SwissACC

Swiss Academic Compute Cloud

Kick-off meeting

Sergio Maffioletti
GC3: Grid Computing Competence Center
http://www.gc3.uzh.ch/

April 19, 2013

Introduction to SwissACC

Main motivations

- Sustain and consolidate community effort established in previous AAA/SWITCH projects
- Consolidate infrastructure available from various projects
- Strength link with user communities
- Expand know-how in both infrastructure operation and community support

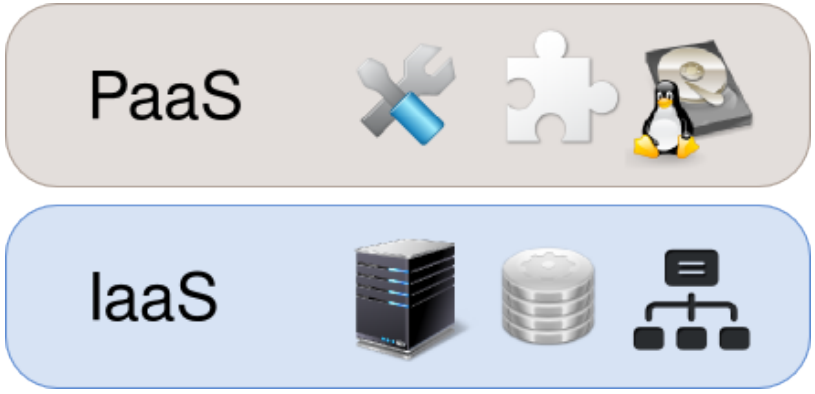
Main questions

IaaS

What could be a possible model for a data center to provide cloud-like services.

Main questions

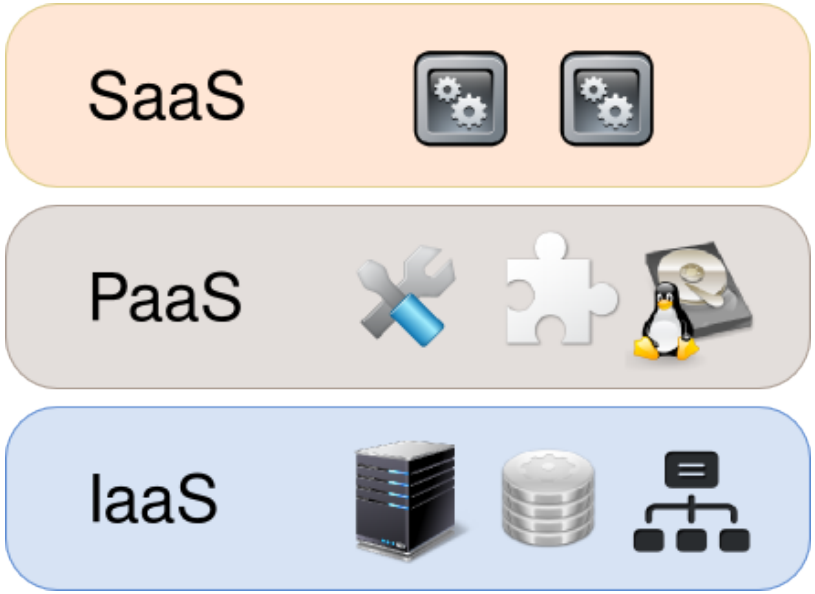
PaaS
What are the tools and services that communities needs for their large-scale data processing



Main questions

SaaS

Can we collaborate actively with both infrastructure providers and research groups for end-to-end solutions



Main questions

Can we lay down the foundation for a Swiss-wide support program based on these approaches ?

