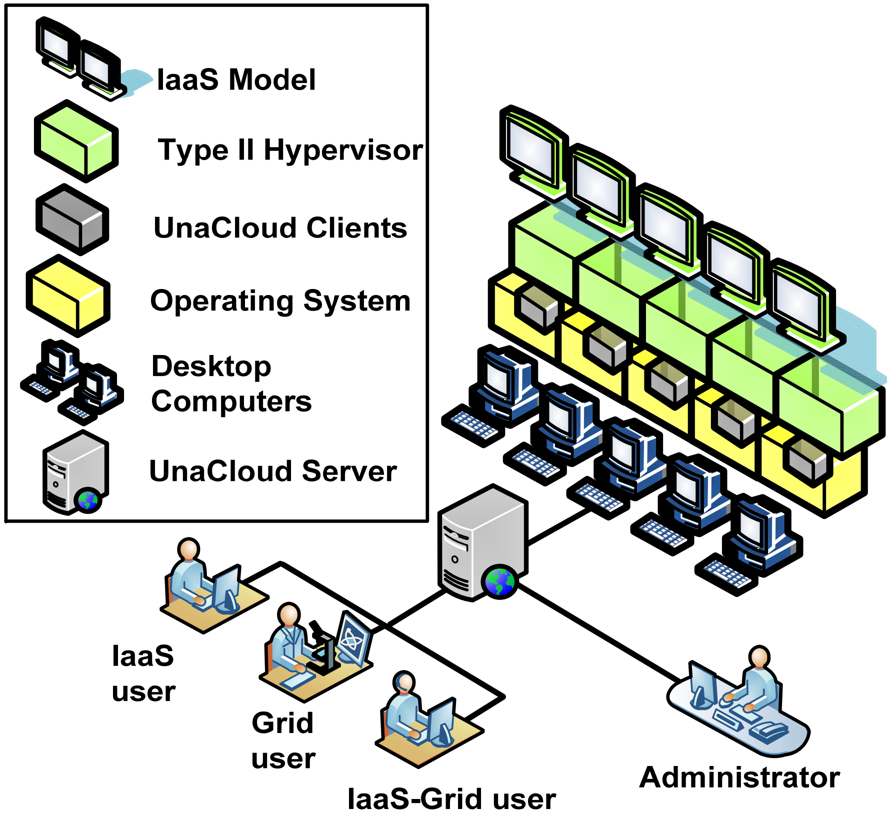
Opportunistic computing for HPC



Harold Castro
hcastro@uniandes.edu.co
Department of Systems and Computing Engineering
Universidad de los Andes
Bogotá, Colombia

