

Cloud in Data Centres

Mobile Cloud Computing course
7th Nov 2012

Jarno Laitinen
Risto Laurikainen
Peter Jenkins *et al.*

CSC – IT Center for science



Today's content

- CSC and its cloud services
- Cloud middleware overview
 - OpenNebula
 - OpenStack
- Data center

Closely related topics



Date ▲	Week	Day	Time	Location	Topic
12 Sep 12	37	Wed	16:15-18:00	30T2	Introduction, Basics and Assignment 1
19 Sep 12	38	Wed	16:15-18:00	30T2	Mobile Networks
26 Sep 12	39	Wed	16:15-18:00	30T2	Open APIs
03 Oct 12	40	Wed	16:15-18:00	30T2	Business in Cloud Computing
10 Oct 12	41	Wed	16:15-18:00	30T2	Scalable Cloud Computing
17 Oct 12	42	Wed	16:15-18:00	30T2	Mobile Cloud
31 Oct 12	44	Wed	16:15-18:00	30T2	Green Cloud Computing and Assignment 2
07 Nov 12	45	Wed	16:15-18:00	30T2	Cloud in Data Centres
IaaS, Jarno Laitinen, CSC OpenStack, Risto Laurikainen, CSC Modular data center, Peter Jenkins, CSC					
14 Nov 12	46	Wed	16:15-18:00	30T2	Cloud Security
21 Nov 12	47	Wed	16:15-18:00	30T2	Industry keynote: Mobile Cloud Computing
28 Nov 12	48	Wed	16:15-18:00	30T2	Advanced topics and Summary



CSC, the Finnish IT center for science



- IT services for Finnish universities, polytechnics, some governmental institutes and companies
 - Collaborates in many [international projects](#)
- Founded 1971. Reorganized 1993 as a company, **Center for Scientific Computing Ltd.**
- Operates on a non-profit principle.
 - All shares owned by the Ministry of Education.
 - CSC's services are in general free of charge for academic end-users in Finland.
- Later look for (summer) jobs: www.csc.fi/jobs
[Interesting place](#) to work!

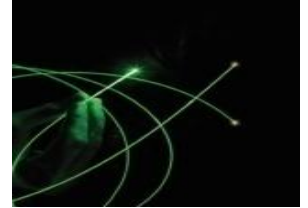


CSC's service areas



Funet network services

Fast research and education network – versatile network services



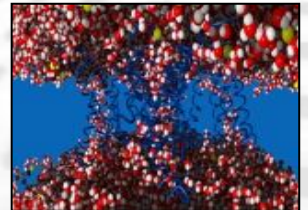
Computing services

High-performance computing and IT consulting services



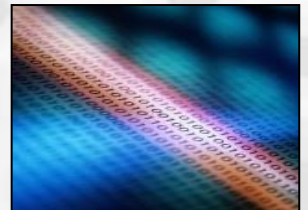
Application services

Support, development and infrastructure for computational science and engineering.