Goals

- Show a convergence in scientific computing infrastructure support and provisioning
- Establish community support model
- Allows other academic providers to include their resources as part of the infrastructure
- Use the defined usecases to show and demonstrate the feasibility of the approach

How are we planning to do this

- **Consolidate infrastructure**: pull together experience from previous projects and merge them.
- **Strength community support**: strength collaboration with research groups and establish a mediumlong term support model

Structure of the project

The project is organized around three main pillars:

- **Infrastructure**: Grid and Cloud providers.
- **Community support**: We'd like establish a collaboration model with research groups.
- **Defined usecases**: Used to demonstrate the feasibility of the approach.

infrastructure provisioning

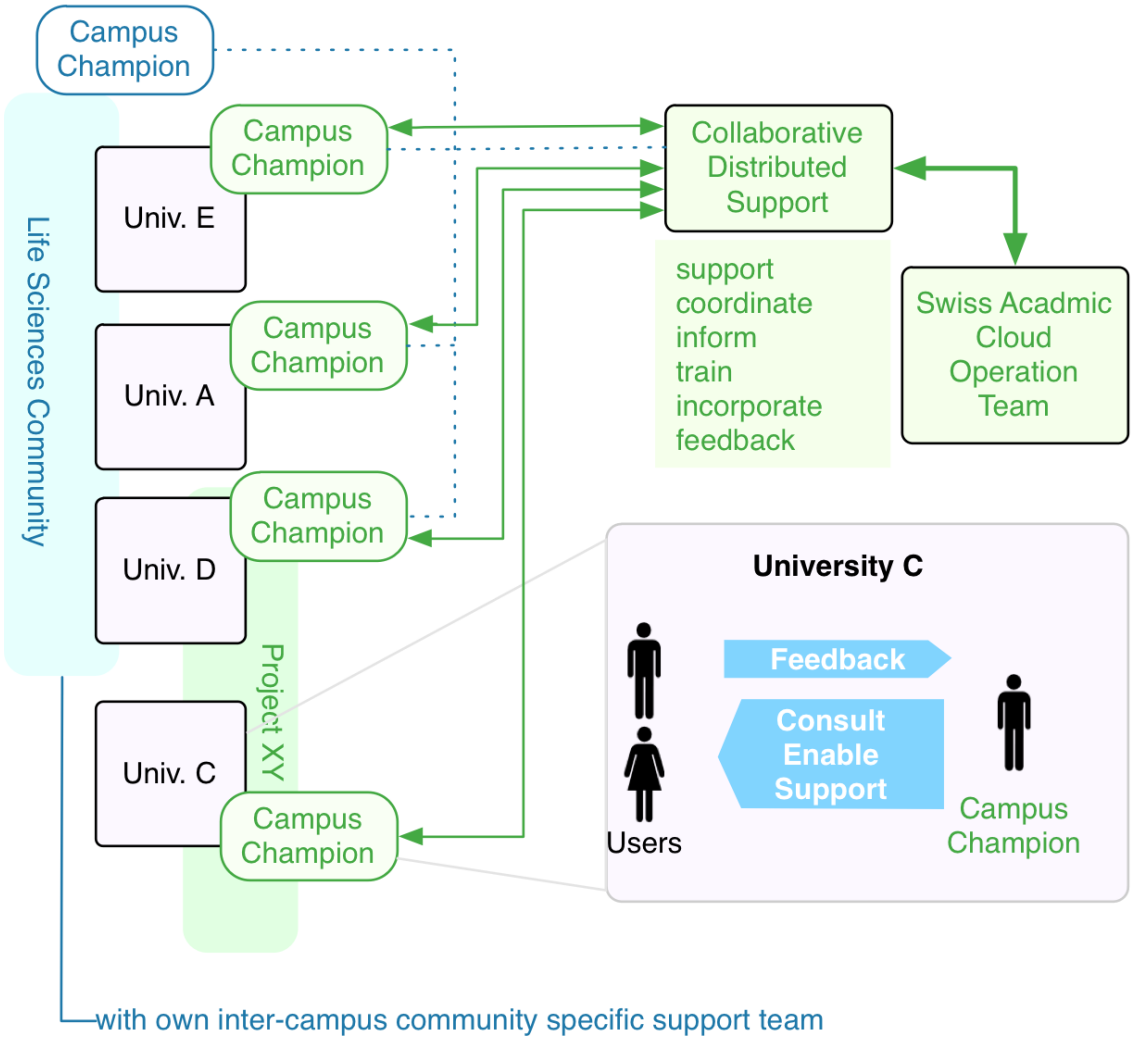
- The project is *NOT* about federated infrastructure.
- It is more about making use of available infrastructure to provide a sense of uniform support for end users.
- Infrastructure is more or less given (SMSCG, Academic Cloud, plus a plethora of related services: VM-MAD, gridcertlib, RS-NAS, elasticluster, . . .)