SOLUTIONS USING NON DEDICATED INFRASTRUCTURES

THE DESIRED SOLUTION

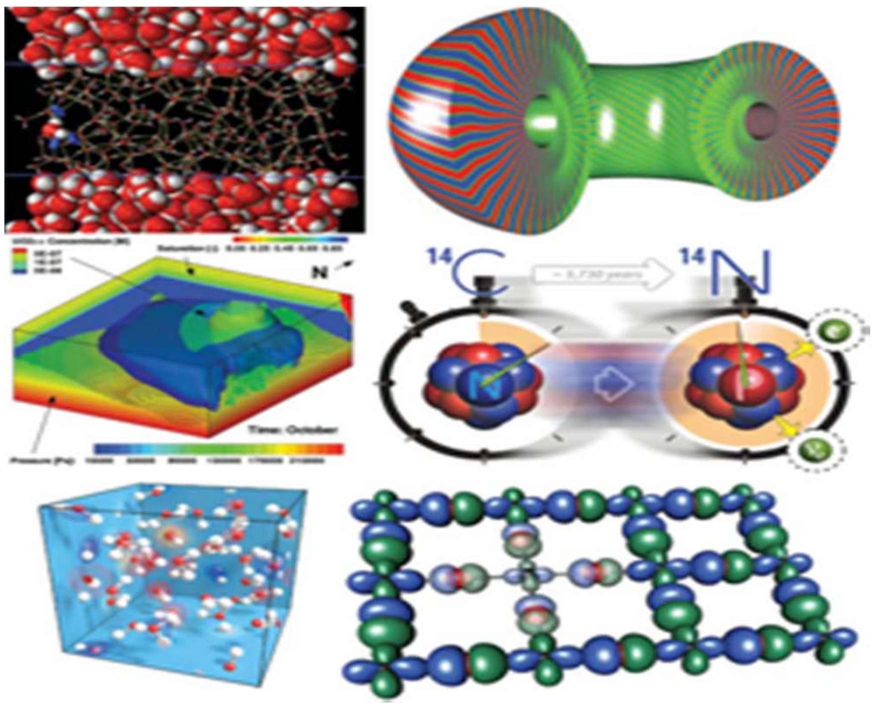
UNACLOUD APPROACH

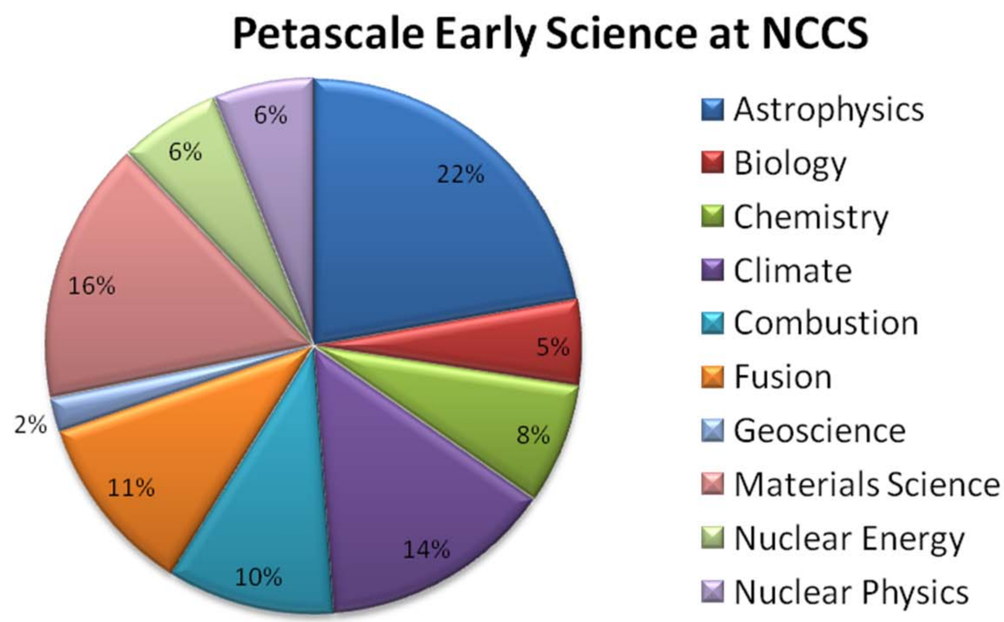
UNACLOUD IMPLEMENTATION

UNACLOUD TESTING AND RESULTS

CONCLUSIONS AND FUTURE WORK

THE PROBLEM

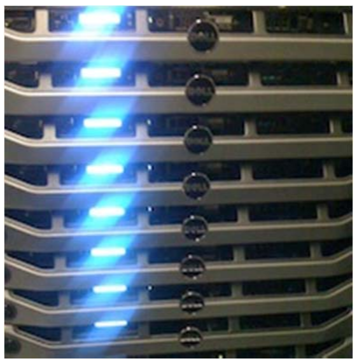


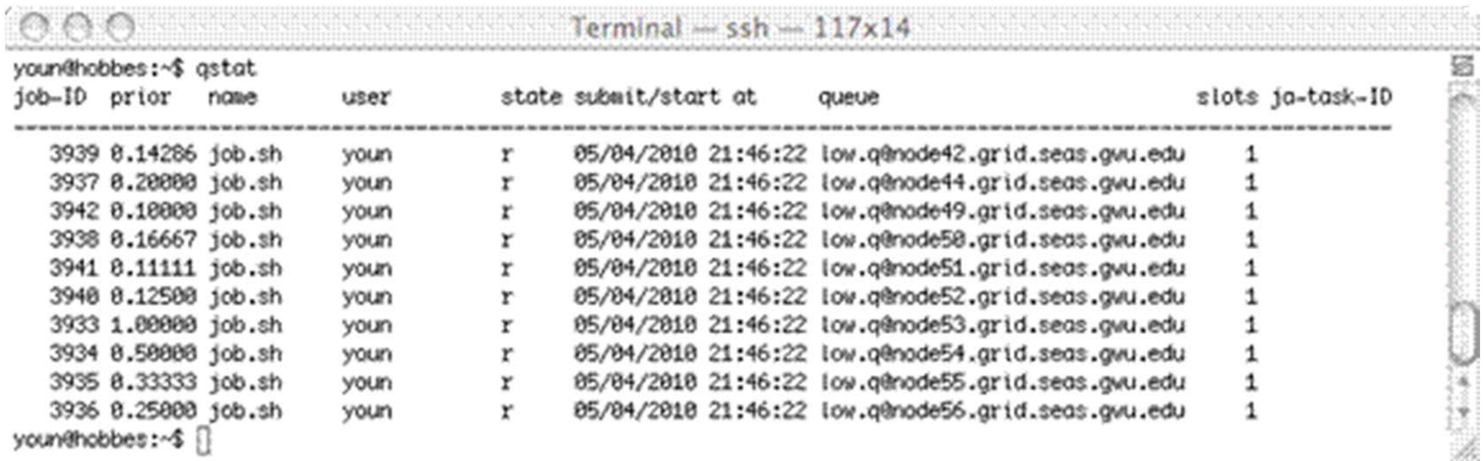


The development of e-Science projects requires large processing capabilities. These capabilities are regularly provided by dedicated *cluster, grid* and *cloud* computing infrastructures.

THE PROBLEM

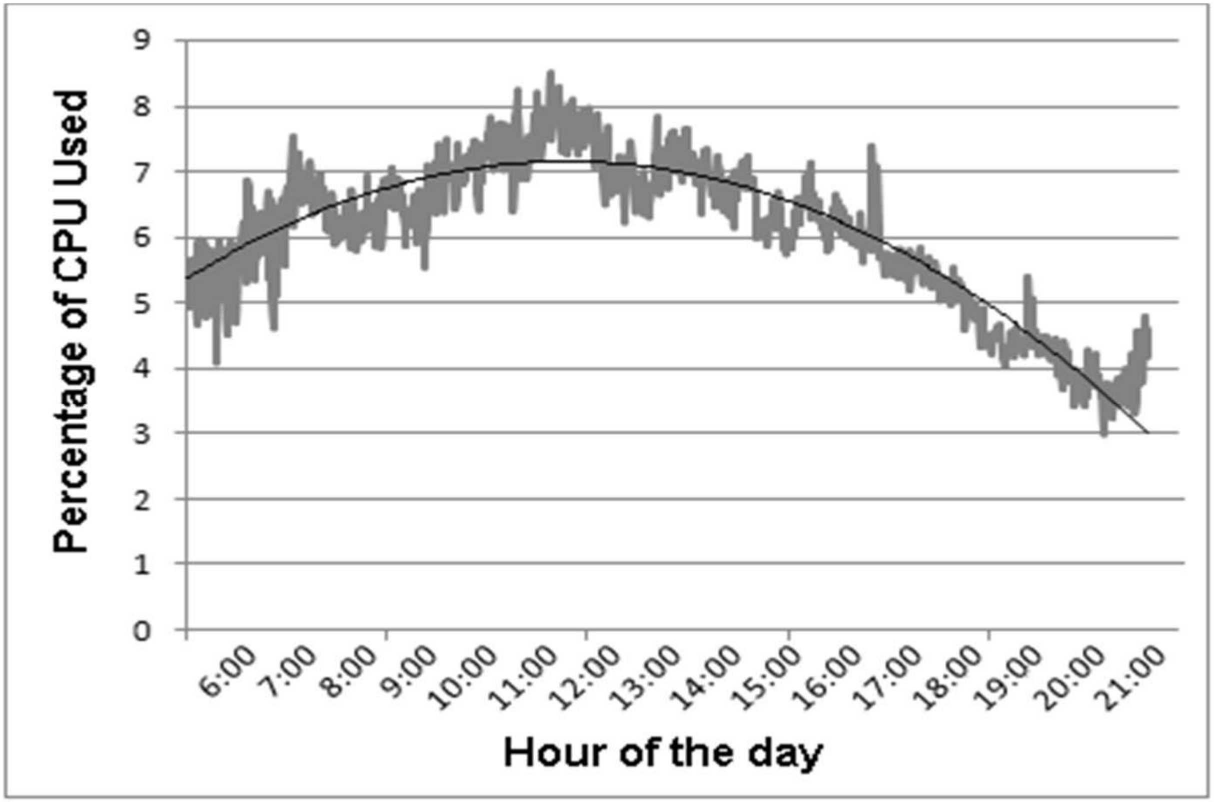
In the research environment of our university campus we find that each research group has its own dedicated clusters and some computer labs for students, so:





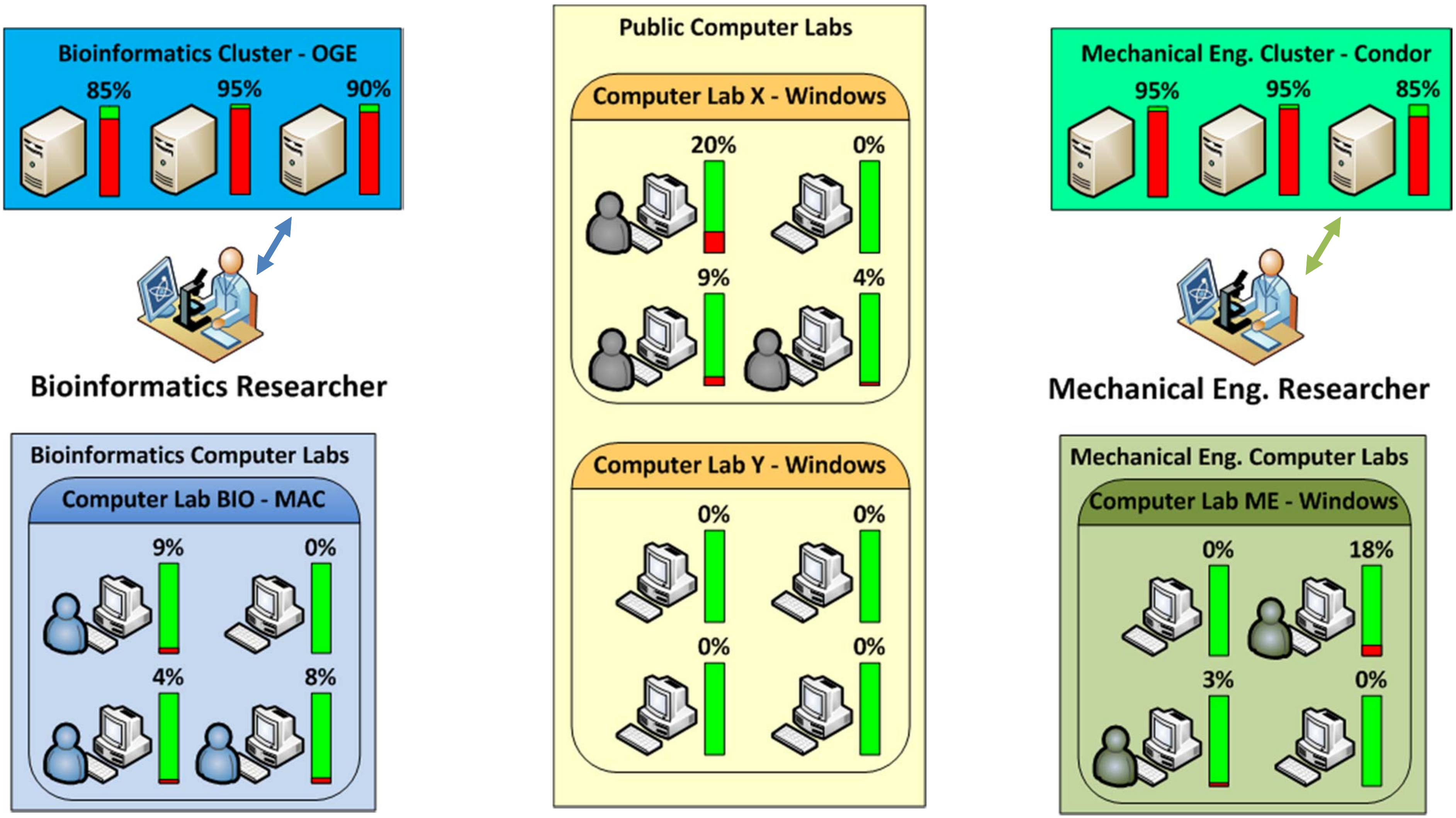
- **Researchers use and have some experience with specific cluster/grid middlewares to distribute load among the nodes of dedicated clusters (OGE, Condor, etc.).**
- **Computer labs have many commodity desktops with different operating systems: Windows (mainly), MAC and Linux, which are idle most of the time.**
- **Most e-Science applications are run on Linux.**

THE PROBLEM



- **Researchers require large HPC/HTC during some peak periods (a project needs to be delivered, the call for paper will finish, etc.).**
- **Additionally, there are a lot of general or public campus computer labs used by students and these labs are also idle most of the time.**

THE PROBLEM



More than 2000 CPU cores

THE CONTEXT: DGVCSs

Great Internet Mersenne Prime Search GIMPS

SETI@HOME

BOINC

InteGrade

distributed.net

Condor High Throughput Computing

OurGrid

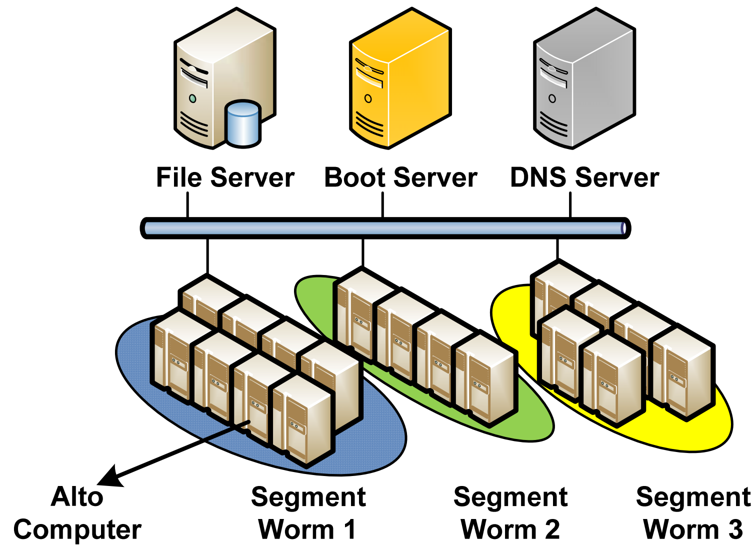
An alternative are Desktop Grids and Volunteer Computing Systems (DGVCS's):