Data services for science and culture

Solutions for data storage, management and analysis needs



Information management services

Reliable, cost-effective and centralized information management services



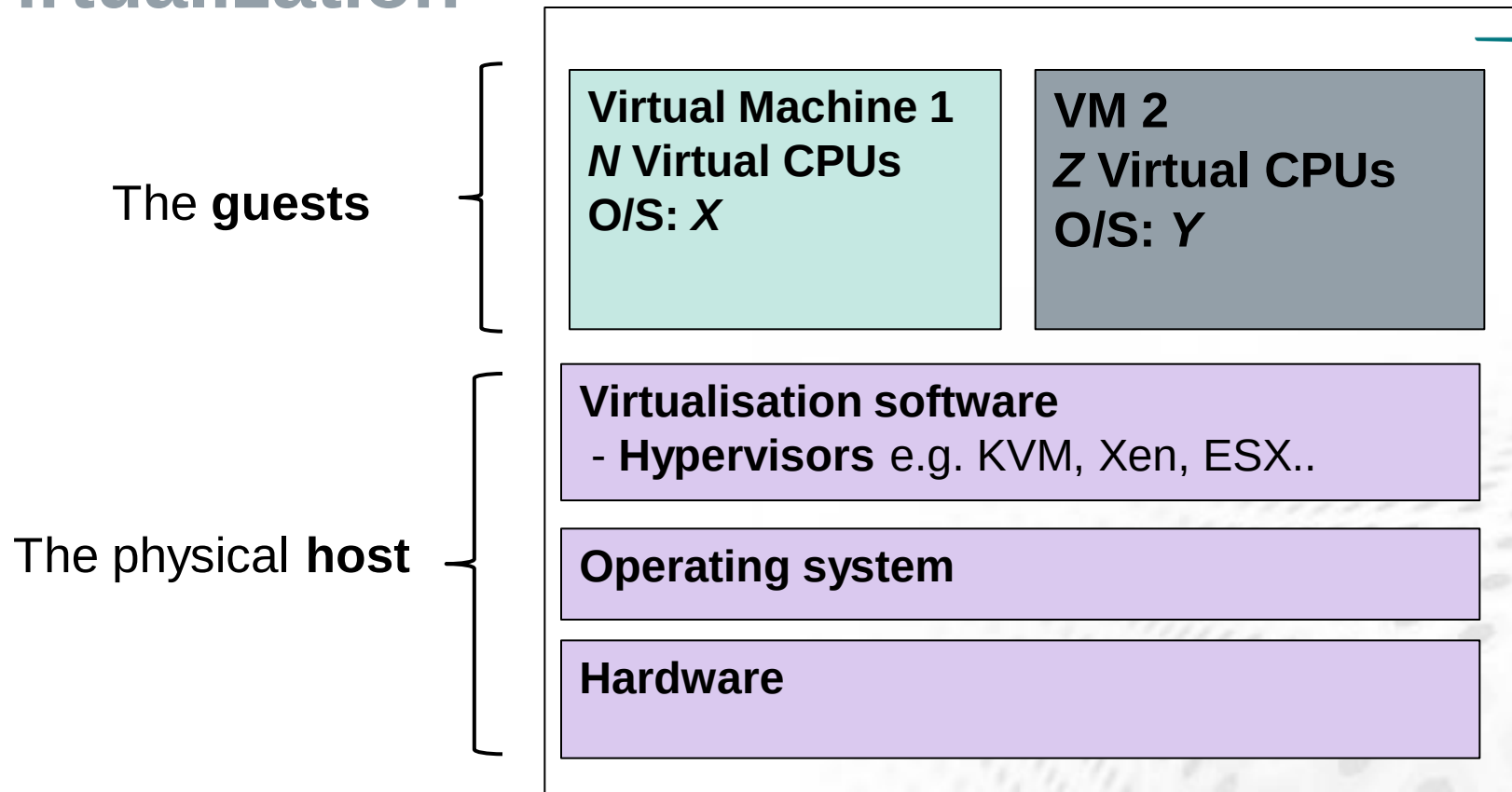
Training: www.csc.fi/courses

CSC News: <http://www.csc.fi/english/csc/publications/cscnews>

Cloud



Virtualization



- Different **virtualization technologies**

http://www.vmware.com/files/pdf/VMware_paravirtualization.pdf

- **For management** of virtualized resources: Cloud **middleware**:
OpenNebula, OpenStack, VMware, Eucalyptus, CloudStack..

Virtual Machine, image, appliance



- VM is launched from an **image file**
- A **VM can be saved to an image into repository**
- Different **image formats** raw, qcow2, vmdk, ami...
 - Cloud environment lock-in? Conversion often possible.
- **Appliance**: image containing ready to use software: firewall, application platform, applications..
 - E.g. Amazon: <https://aws.amazon.com/amis>,
<http://thecloudmarket.com/>
 - OpenNebula: <http://marketplace.c12g.com/appliance>
- **Image Marketplace** is the portal of appliance images with some description, maybe statistics, voting..

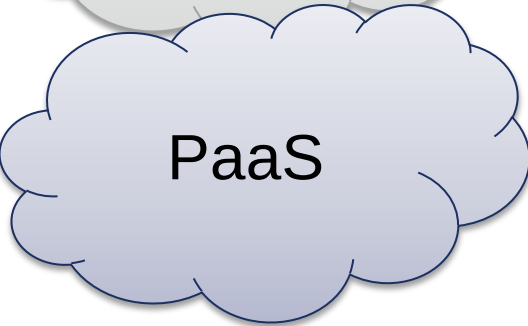
Cloud classification applied to CSC

A light gray cloud icon representing Infrastructure as a Service (IaaS).

