OpenNebula

Latest Innovations in Private Cloud Computing

Ignacio M. Llorente

OpenNebula Project Director

OpenNebula.org

This presentation is about innovation

... in cloud provisioning

... in cloud consumption

and in two cloud use cases

... HPC and science applications

... Pro-active autonomic applications

What is OpenNebula?

Simple but feature-rich, production-ready, customizable solution to build clouds

SIMPLE

Easy to operate, install and upgrade, with packages for the main Linux distributions

FLEXIBLE

Really open-source and customizable to fit into any data center and policies

ROBUST

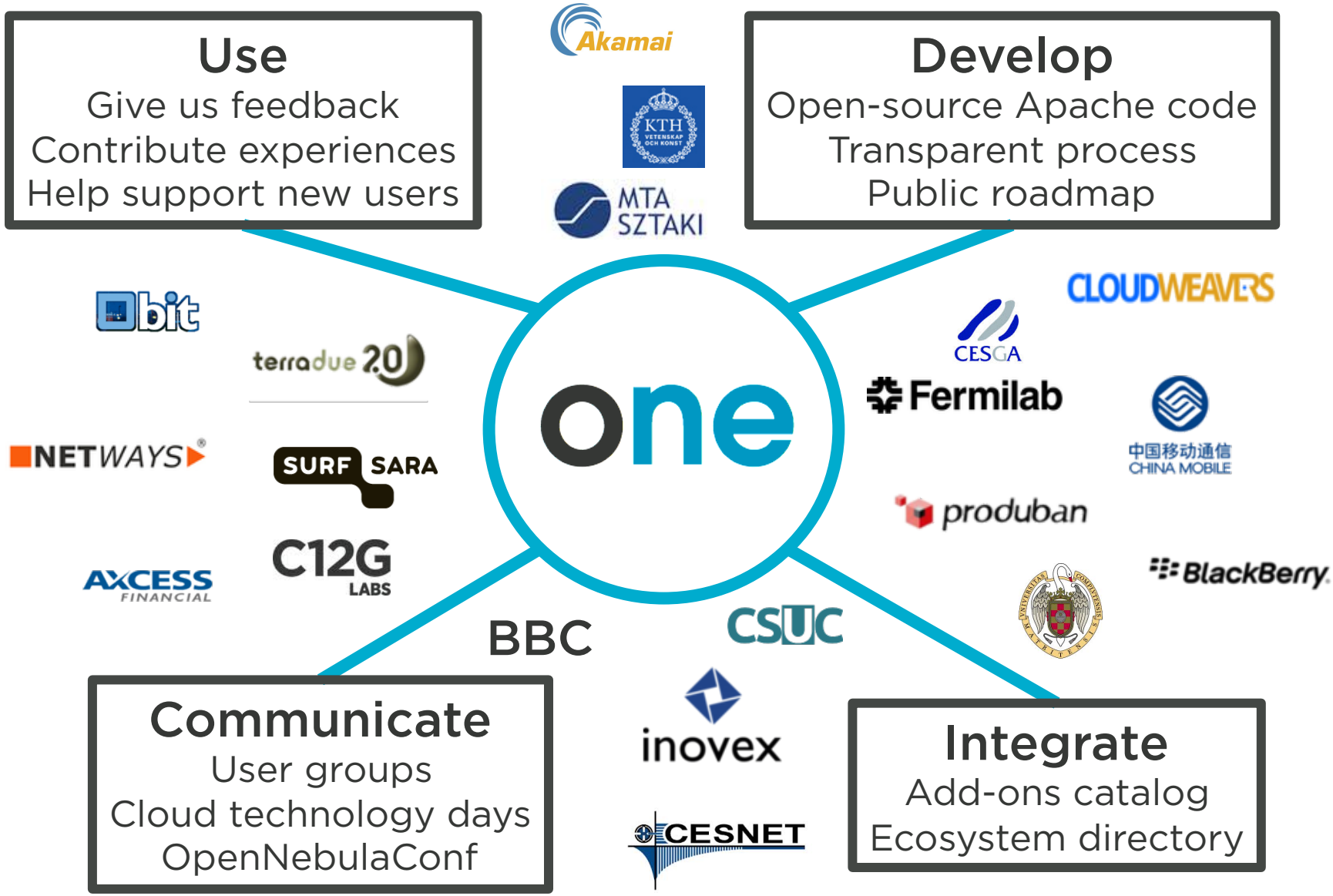
Production-ready, mature, reliable and commercially supported

POWERFUL

Innovative functionality for enterprise clouds and data center virtualization

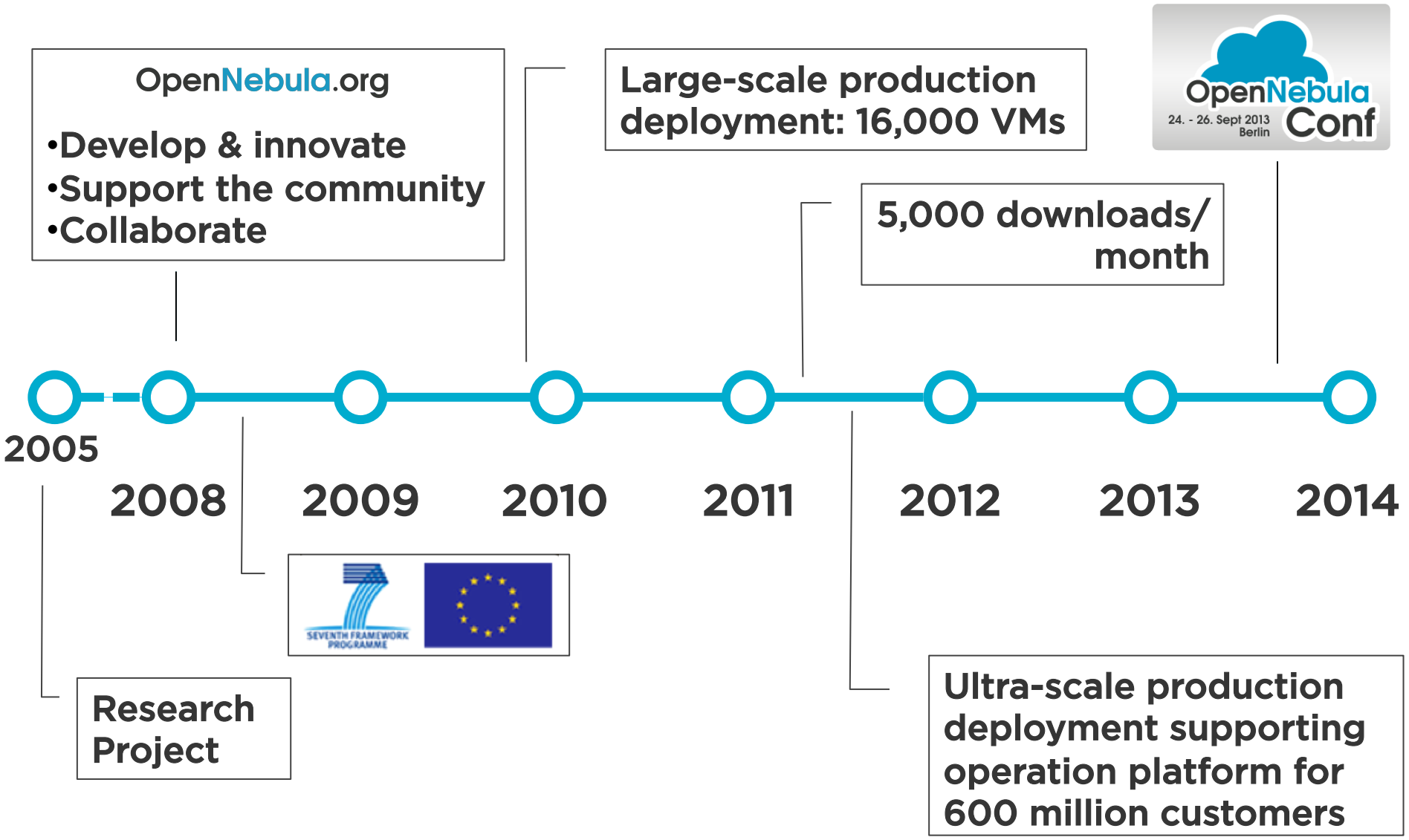
What is OpenNebula?

An Open Community Driven by Users



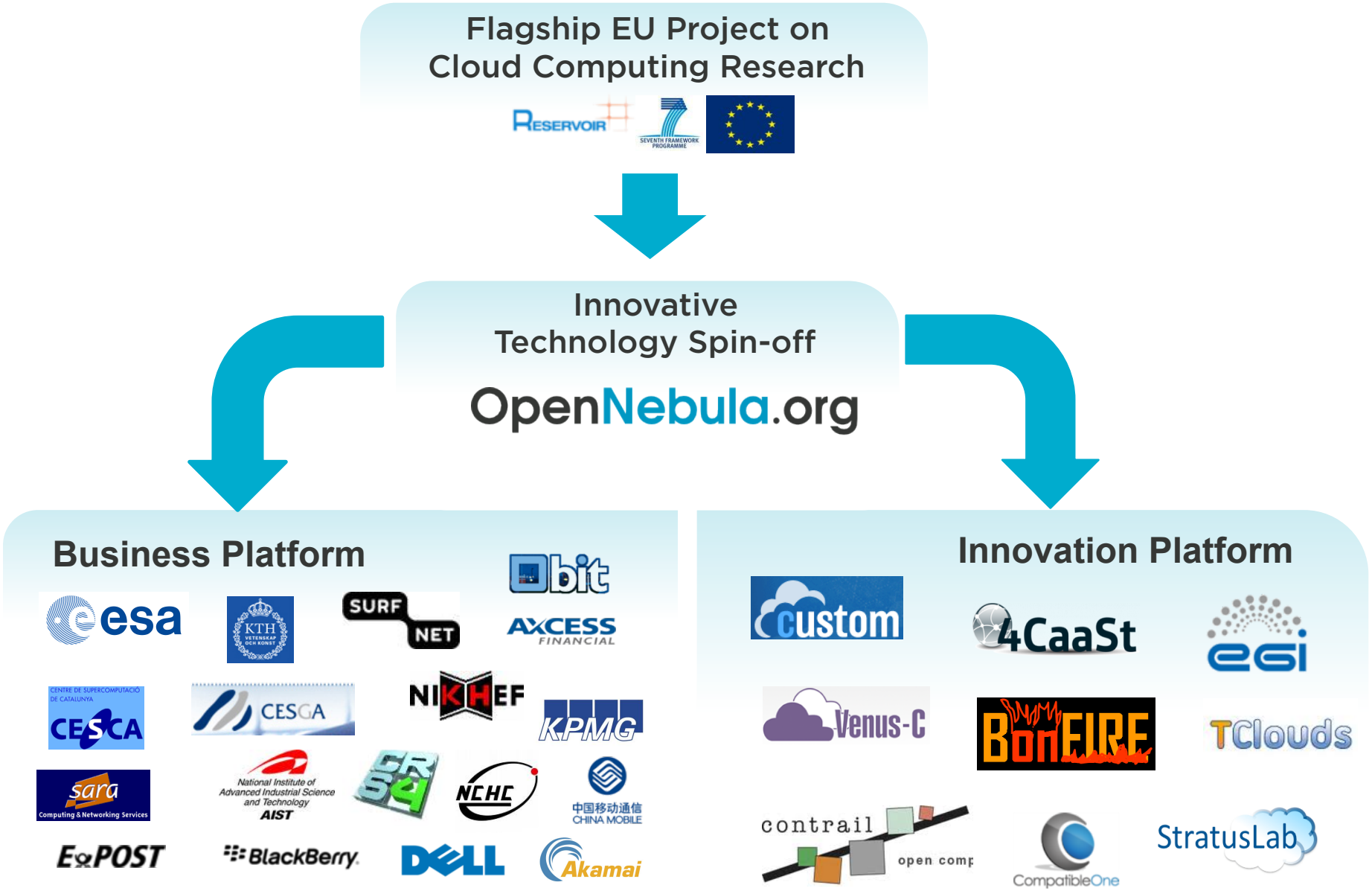
What is OpenNebula?