- **Offer large scale computing infrastructures at low cost.**
- **Use inexpensive resources, most of them underutilized desktop computers.**
- **Interconnect thousands of computing resources available through Internet or Intranet environments.**
- **Are based on resources that are *non-dedicated*, distributed, highly heterogeneous, and part of independent administrative domains.**

DVGC over LANs

Condor High Throughput Computing

Worm XEROX



Pioneers DGVCS's:

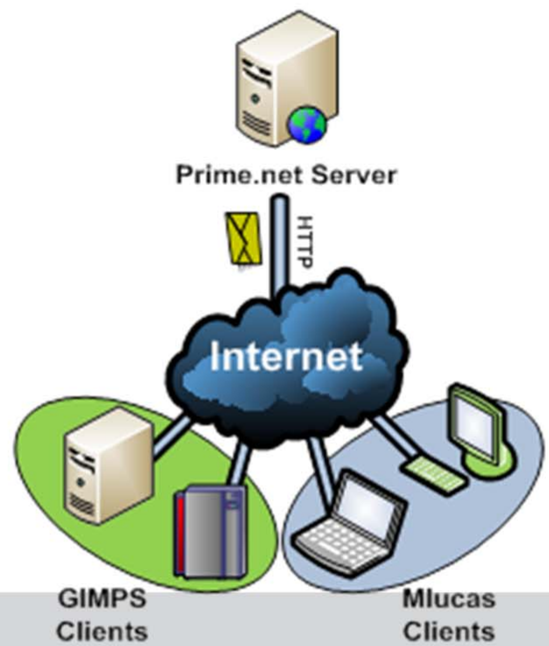
- Early 80s
- Low distribution and heterogeneity
- LAN scope
- Non dedicated resources
- Use of idle resources
- Failures aware

Single-Purpose DGVCSs over Internet

SETI@HOME

Great Internet Mersenne Prime Search

GIMPS



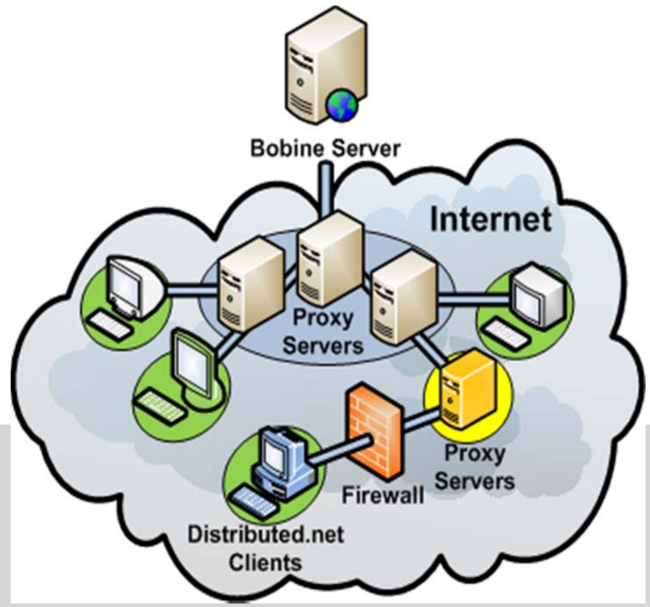
First DGVCS's at world scale:

- **2000s**
- **Highly distributed and heterogeneous computing: lightweight agents**
- **Independent administrative domains**
- **Petaflops capacities**
- **Restrictions on size of messages and files**
- **Redundancy and prizes**

General purpose DGVCs over Internet

DGVCS's for almost anything:

- **Huge processing capabilities applied to BoT applications (one at a time)**
- **Pre-defined projects**
- **Sharing resources configuration**
- **Security enforcement**



DGVCSs in grid computing environments

Bayanihan Computing.NET

OurGrid

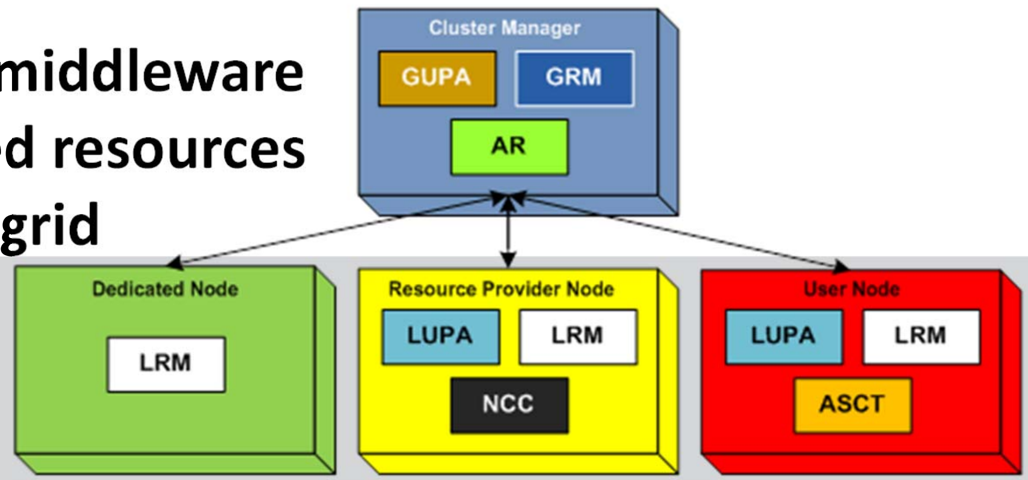
CondorG

Integrade

UnaGrid

DGVCS's participating in large scale projects:

- Variable scalability
- Deployment of a particular middleware
- Dedicated and non dedicated resources
- Integration to international grid
- Use of Web Services
- Clusters hierarchy