IaaS

Infrastructure: Cloud user interface to launch Virtual Machines where the user can install choose operating system and administrate it.

Storage: either via VMs or iRODS middleware.

A light blue cloud icon representing Platform as a Service (PaaS).

PaaS

Platform: Virtual Machine images with *preinstalled software* e.g. Hadoop, cluster tools, code development tools, science discipline tailored applications.

A light purple cloud icon representing Software as a Service (SaaS).

SaaS

Software: Some applications possible to use via web UI. Integrated also e.g. file system view, job submission. Traditional way to use resources and application is ssh login.

A light pink cloud icon representing a Community cloud.

Community
cloud

Target customers are CSC's existing customers.
Not a public or a private cloud.

Cloud middleware: OpenNebula

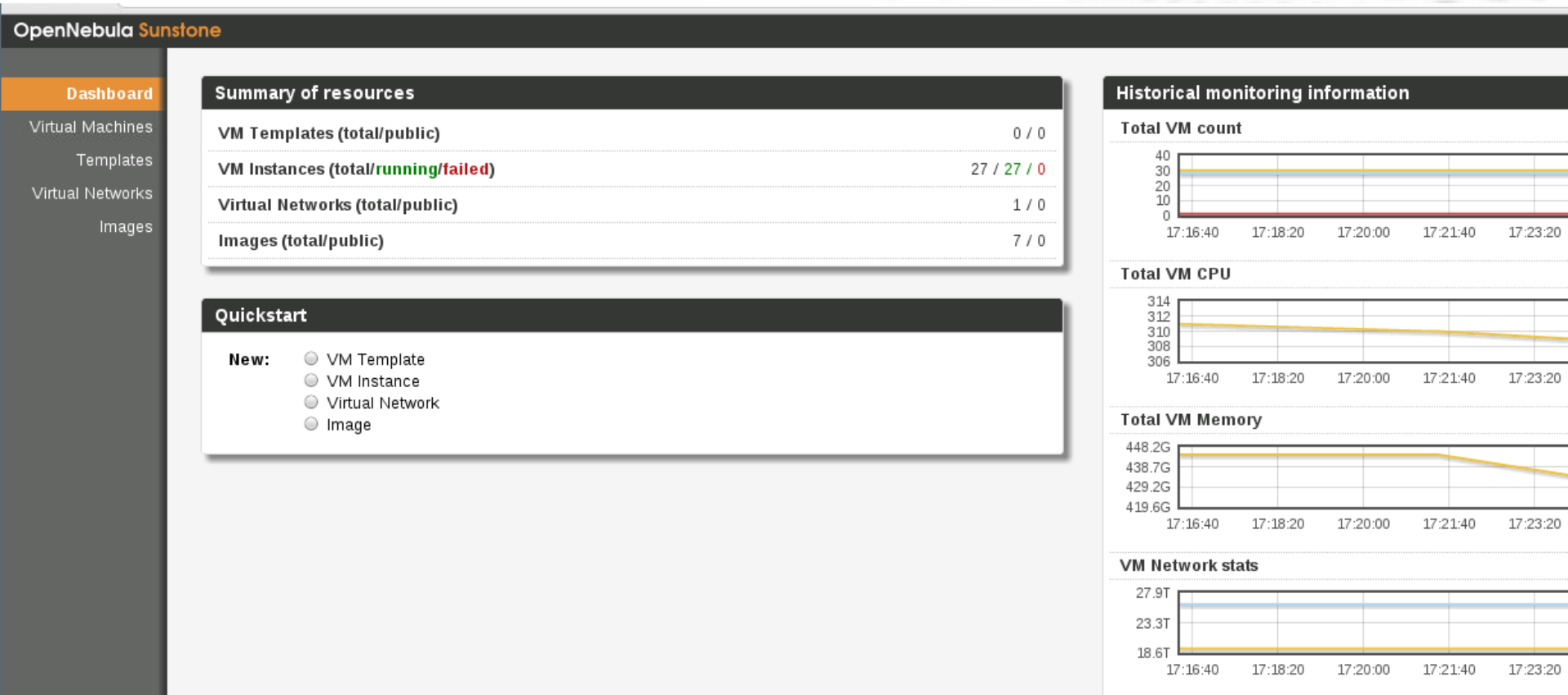
<http://opennebula.org/about:about>

<http://opennebula.org/users:users>

- First public release 2008
- OpenSource
- Project managed by C12G Labs founded by the developers.