From Research Project to Open-source Project for Enterprise



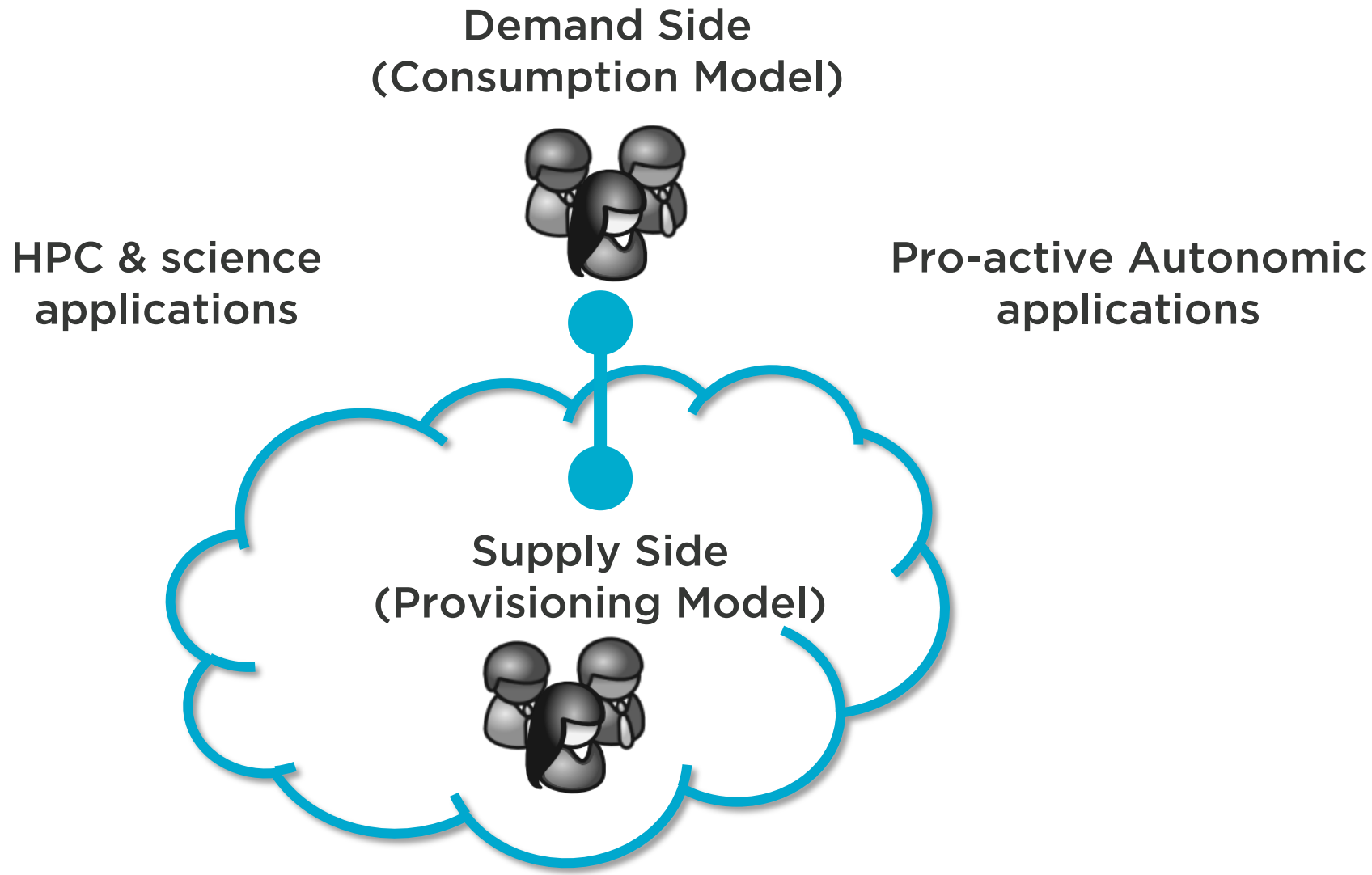
What is OpenNebula?

EU Success Story in Cloud Computing Research and Innovation



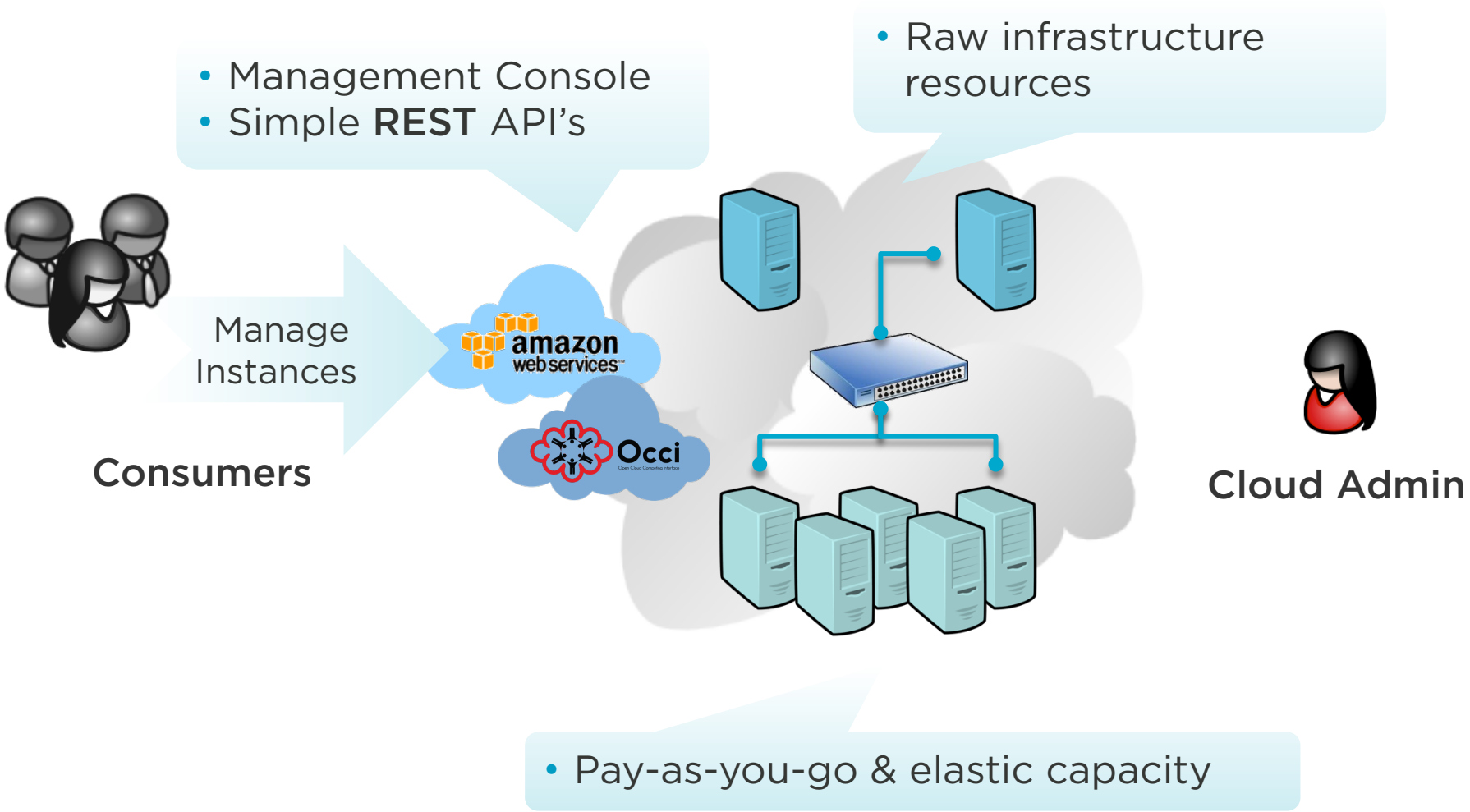
EU Success Story in Exploitation of FP7 Results