DGVCSs in cloud computing environments

CernVM: LHC@Home

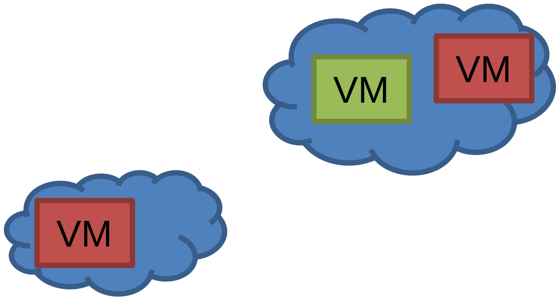
Cloud@Home

clouds@home

UnaCloud

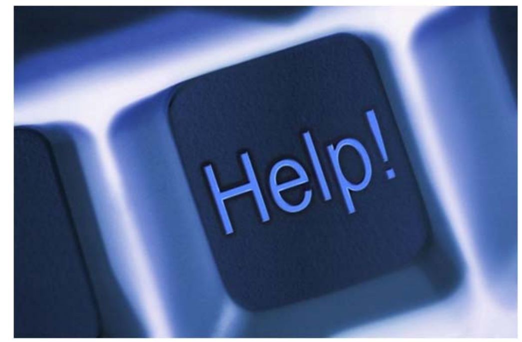
Virtual machines to emulate user's environment:

- Reaction to failures: live migration
- More concepts than implementations
- Extension to public clouds
- VM management



THE PROBLEM

When a research group wants to use a DGVCS it regularly find that:



- **They will need to recode, modify or adapt every application that is going to be executed on the DGVCS, for several research groups and tens of existing applications it is a complex process.**

- **The installation, configuration, maintaining and use of most DGVCSs require of people with some/advanced skills in applications and IT infrastructures.**

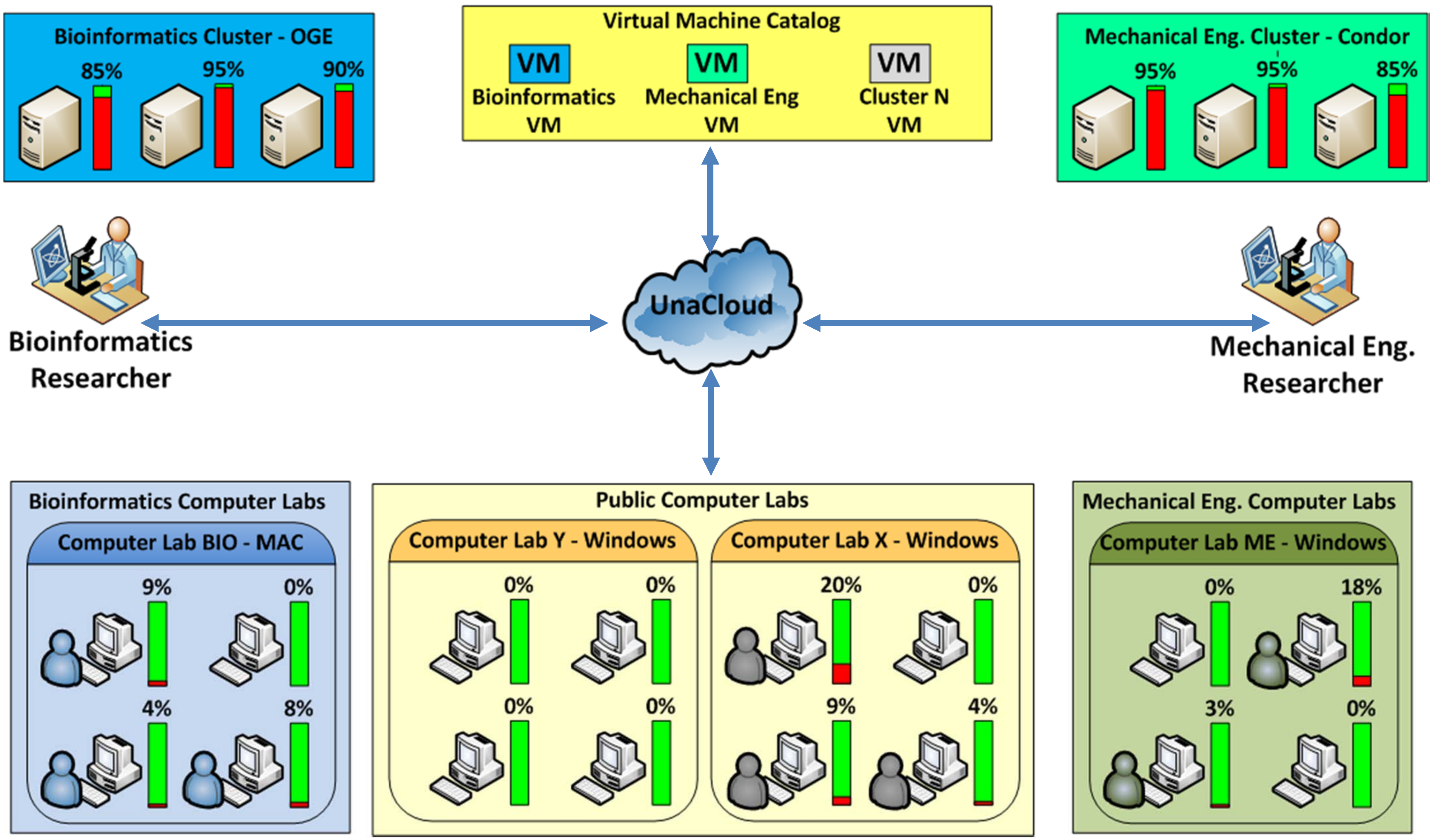
- **For using the idle processing capabilities of hundreds of commodity desktops, they will need to configure manually every desktop computer with the DGVCS software.**

