
  – GC3Pie
  – Elasticluster

– Development to implement large-scale data analysis solutions

– and the infrastructure where to run it.

Researcher's FAQ

- How can I run this data analysis on 1000 cores since on my laptop is too slow? (btw, I need to submit for publication by end of this month)

- Where can I put this 100TB of data that I need to analyze and share with my colleagues? (did I tell you I have a deadline end of this month?)

- How can I automate all of this? Can you do it for me? BTW: I only know Matlab.

- Do I need to adapt my application to run on your system? Can you do it for me?

Researchers care about speed

but what is speed?

Researchers care about speed
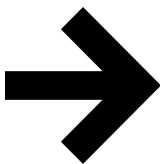
speed is *time to solution*

- actual computational time can be a small part
- not providing an infrastructure where the researcher can run means *no research*
- learning curve of adapting an application can be a blocking factor

What is a cloud?

An infrastructure to provide users with the most flexible way to allocate computational power and storage space.

- **self-provision** of resources when needed
- **customization** of the infrastructure to the use case
- **automation** of the provisioning of the infrastructure *programmatically* (via RESTful APIs)
- **scalability** of the infrastructure
- **Highly Available** infrastructure

Services a cloud can offer (1/2)

- **Compute**: start a VM somewhere

- **Block Storage**: create a block device and attach it to your VMs

- **Object Storage**: a infinite, distributed, highly available storage accessible via HTTP (with ACLs)

- **Autoscaling**: the ability of automatically spawn or destroy VMs based on triggers

- **Network**: ability to create complex network configurations in the cloud, possibly integrated cloud resources with your own network

Services a cloud can offer (2/2)

- – Network File system: an elastic POSIX network filesystem that scales on demand

- – Relational databases: create and manage a scalable relational database

- – NoSQL databases: create and manage a scalable NoSQL database

- – MessageQueue systems