Different Perspectives to Present Innovations in Cloud Computing



The OpenNebula Provisioning Model

Existing Amazon-like Provisioning Model

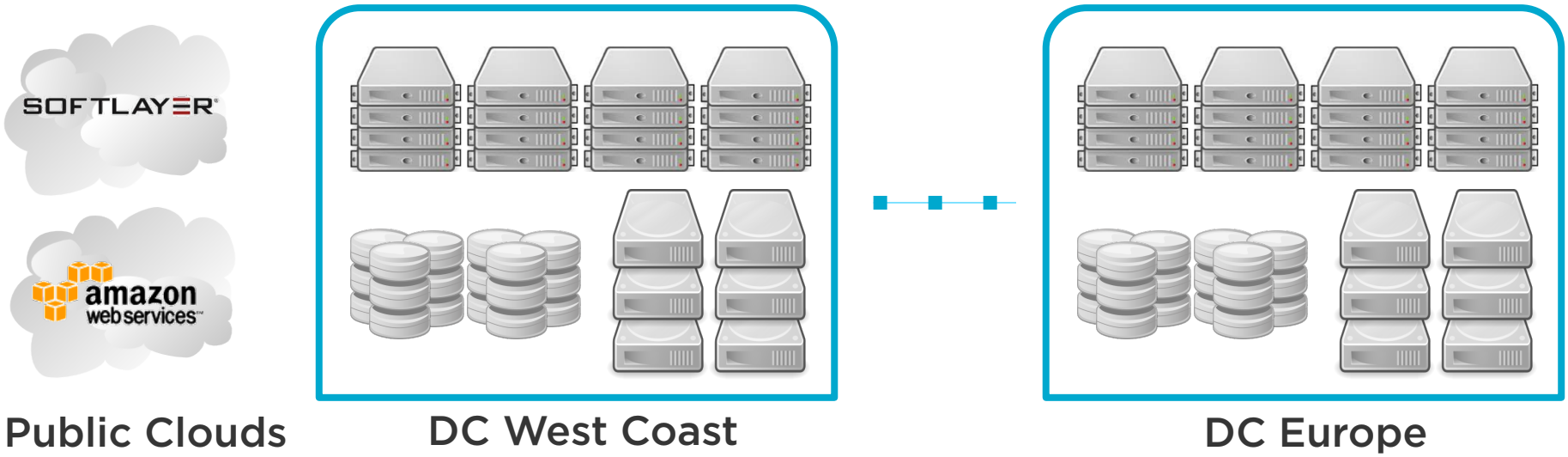


The OpenNebula Provisioning Model

Challenges from the Infrastructure Perspective

Comprehensive Framework to Manage Infrastructure Resources

- **Scalability**: Several DCs with multiple clusters
- **Outsourcing**: Access to several clouds for cloudbursting
- **Heterogeneity**: Different hardware for specific workload profiles



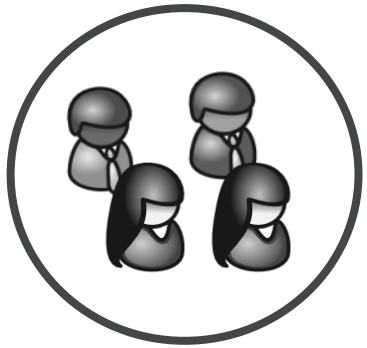
The OpenNebula Provisioning Model

Challenges from the Organizational Perspective

Web Development

HPC Simulations

Big Data Analysis



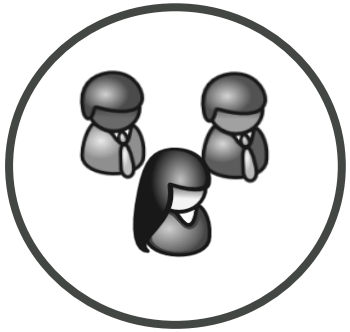
Comprehensive Framework to Manage User Groups

- Several divisions, units, organizations...
- Different performance and security requirements
- Dynamic groups that require admin privileges

The OpenNebula Provisioning Model

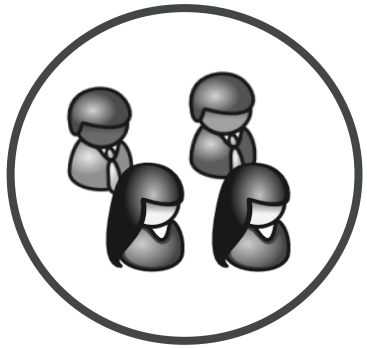
The Goal: Dynamic Allocation of Private and Public Resources to Groups of Users

Web Development

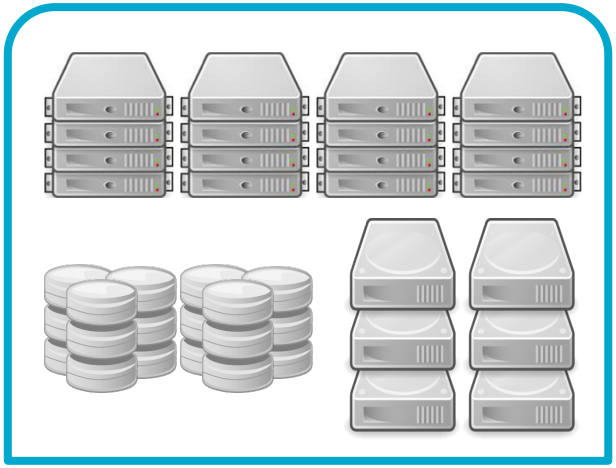


HPC Simulations

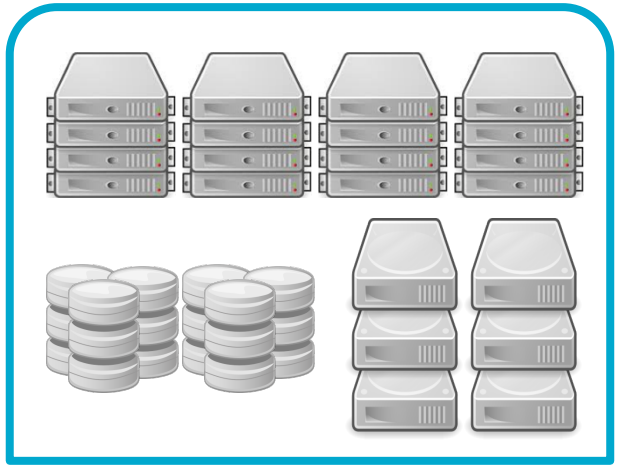
Big Data Analysis



Public Clouds



DC West Coast



DC Europe

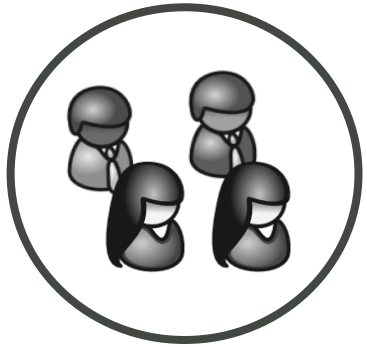
The OpenNebula Provisioning Model

Definition of Clusters (Resource Providers)