THE PROBLEM

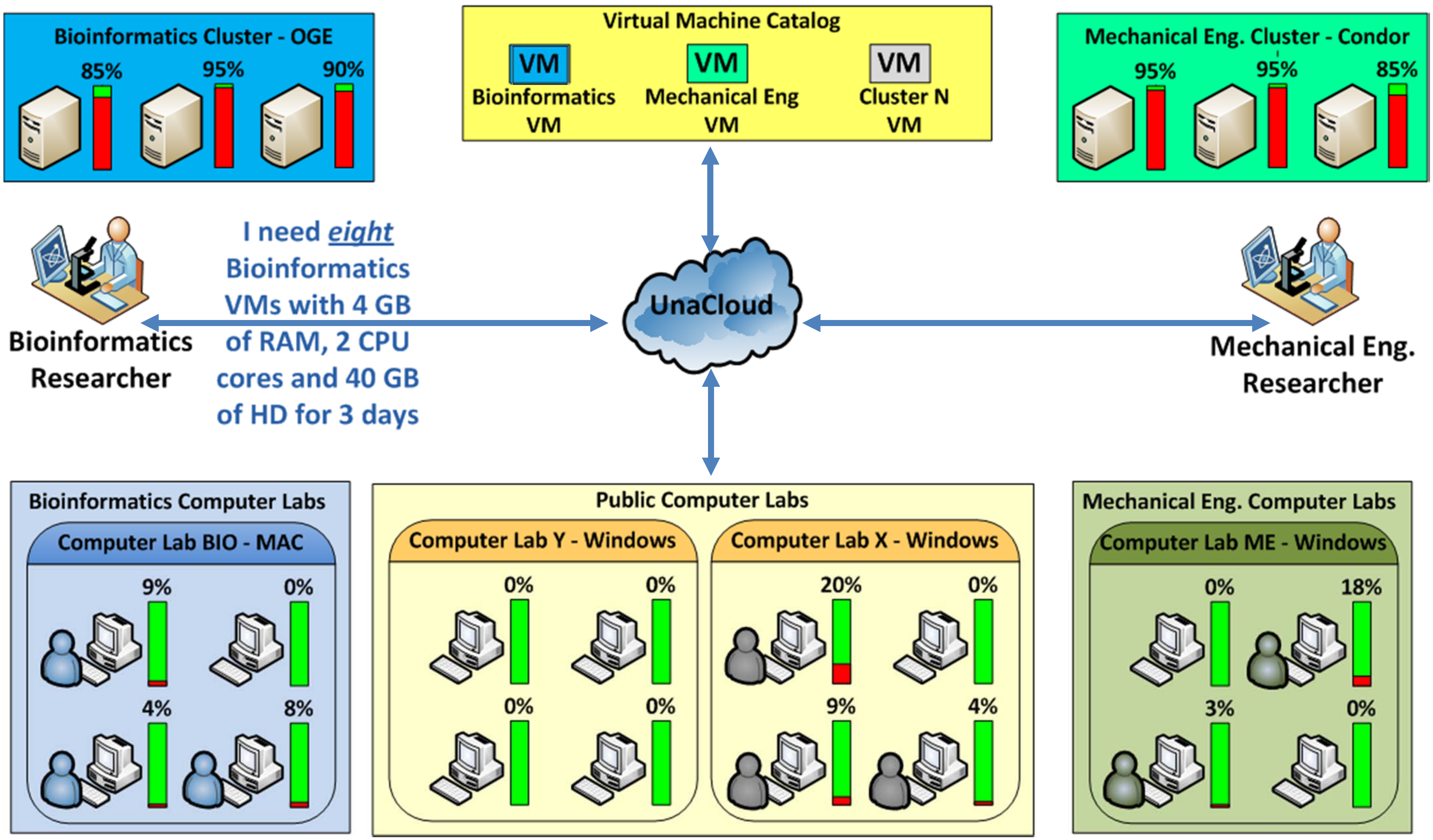
When a research group wants to use a DGVCS it regularly find that:

- **Administrators of different computer labs do not want that external people modify the configurations of the physical machines.**
- **Most of the physical desktops (99%) machines available in computer labs have Windows operating systems.**
- **They would like to share easily with other research groups the idle capabilities available in computer labs using a shared model.**