Community Support

- Establish local contact person(s) in each partner's institution
- Ideally should be endorsed by the local informatics service group
- Support teams will be linked together with a national representation (UZH/GC3)
  - this is just to make sure we all have means to discuss the support activities
  - and, in case, we can lever each others knowledge

Strength Community Support

Collaborative Distributed Support (CDS)



Every member institution is establishing a local support unit focused on enabling local usecases

In summary

- We are *NOT* aiming for a uniform federation of resources
- We will evaluate on a case-by-case
- Depending on the usage scenario and on the possibilities
- Support activity will become crucial (otherwise users will be presented with chaotic set of sparse resources)

But this is also a collaboration project

How will we work together ?
https://wiki.systemsx.ch/display/SwissACC/Services

Financial Flow

- Project money managed by SWITCH
- Use the AAA/SWITCH model
- Your institution's AAA/SWITCH contact person should have been already informed
- You have to contact your local contact person
- 2 trances: April/May and December
- In December 2013 we will have to make a financial report

Financial Flow

Remember...

This is project is funded with a 50% matching model. The federal contribution will have to be matched by an own contribution.

Institution	Federal funds CHF	Own funds CHF
UZH/GC3	20'000	20'000

If you have questions, we will collect them and get back to CRUS for clarifications.

GEOTop usecase

GEOTop

- Ongoing collaboration with UZH/GEO from the SMSCG project
- Already levering the SMSCG infrastructure for medium-size data analysis
- Interested in porting the application on a cloud infrastructure because of the application and post-processing requirements

GEOTop

- Allocated Manpower: 4PMs from UZH/GC3
- Milestone: WP2-GEO: GEOtop integration
  - Due to: August 2013
  - Deliverable: Appliance, GC3Pie driver script, processed dataset
- UZH/GEO will provide dataset to compute
- UZH and ETH cloud infrastructure will be used

GEOTop

Tasks

- Extend current *ggeotop* to run on a cloud infrastructure (0.5PM)
- Add usability functionality (discussed with UZH/GEO) (1PM)
- Assist UZH/GEO in porting their current codebase on an OpenSource project with continuous integration support (e.g. build system, track system, repository, unittest, . . .) (2.5PM)

TimeLine

- **May - August** used to implement tasks (4PM)
- **August - December** process datasets

GC3Pie tool

GC3Pie

https://code.google.com/p/gc3pie/

GC3Pie is a Python framework for orchestrating the execution of external commands over diverse computing resources. You write your workflow using a set of Python objects, and GC3Pie translates it into commands appropriate for the accessible computing resources. Currently, GC3Pie can execute processes on cloud-based VMs, batch-queueing clusters, computational Grids, and -of course- any Linux/UNIX host where you can SSH into.

GC3Pie

- Allocated Manpower: 3PMs from UZH/GC3
- Milestone: WP2-INT: User interface for new computational models available.
  - Due to: July 2013
  - Deliverable: Software release
- This milestone affects WP2-GEO

GC3Pie

Tasks

- Enable scalable orchestration on VM on a cloud infrastructure (1PM)
- Support most of the usecases with this approach
- Re-design GC3Pie internal structure to better scale on large datasets (2PM)
- Fully transparent for supported applications