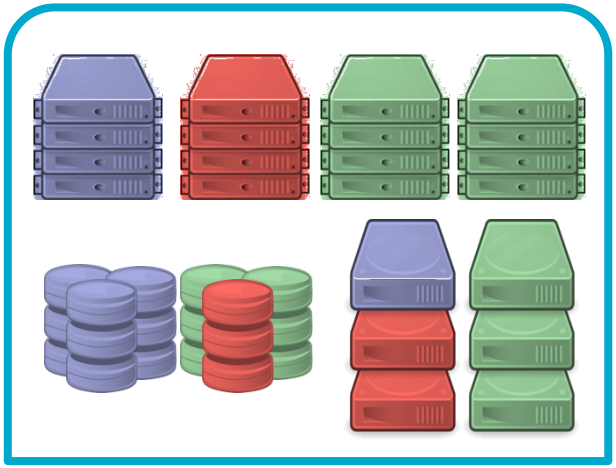
Web Development

HPC Simulations

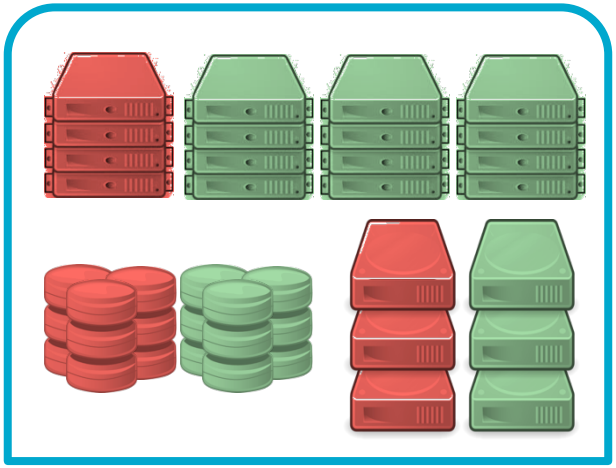
Big Data Analysis



Public Clouds



DC West Coast



DC Europe

The OpenNebula Provisioning Model

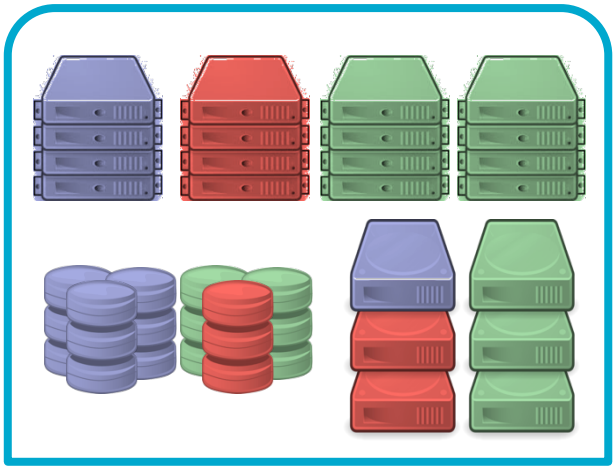
Definition of vDCs

Web Development

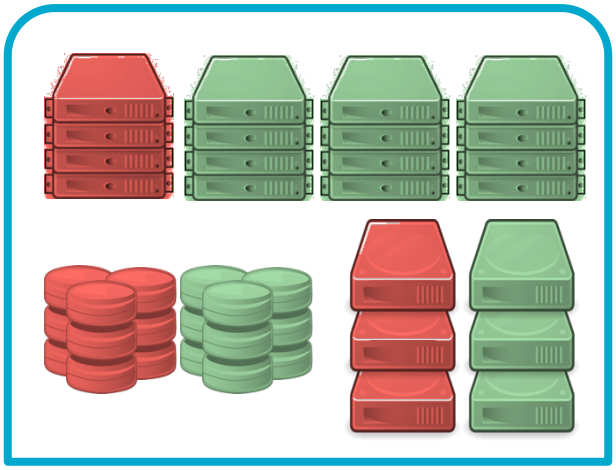
HPC Simulations

Big Data Analysis

Public Clouds



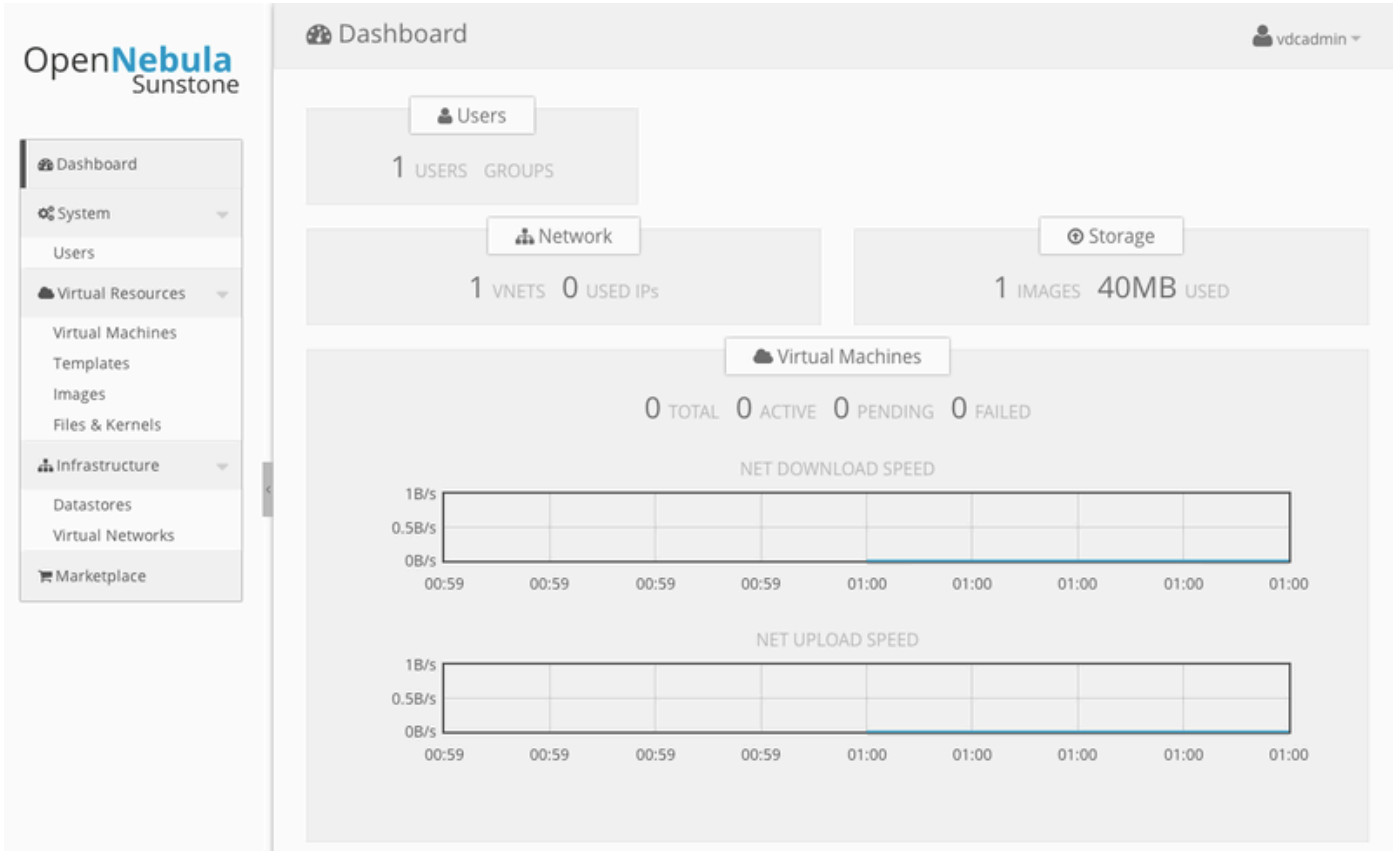
DC West Coast



DC Europe

Admins in each Group/vDC Manage to its Own Virtual Private Cloud

- Each vDC has an **admin**
- **Delegation of management** in the VDC
- Only **virtual resources**, not the underlying physical infrastructure