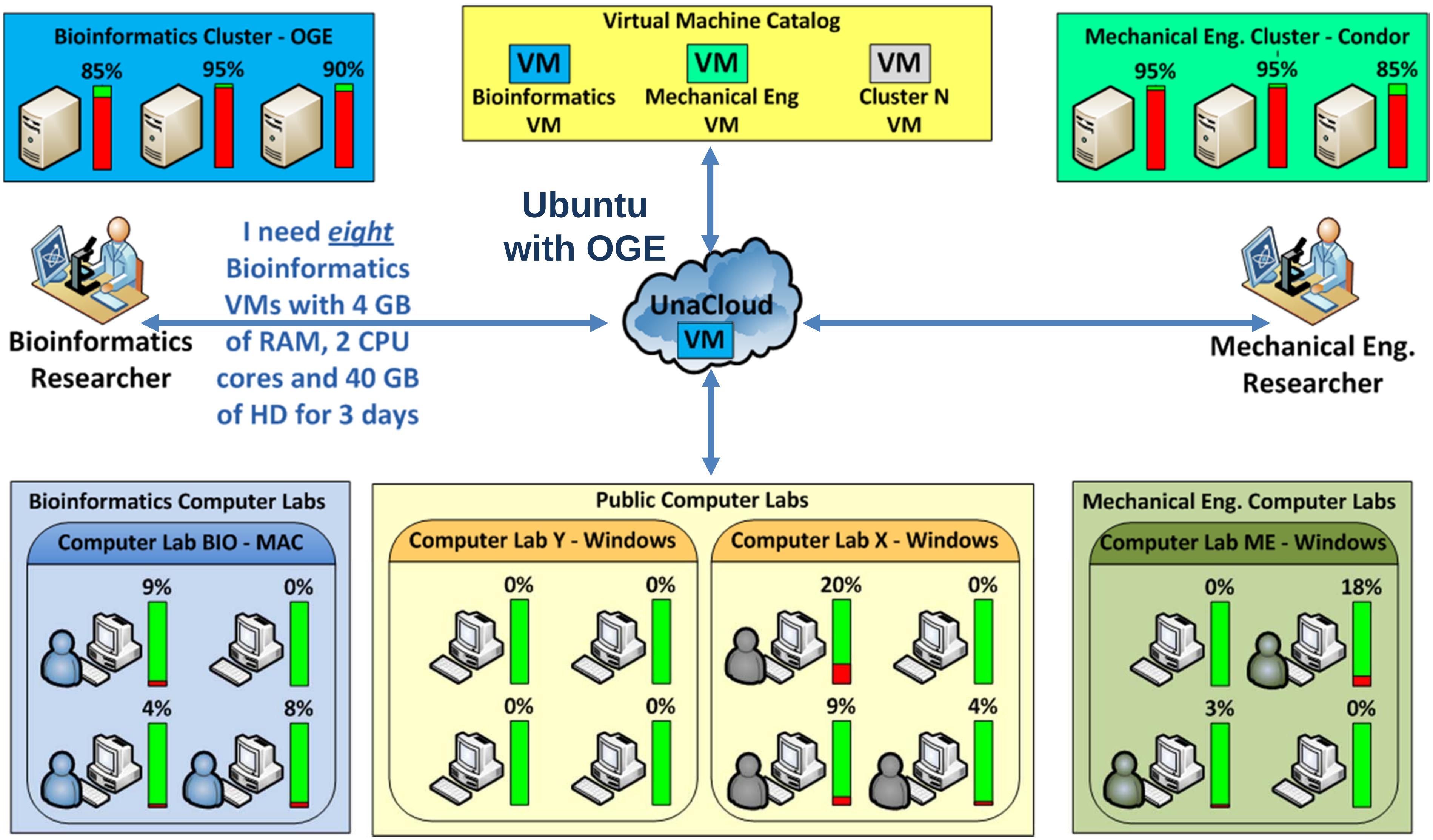
THE DESIRED SOLUTION



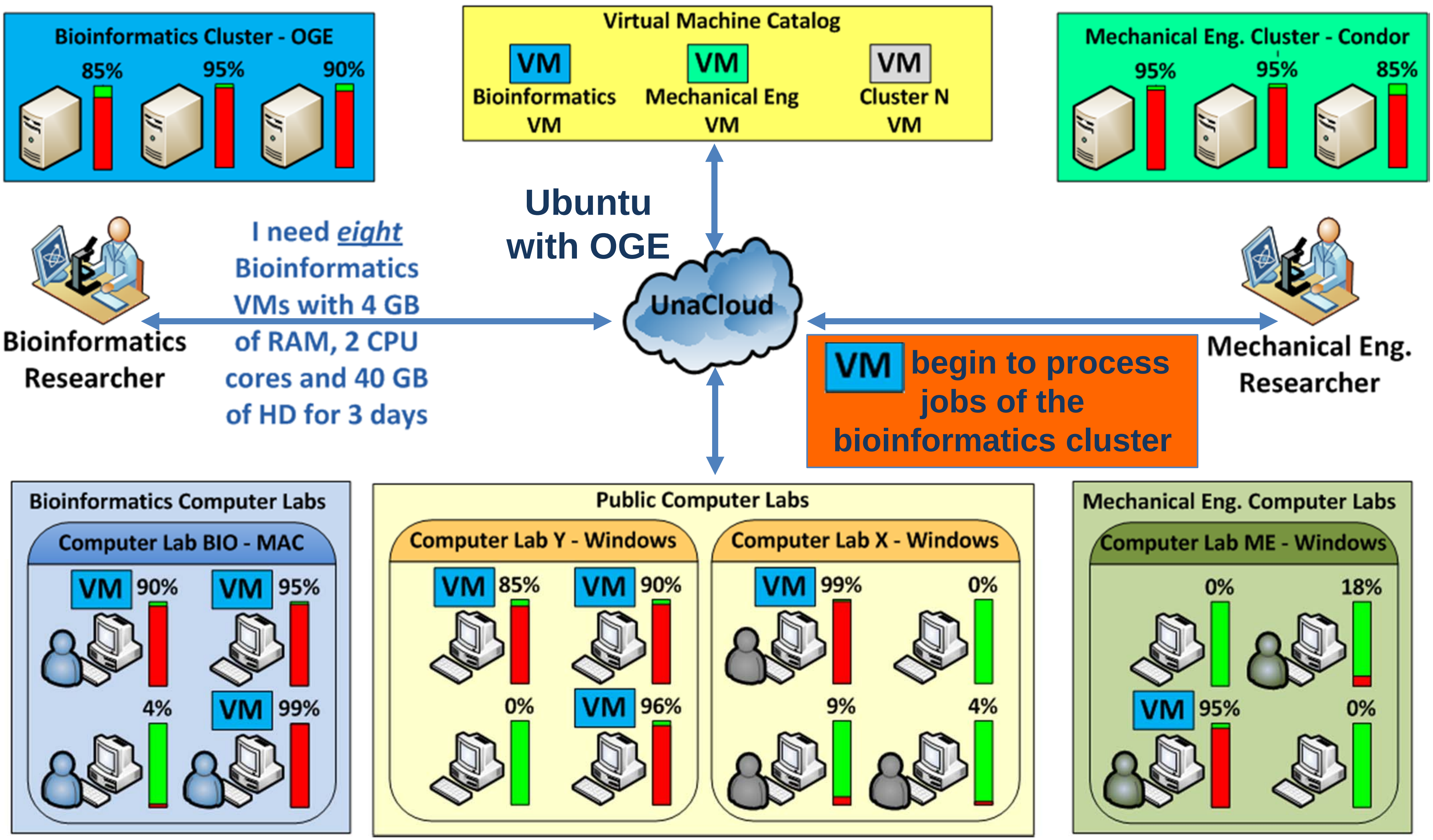
THE DESIRED SOLUTION