User interfaces

- ➡ Different roles: cloud admin vs. normal user.
- ➡ Command line
- ➡ Web UI
 - Limited functionality but VNC console, log



OpenNebula commands examples



- To create a new VM:

```
onevm create <template>
```

- To list VMs:

```
onevm list ; onevm top
```

- To show details of a VM:

```
onevm show <vm id>
```

- Save a disk of a VM:

```
onevm saveas <vm id> <disk id> <image name>
```


OpenNebula demo



OpenNebula's APIs



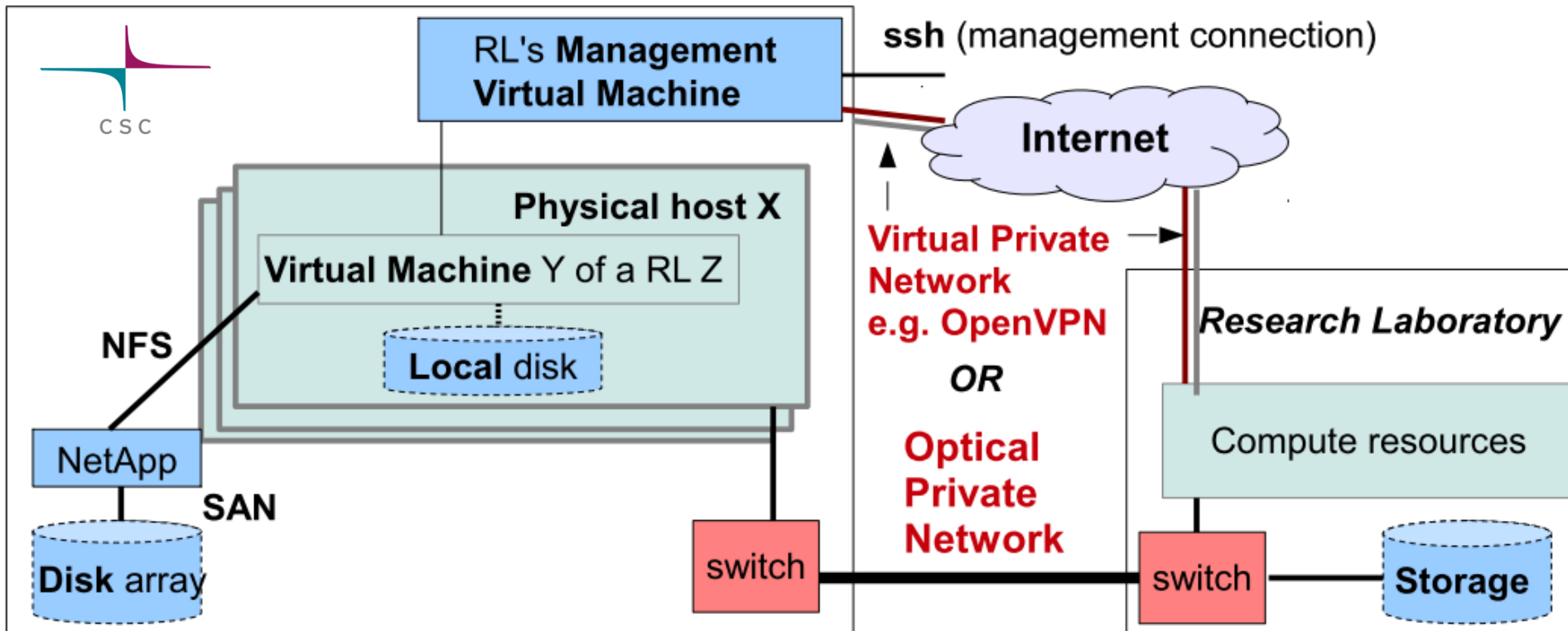
- **XML-RPC** interface OpenNebula specific.
 - <http://opennebula.org/documentation:rel3.8:introapis>
- Open Cloud Computing Interface (**OCCL**) [support evolving](#). [User Guide](#) (cli tools, XML).
- Supports KVM, Xen and VMware **hypervisors**.
- **Hybrid** cloud
 - Extend your private OpenNebula cloud into public cloud with Amazon APIs. "Cloud bursting".
 - Enable to create public cloud managed by OpenNebula
 - <http://opennebula.org/documentation:rel3.8:ec2qug>
- **Ecosystem**: various plugins, tools, documentations..
 - <http://opennebula.org/software:ecosystem>