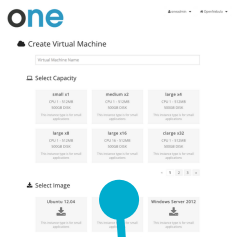


vDC Admin View

The OpenNebula Provisioning Model

Users in each Group/vDC Access to its Own Virtual Private Cloud

Cloud API

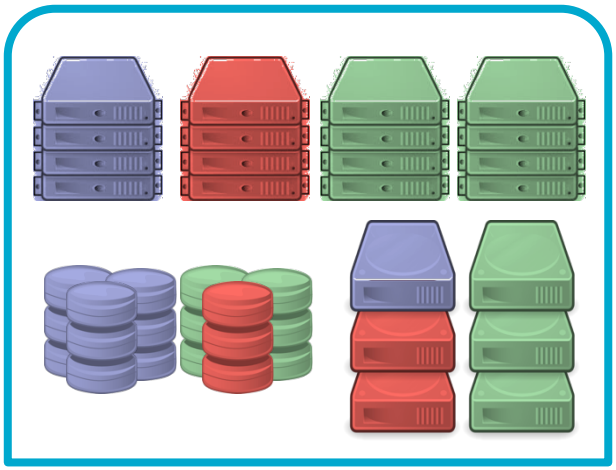


Web Development

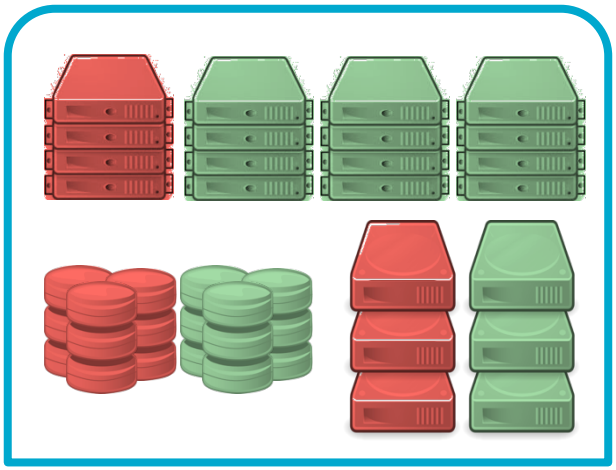
HPC Simulations

Big Data Analysis

Public Clouds



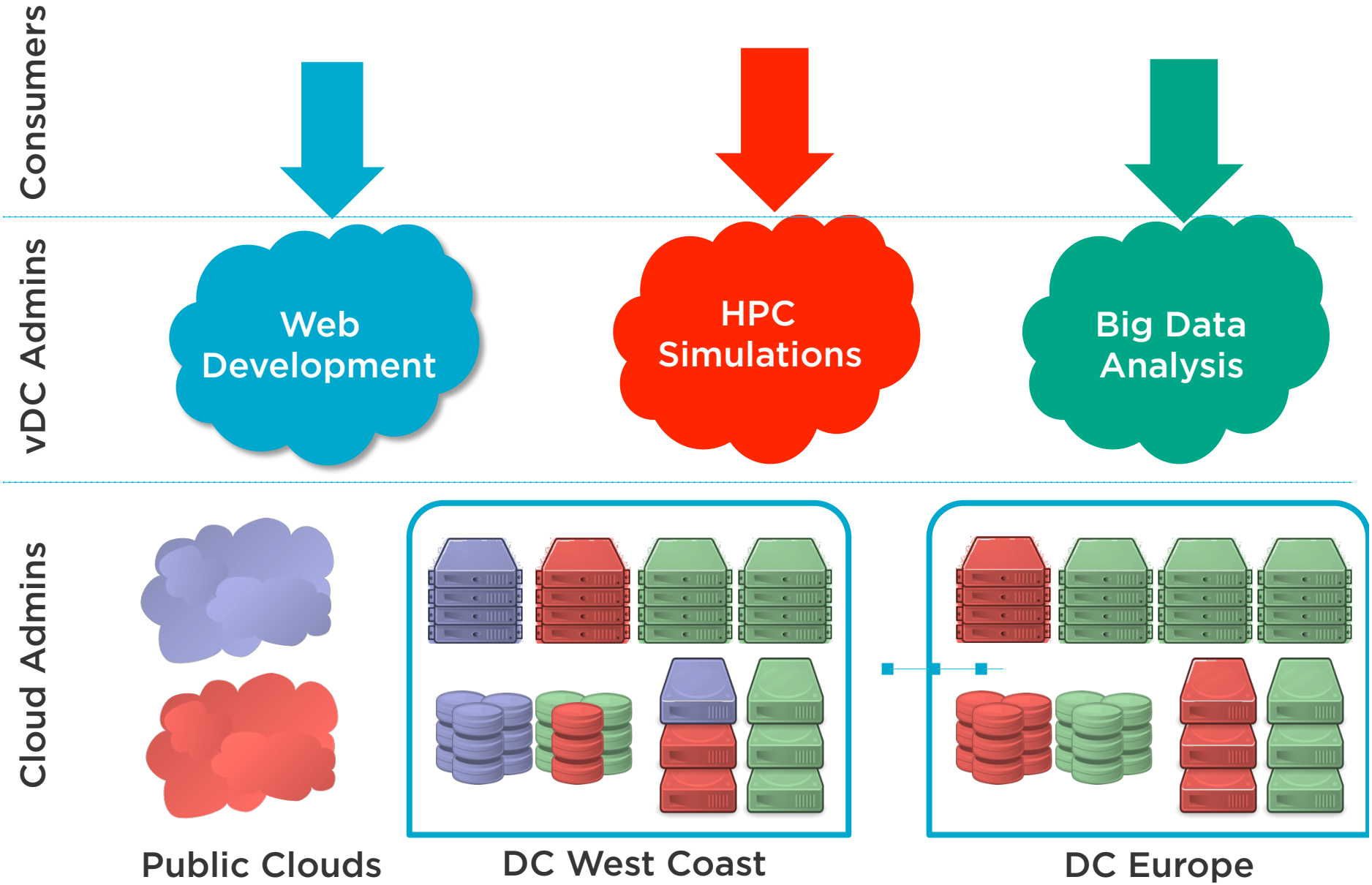
DC West Coast



DC Europe

The OpenNebula Provisioning Model

New Level of Provisioning: IaaS as a Service



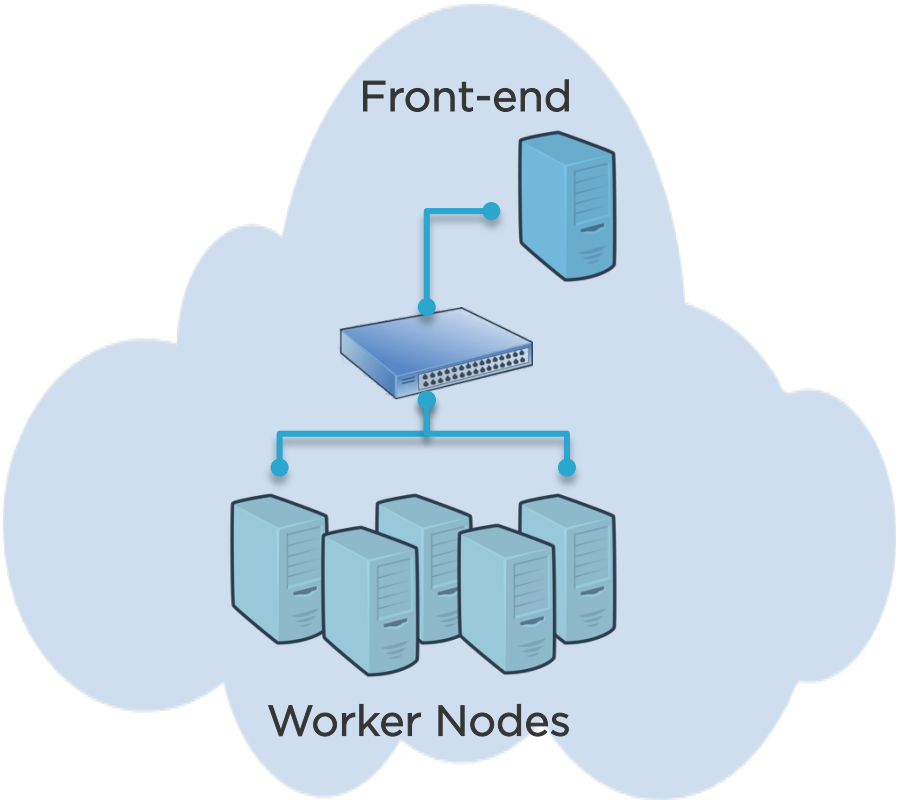
Benefits

- **Partition of cloud resources**
- Complete **isolation** of users, organizations or workloads
- Allocation of Clusters with different levels of **security, performance** or **high availability** to different groups with different workload profiles
- **Containers** for the execution of virtual appliances (SDDCs)
- Way of **hiding physical resources** from Group members
- Simple **federation** and **scalability** of cloud infrastructures beyond a single cloud instance and data center

How Do I Manage a Multi-tier Service?

A Comprehensive Framework to Manage Complex Applications

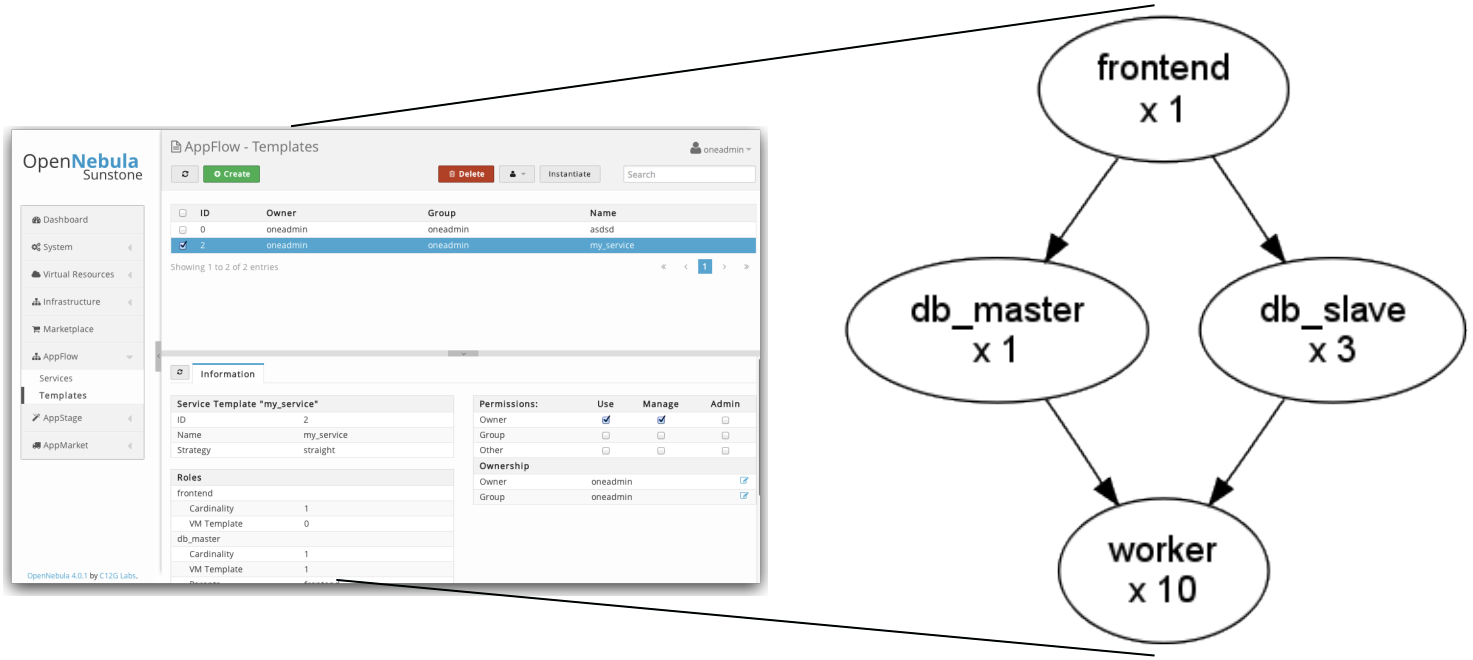
- Several tiers
- Deployment dependencies between components
- Each tier has its own cardinality and elasticity rules
- **...**



Multi-tier Application Management and Catalog

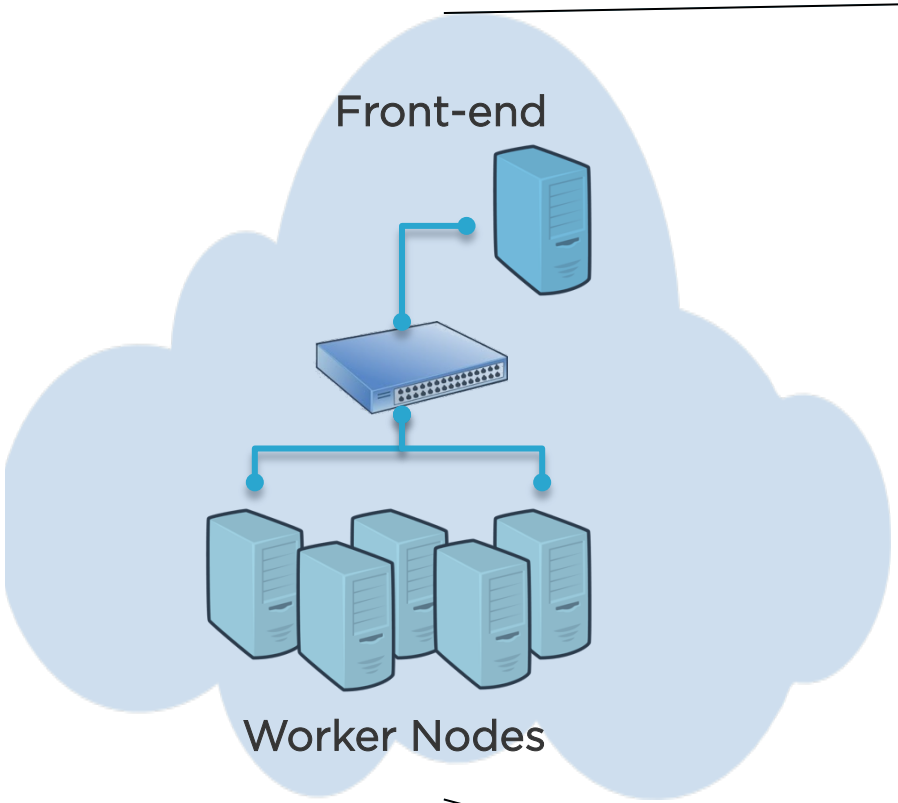
Management of interconnected multi-VM applications:

- Definition of **application flows**
- **Catalog** with pre-defined applications
- **Sharing** between users and groups
- Management of **persistent scientific data**
- Automatic **elasticity**



The OpenNebula Consumption Model

An Example of Flow




```
{ "name": ”Computing_Cluster",
  "deployment": "straight",
  "roles": [
    {
      "name": "frontend",
      "vm_template": 0
    },
    {
      "name": "worker",
      "parents": frontend,
      "cardinality": 2,
      "vm_template": 3,
      "min_vms" : 1,
      "max_vms" : 5,
      "elasticity_policies" :
        {
          ”expressions" : ”CPU> 90%”,
          "type" : "CHANGE",
          "adjust" : 2,
          "period_number" : 3,
          "period" : 10
        }, …
```

Let us Talk in Services Terms

Types of Elasticity Rules

- Auto-scaling Based on Metrics

```
"elasticity_policies" : [
  {
    "expression" : "ATT > 50",
    "type" : "CHANGE",
    "adjust" : 2,

    "period_number" : 3,
    "period" : 10
  },
  ...
]
```