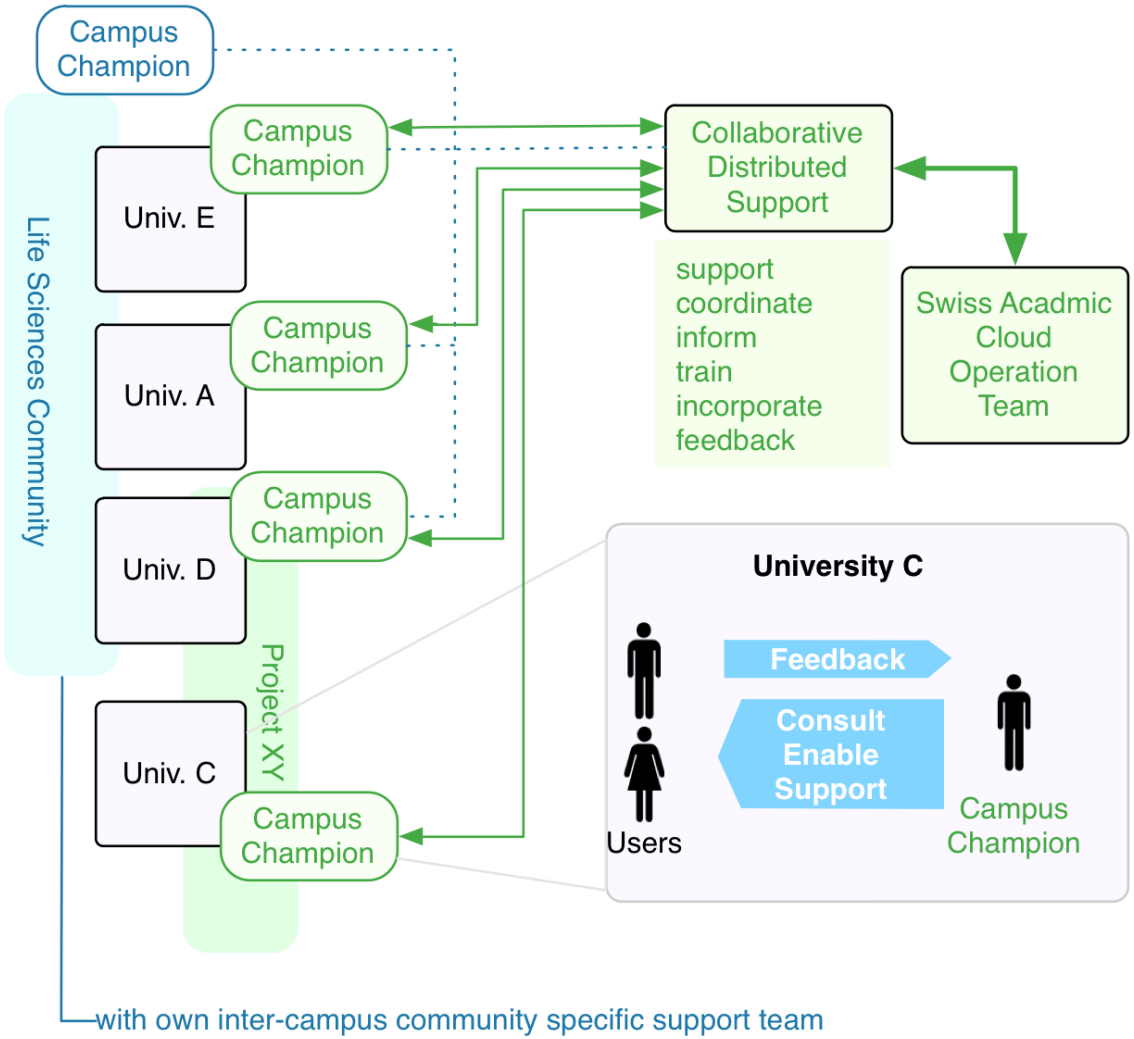
TimeLine

- **May - June** scalable orchestration of VMs (1PM)
- **August - December** Re-design internal structure (2PM)

Collaborative Distributed Support (CDS)

Strength Community Support

Collaborative Distributed Support (CDS)



Every member institution is establishing a local support unit focused on enabling local usecases

Collaborative Distributed Support (CDS)

- Distributed User Support model
- One of the main added value of this project

Main goal

Establish a collaboration model between end-users and service providers to support scientific usecases

Collaborative Distributed Support (CDS)

Assumptions

- We could help enabling/scaling scientific usecases on our infrastructure
- We may not know *everything* but we provide a forum-like approach were solutions could be discussed and reviewed (as opposite to end-users doing everything by themselves)
- This is not a new concept; Vital-IT, CSCS have been implementing a similar model already. At the international level it is an approach followed by many.

Collaborative Distributed Support (CDS)

Point to discuss in the next meeting

Switch to Wiki