Community cloud



- National project for **biomedical community**.
- **Res. lab's IT admins** manage. For the end-users (scientist) cloud can be "invisible".
- Cloud middleware: **OpenNebula**.
- Runs **since fall 2010**.
- Currently **~2800 cores**, up to 96 GB RAM (per 24 cores node), **350 TB NFS**.



Human factor is important



- Establishing trust, discussion of risks and security.
- Assistance in the integration of the resources and in usage.

IaaS cloud expert

- **Provides**
 - **Resources (compute, storage)**
 - **Interfaces** to access the system
- **Supports** usage of the cloud, but does not necessarily manage Virtual Machines (VM)
 - Does not know what is running on the VMs



Local IT admin

- **Can connect the existing compute / storage resources through the private network solution**
- **Manages Virtual Machines**
 - **root permission** for VMs
 - **Installs and maintains** Operating System and other software for VMs
 - Pays the software licenses

Cloud as a Science Platform?



- More flexible "build-your-own" environment.
- Cheaper price – also for small units.
 - Outsource IT infra to concentrate on doing science.
 - Scalability.
- Using image to save and share an environment.
 - Use ready environment of your science / collaborator.
- Open data trend.
 - Funder of the research might require publishing the output data.

The extra slides



Mobility in cloud?



- Compute, storage, services in the cloud – less limits.
- Allows mobility of the users, multiple end-devices.
- Data sharing and collaboration features.
- Mobile networks already reaching Mb/s rate.
- Danger: lock-in for certain interfaces and vendor.
- Challenges in network
 - secure communication (encryption)
 - firewall restrictions.
 - Speed and robustness of the mobile network?

A real mobile cloud with data center?

Case Facebook



Statistics from Oct 2012:

- 500 TB new data per day
- 7 PB photos / month
- over 1'000'000'000 users
- 600 million use a mobile device
- Some articles regarding its data centers:

<http://www.wired.com/wiredenterprise/2011/12/facebook-data-center/all/1>

<http://datacenterdesign.datacenterreports.com/facebook-and-the-open-compute-project-thankyou/>

Facebook's 3rd data center opening in 2013. Now in Sweden:

<http://gigaom.com/cloud/facebook-has-220-billion-of-your-photos-to-put-on-ice/>

References:

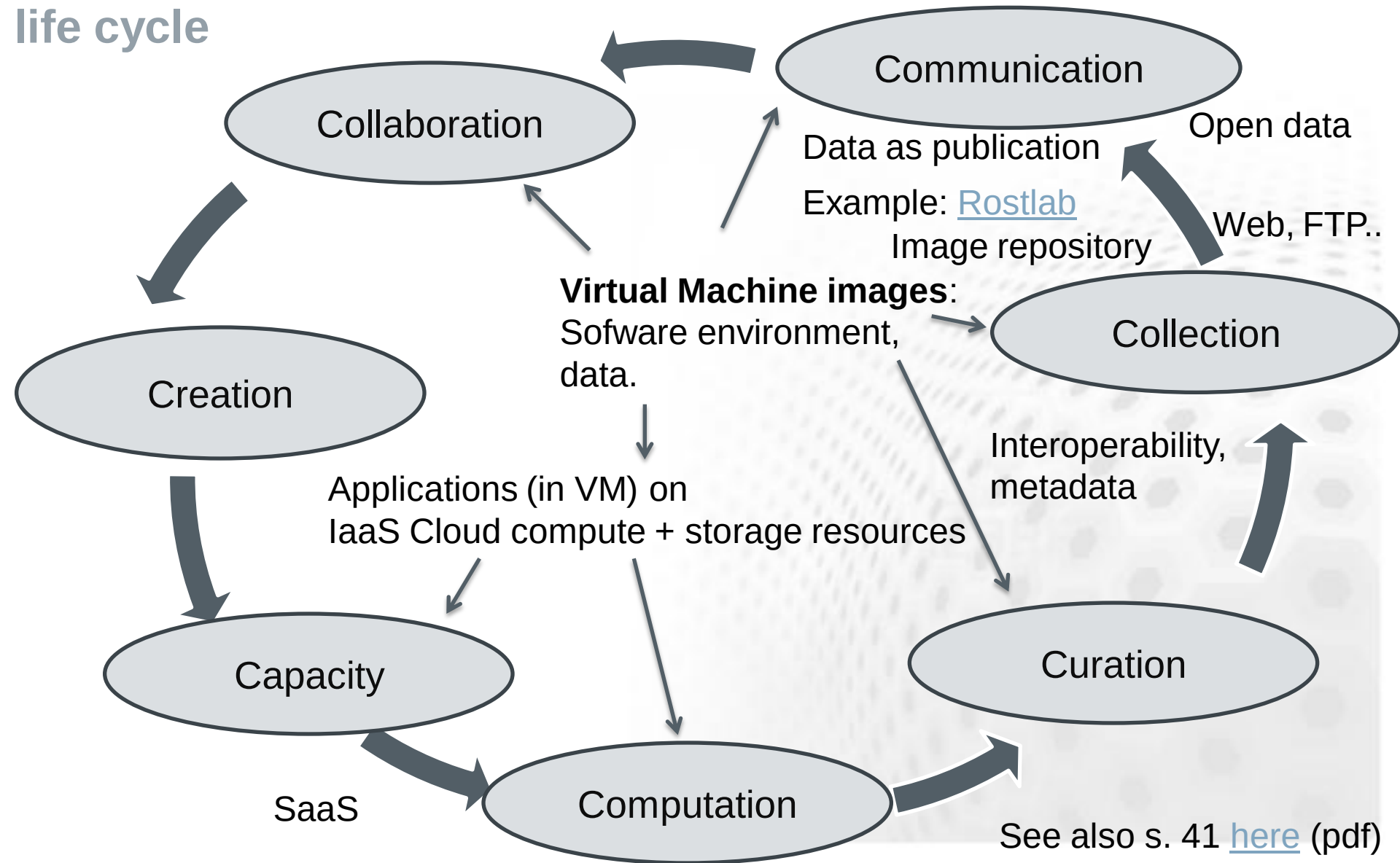


- Scientific clouds e.g. [SARA](#) (NL), [Nectar](#) (Australia)
- EU related projects and publications:
 - [StratusLab](#). Also Image Marketplace.
 - [Venus-C](#) including business partners
 - EGI [Federated clouds](#) taskforce
 - [EU](#): "Research in future of cloud computing"
- Other
 - CISCO network vendors [marketing forecasts](#)
 - [NIST](#): "Cloud Computing Synopsis and Recommendations"
 - IaaS company CloudSigma [blog](#) of outsourced data center.
 - TERENA: network conference [cloud session](#)
 - <http://www.elixir-europe.org/researchers/data-deluge>
 - www.eduserv.org.uk/newsandevents/events/2012/symposium#Agenda
 - [Cloud middleware community growth](#) [blog](#) posting.

7 C's from: Bicarregui, J. Terena conference 2012

Building an Open Data Infrastructure for Science: Policy and Practice + my quick thoughts of Cloud's role in data life cycle

CSC



Biomedinfra cluster



- **Three rack cluster. Dell C6100s as compute nodes.**
- **116 nodes with 24/48 GB RAM, 300..1000 GB disk**
- **4 hosts 96 GB RAM, 6*600 GB SAS disks**
- Intel Xeon processors, 2*12 cores, 24 with H-T.

→ **~ 2880 cores (with Hyper-Threading)**

- + 4 Frontend nodes
- **68 nodes connected with 10 GbE and 48 with 1 GbE**
- **Key issue: storage (NFS) capacity.**
- **Customer networks are separated with virtual LANs**

CSC's computing servers 2012



Louhi: CRAY XT4/XT5

- 10864 cores
- 1/2 GB memory / core
- Cray SeaStar2 interconnect
- Profile: Massively parallel computing



Hippu: HP ProLiant application server

- 128 cores
- 3 TB shared memory
- Profile: Interactive use, large memory, plenty of applications



Vuori: HP ProLiant supercluster

- 2880 cores
- 16 GPGPUs
- 1.33/2.66 GB/core
- Infiniband
- Profile: mid-range computing, GPU computing