- Auto-scaling Based on a Schedule

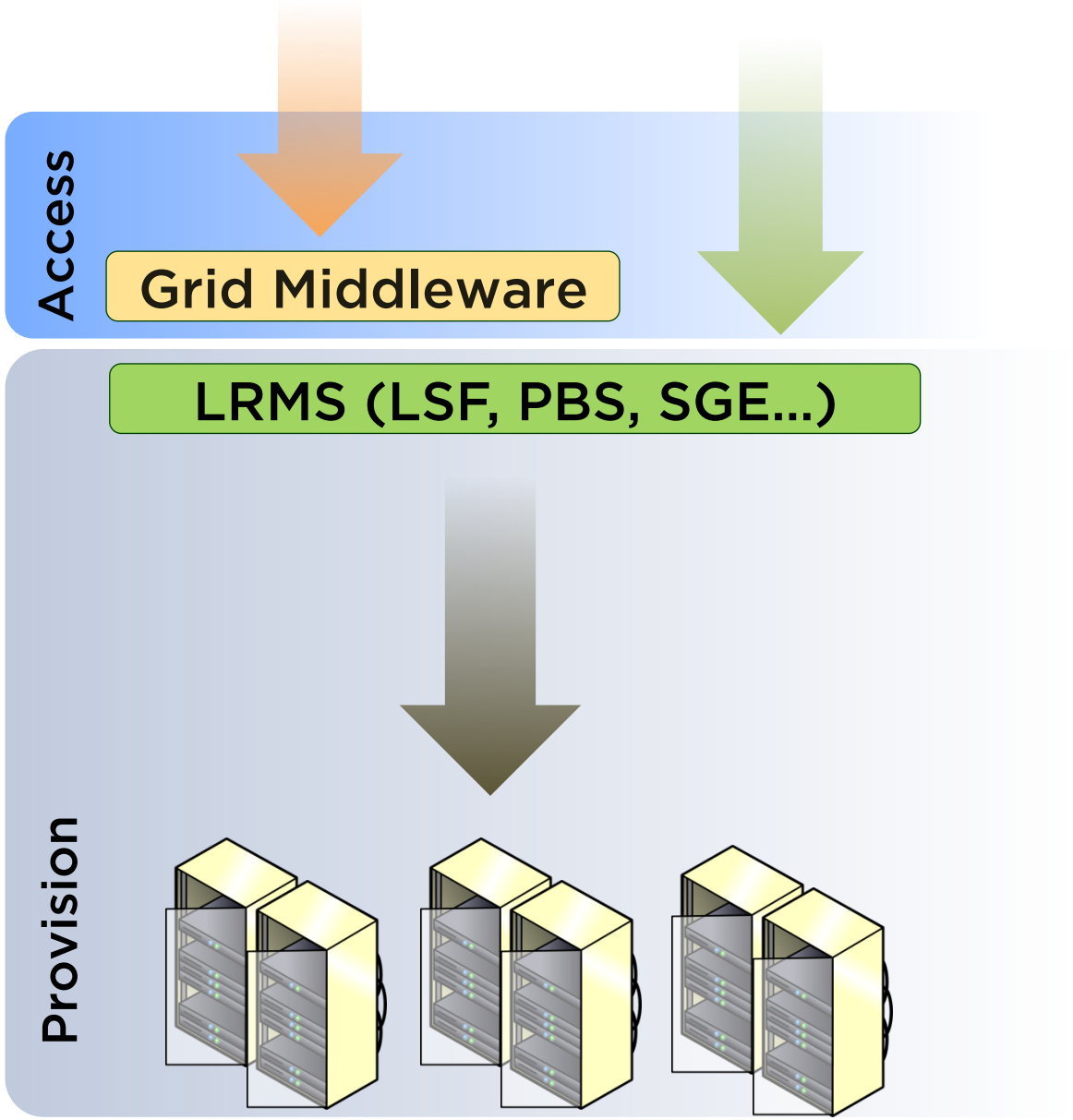
```
"scheduled_policies" : [
  {
    // Set cardinality to 2 each 10 minutes
    "recurrence" : "*/10 * * * *",

    "type" : "CARDINALITY",
    "adjust" : 2
  },
  {
    // +10 percent at the given date and time
    "start_time" : "2nd oct 2013 15:45",

    "type" : "PERCENTAGE_CHANGE",
    "adjust" : 10
  }
]
```

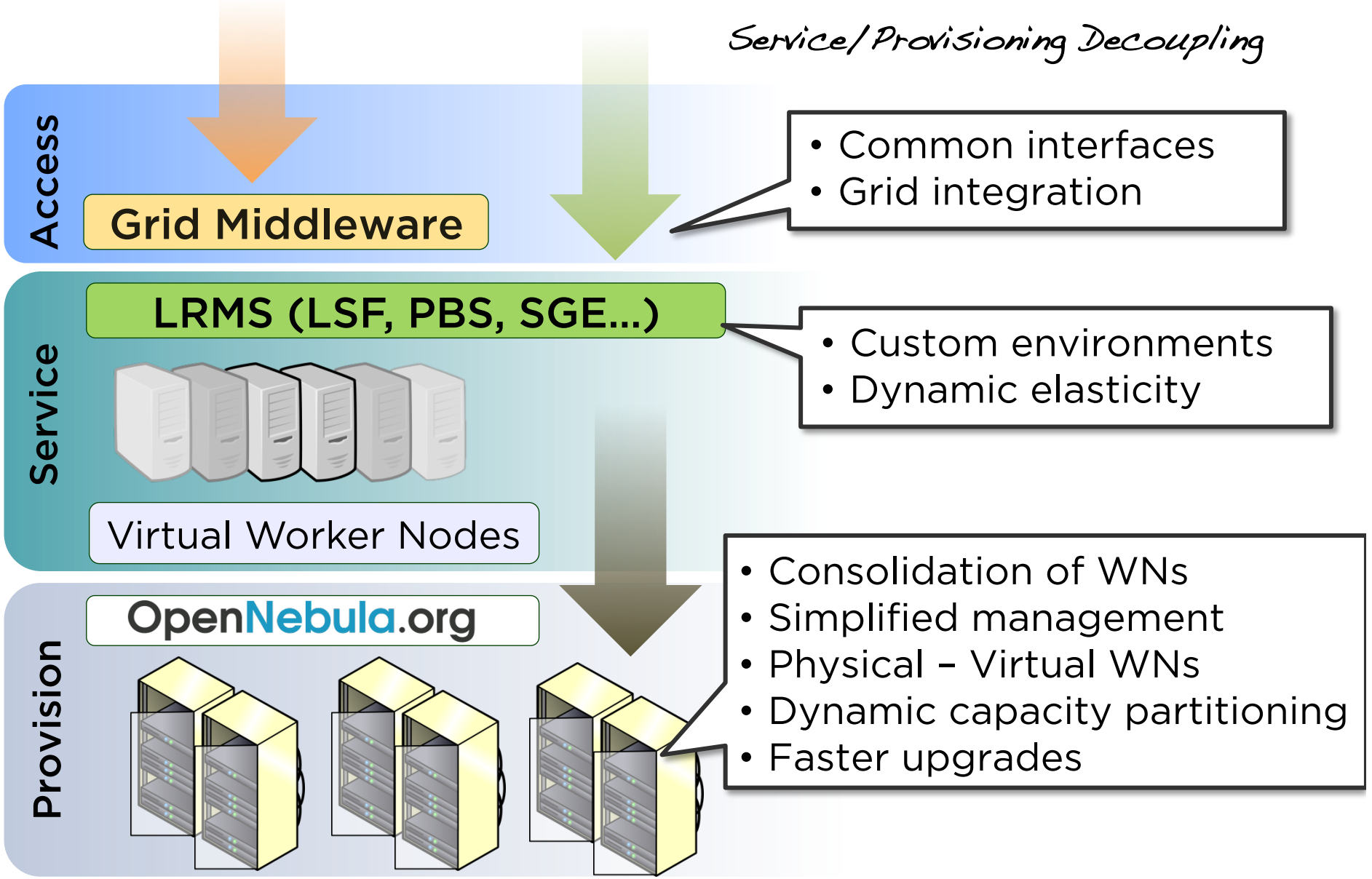
The Private HPC and Science Cloud Use Case

The Pre-cloud Era



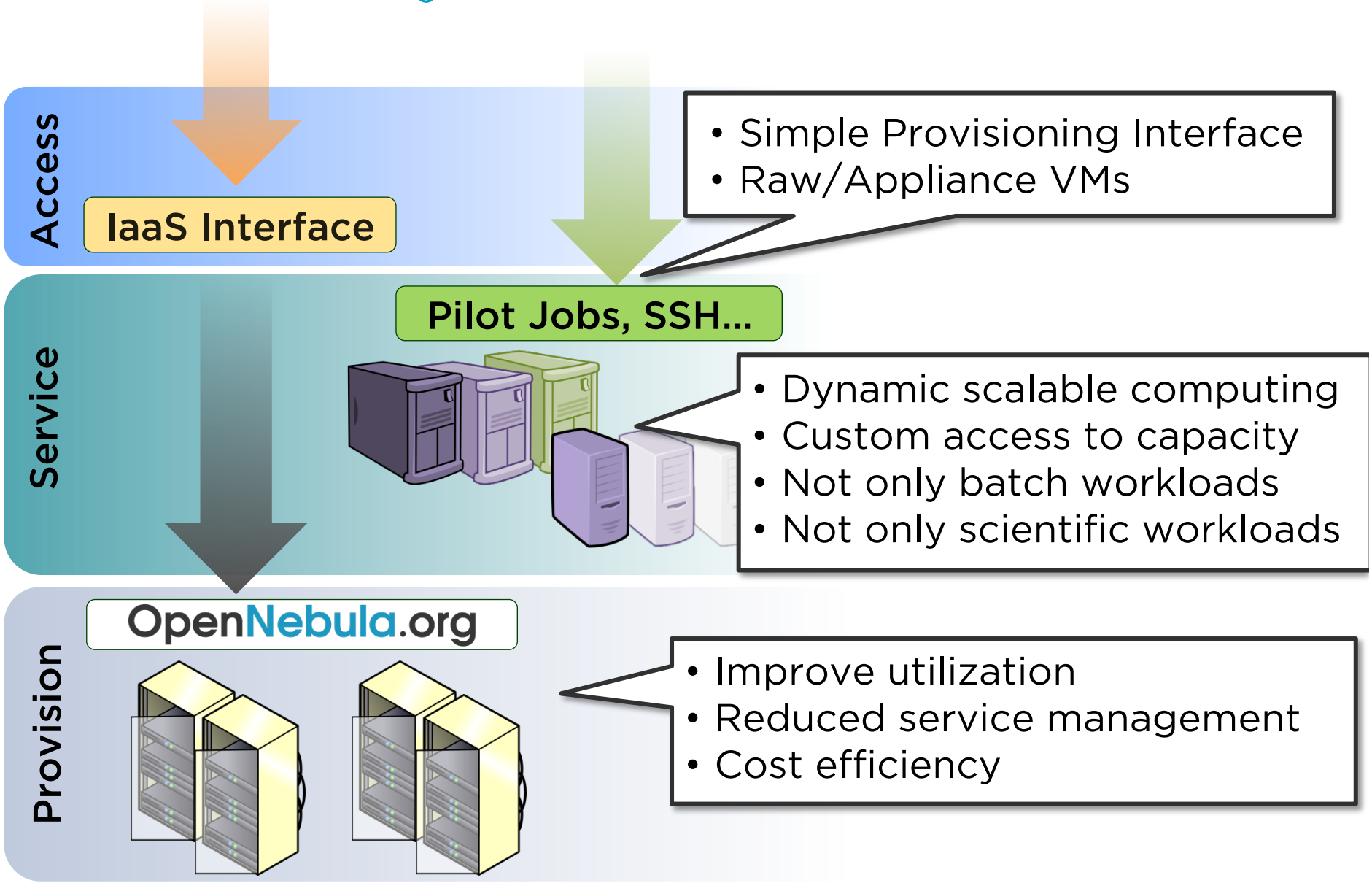
OpenNebula as an Infrastructure Tool – Enhanced Capabilities

Service/Provisioning Decoupling



The Private HPC and Science Cloud Use Case

OpenNebula as an Provisioning Tool - Enhanced Capabilities



The Private HPC and Science Cloud Use Case

Using the Cloud – Performance Penalty as a Small Tax You Have to Pay

Overhead in Virtualization

- Single processor performance penalty between **1% and 5%**
- NASA has reported an overhead between **9% and 25%** (HPCC and NPB)[1]
- Growing number of users demanding containers (**OpenVZ** and **LXC**)

Overhead in Input/Output

- Growing number of **Big Data apps**
- Support for **multiple datastores including automatic scheduling**

Need for Low-Latency High-Bandwidth Interconnection

- Lower performance, **10 GigE** typically, used in clouds has a significant negative (**x2-x10,** especially latency) impact on HPC applications[1]
- FermiCloud has reported MPI performance (HPL benchmark) on VMs and SR-IOV/**Infiniband** with **only a 4%** overhead[2]
- The Center for HPC at CSR has contributed the **KVM SR-IOV Drivers for Infiniband**[3]

(1) An Application-Based Performance Evaluation of Cloud Computing, NASA Ames, 2013
(2) FermiCloud Update, Keith Chadwick!, Fermilab, HePIX Spring Workshop 2013
(3) http://wiki.chpc.ac.za/acelab:opennebula_sr-iov_vmm_driver , 2013

The Private HPC and Science Cloud Use Case

Operating the Cloud – Innovations in Resource Management

Optimal Placement of Virtual Machines
- Automatic placement of VM near input data
- Striping policy to maximize the resources available to VMs

Fair Share of Resources
- Resource quota management to allocate, track and limit resource utilization

Isolated Execution of Applications
- Full Isolation of performance-sensitive applications

Management of Different Hardware Profiles
- Resource pools (physical clusters) with specific Hw and Sw profiles, or security levels for different workload profiles (HPC and HTC)

Hybrid Cloud Computing
- Cloudbursting to address peak or fluctuating demands for no critical and HTC workloads

Provide VOs with Isolated Cloud Environ
- Automatic provision of Virtual Data Centers

One of Our Main User Communities

Supercomputing Centers

NCHC, SURF SARA, KTH Vetenskap och Konst, Centre de Supercomputació de Catalunya CESCA, PIC port d'informació científica, CESGA

Industry

MONTE DEI PASCHI DI SIENA Banca dal 1472, AIRBUS, produban, Santander

Research Centers

IPB, esa, CSIRO, MTA SZTAKI, KISTI www.kisti.re.kr, KIT Karlsruher Institut für Technologie, DESY, Fermilab, VERITAS, NIKHEF

Distributed Computing Infrastructures

egi, PRAGMA, Future Grid, BonFIRE, meGha, Ochain reds

The Private HPC and Science Cloud Use Case

FermiCloud

Fermilab

http://www-fermicloud.fnal.gov/

Nodes	**KVM** on 29 nodes (2 TB RAM – 608 cores) Koi Computer
Network	Gigabit and **Infiniband**
Storage	CLVM+**GFS2** on shared 120TB NexSAN SataBeats
AuthN	**X509**
Linux	**Scientific Linux**
Interface	**Sunstone** Self-service and **EC2** API
App Profile	Legacy, HTC and **MPI HPC**

Typical Workloads

- Production VM-based batch system via the "econe" EC2 emulation
- Scientific stakeholders get access to on-demand VMs
- Developers & integrators of new Grid applications

CESGA Cloud

http://cloud.cesga.es/

Nodes	**KVM** on 35 nodes (0.6 TB RAM – 280 cores) HP ProLiant
Network	2 x **Gigabit** (1G and 10G)
Storage	ssh from remote **EMC** storage server
AuthN	**X509** and core password
Linux	**Scientific Linux 6.4**
Interface	**Sunstone** Self-service and **OCCI**
App Profile	Individual VMs and **virtualised computing clusters**

Typical Workloads

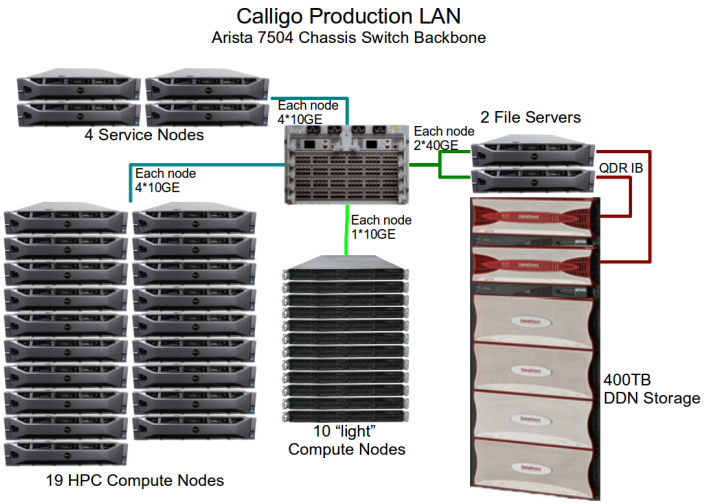
- 160 users
- Genomic, rendering…
- Grid services on production at CESGA
- Node at FedCloud project
- UMD middleware testing

SARA Cloud

SURF SARA

ww.cloud.sara.nl

Nodes	**KVM** on 30 HPC nodes (256 GB RAM 960 cores + 2 TB High-memory node) Dell PowerEdge and 10 “light” nodes (64 GB RAM 80 cores) Supermicro
Network	2 x **Gigabit** (10G) with Arista switch
Storage	**NFS** on 500 TB NAS for HPC and ssh for “light”
AuthN	**Core password**
Linux	**CentOS**
Interface	**Sunstone** and **OCCI**
App Profile	MPI clusters, **windows clusters** and independent VMs



Typical Workloads

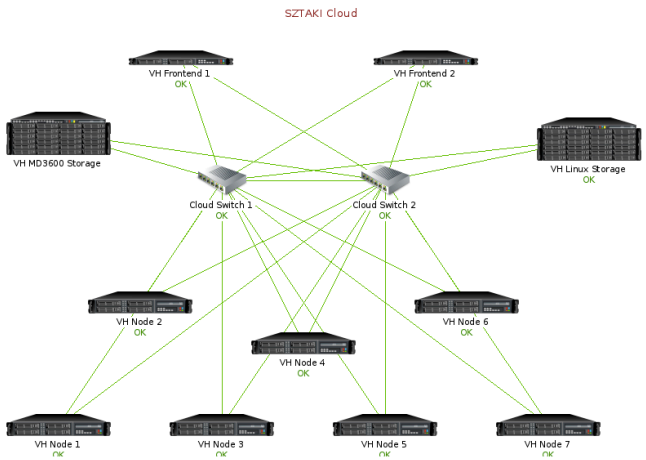
- Ad-hoc clusters with MPI and pilot jobs
- Windows clusters for Windows-bound software
- Single VMs, sometimes acting as web servers to disseminate results

SZTAKI Cloud

MTA SZTAKI

http://cloud.sztaki.hu/

Nodes	**KVM** on 8 nodes (2 TB RAM – 512 cores) DELL PowerEdge
Network	Redundant **10Gb**
Storage	Dell storage servers: **iSCSI** (36TB) and **CEPH** (288 TB)
AuthN	**X509**
Linux	**CentOS 6.5**
Interface	**Sunstone** Self-service, **EC2** and **OCCI**
App Profile	Individual VMs and **virtualised computing cluster**



Typical Workloads

- Run standard and grid services (e.g.:web servers, grid middlewares…)
- Development and testing of new codes
- Research on performance and opportunistic computing

KTH Cloud

http://www.pdc.kth.se/

Nodes	**KVM** on 768 cores (768 GB RAM) HP ProLiant
Network	**Infiniband** and Gigabit
Storage	NFS and **LVM**
AuthN	**X509** and core password
Linux	**Ubuntu**
Interface	**Sunstone** self-service, **OCCI** and **EC2**
App Profile	Individual VMs and **virtualised computing cluster**

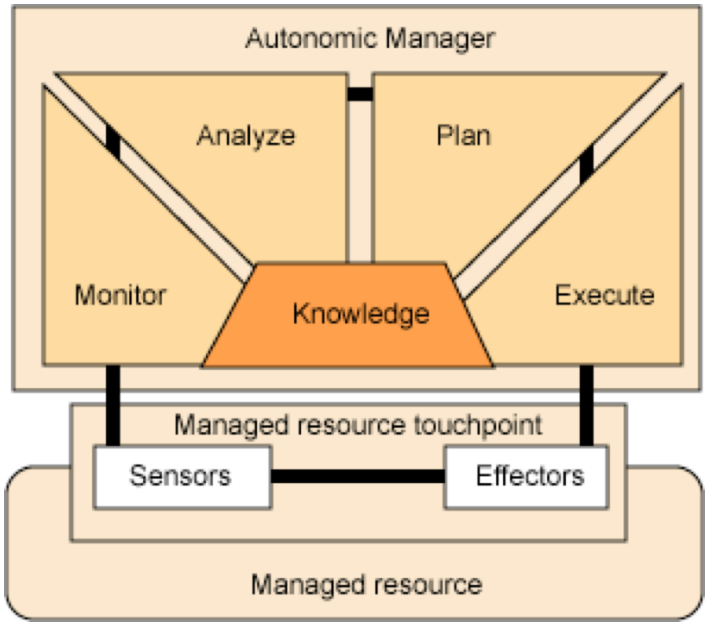
Typical Workloads

- Mainly BIO
- Hadoop, Spark, Galaxy, Cloud Bio Linux...

Autonomic Management Properties

Properties of Autonomic Systems:

- **Self-configuring**: Ability to define themselves "on-the-fly"
- **Self-healing**: Discover, diagnose, and react to disruptions
- **Self-optimizing** (self-adapting): Monitor and tune resources automatically
- **Self-protecting**: Detect, identify and protect themselves against threats



IBM architecture for autonomic computing

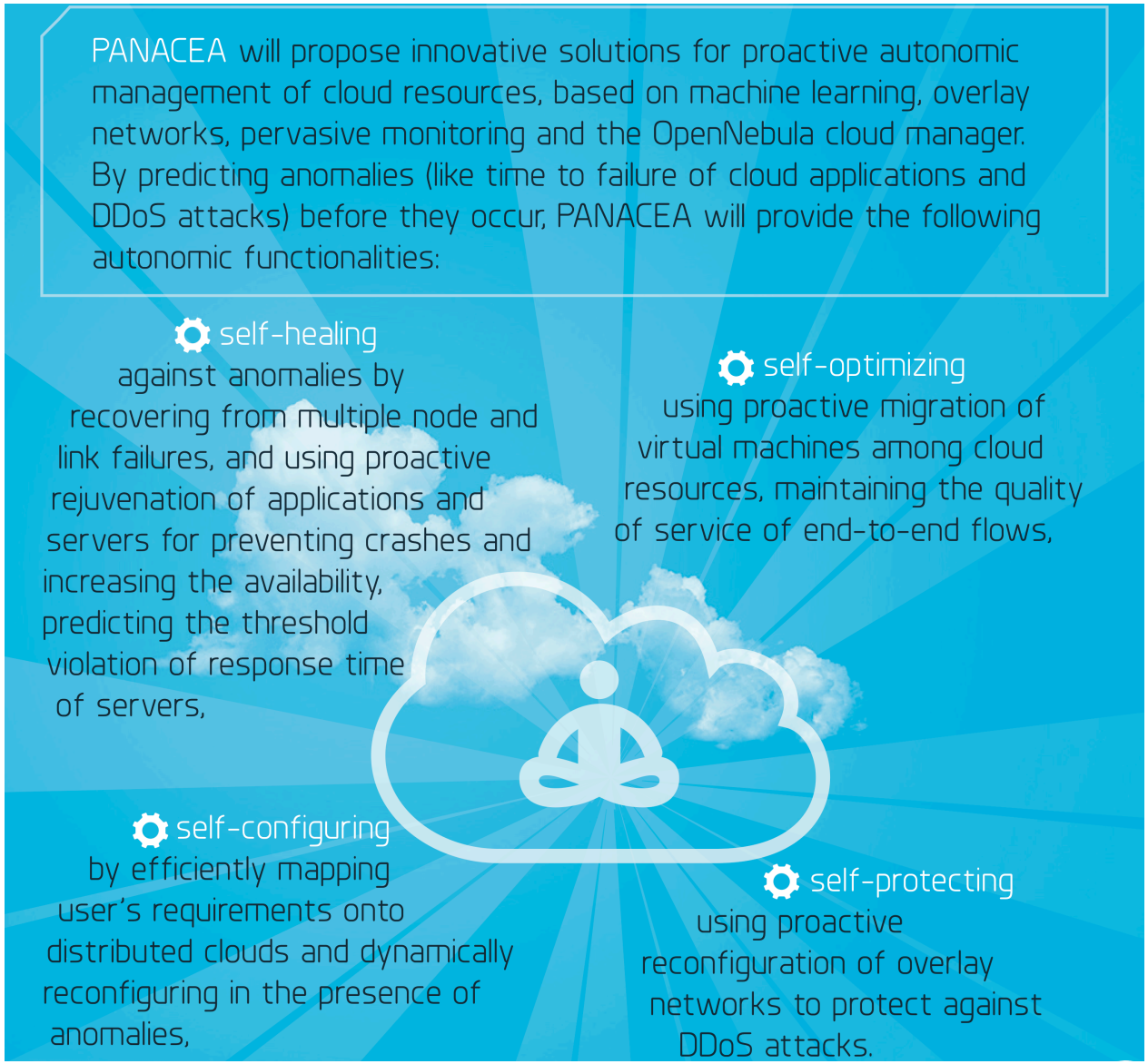
Proactive Autonomic Management Use Case

Upcoming Innovations in Automatic Management Support

Configure	Systems can change parameters (capacity, placement,...) at runtime
	Systems can connect new devices at runtime (hot plugging)
Healing	Systems can detect faults and recover from them
	Systems can perform software rejuvenation
	Systems can manage spares for application live migration
Optimizing	Systems can get information of the environment at runtime and adapt to it
	Systems can manage elasticity
Protecting	Systems can define and manage user access
	Systems can perform backup and recovery

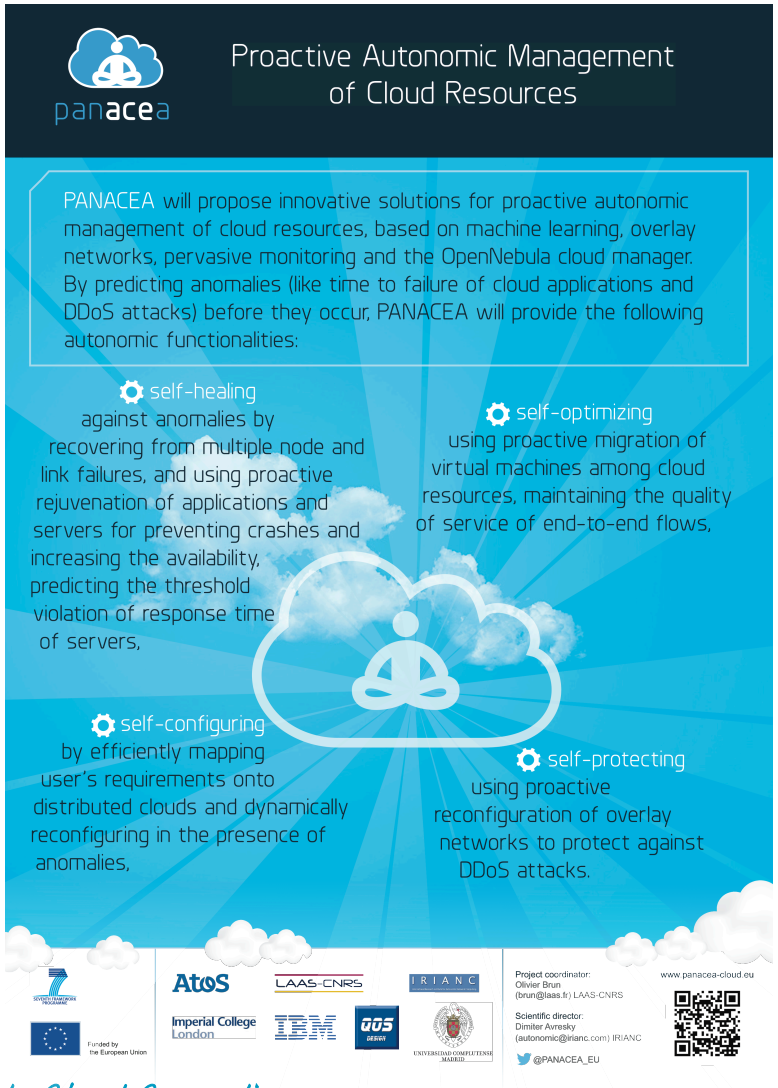
Proactive Autonomic Management Use Case

The PANACEA FP7 Project: http://panacea-cloud.eu



The PANACEA Poster at CLOSER 2014

Dimiter Avresky will talk about PANACEA today at 16:45
Session about EU Projects (MR09)



Research References

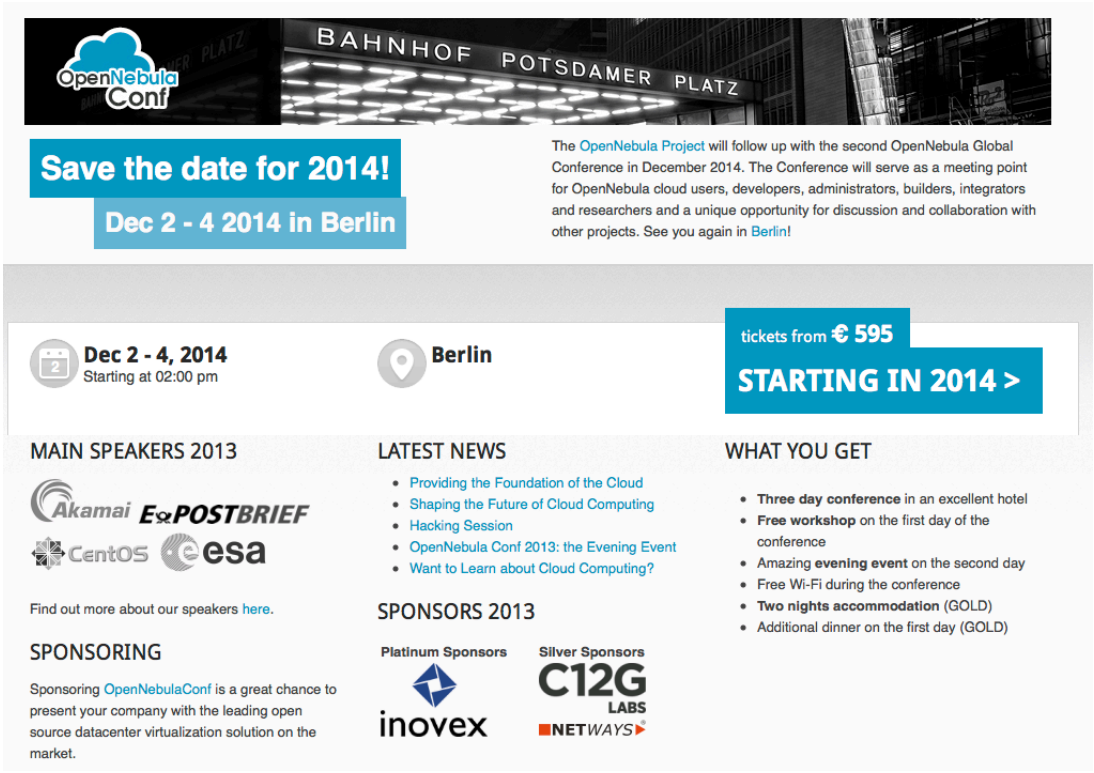
This is a Research Event...

Innovation in Cloud Architecture

- B. Sotomayor, R. S. Montero, I. M. Llorente and I. Foster, "Virtual Infrastructure Management in Private and Hybrid Clouds", **IEEE Internet Computing**, September/October 2009 (vol. 13 no. 5)
- Rafael Moreno-Vozmediano, Ruben S. Montero, Ignacio M. Llorente, "Multi-Cloud Deployment of Computing Clusters for Loosely-Coupled MTC Applications", **IEEE Transactions on Parallel and Distributed Systems**, 22(6): 924-930, April 2011
- Rafael Moreno-Vozmediano, Ruben S. Montero, Ignacio M. Llorente, "IaaS Cloud Architecture: From Virtualized Data Centers to Federated Cloud Infrastructures", **IEEE Computer**, 45(12):65-72, December 2012
- Rafael Moreno-Vozmediano, Ruben S. Montero, Ignacio M. Llorente, "Key Challenges in Cloud Computing to Enable the Future Internet of Services", **IEEE Internet Computing**, 17(4):18-25, 2012.

Upcoming Community Events

Other Cloud TechDays scheduled in Florida, Bay Area, Berlin, Aveiro, Chicago…



Questions?

We Will Be Happy to Answer Your Questions

OpenNebula.org @OpenNebula