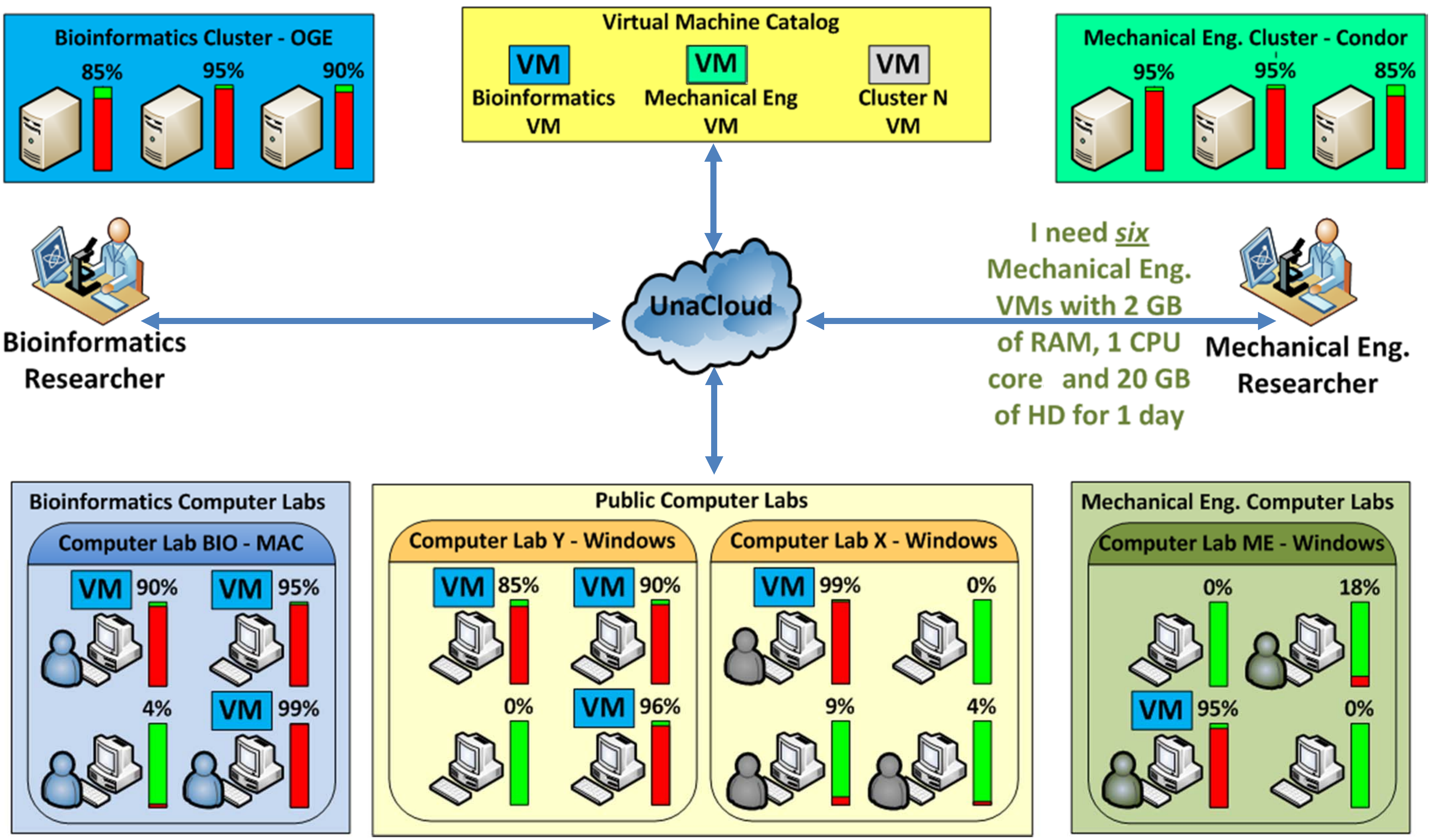
THE DESIRED SOLUTION



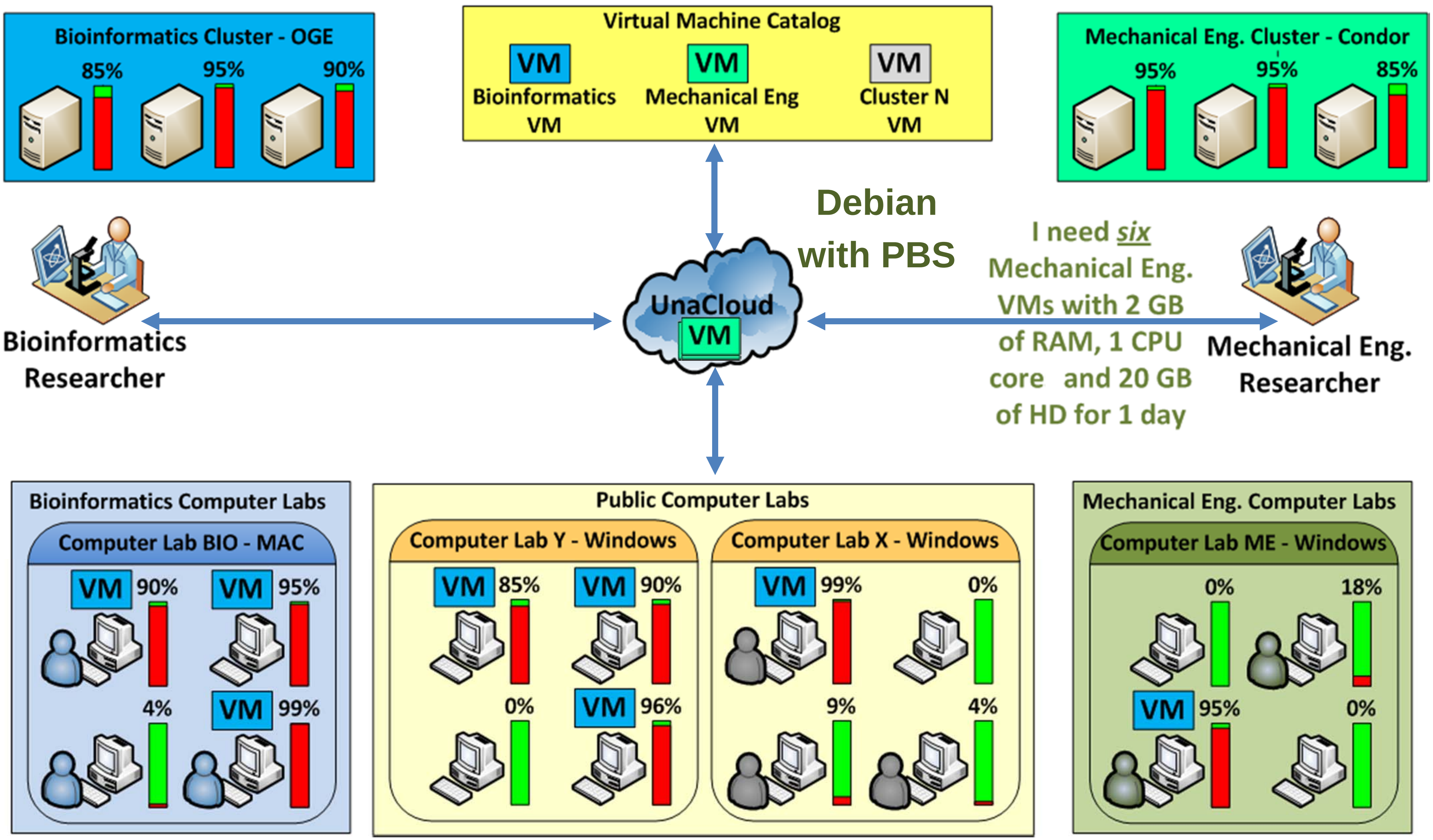
THE DESIRED SOLUTION



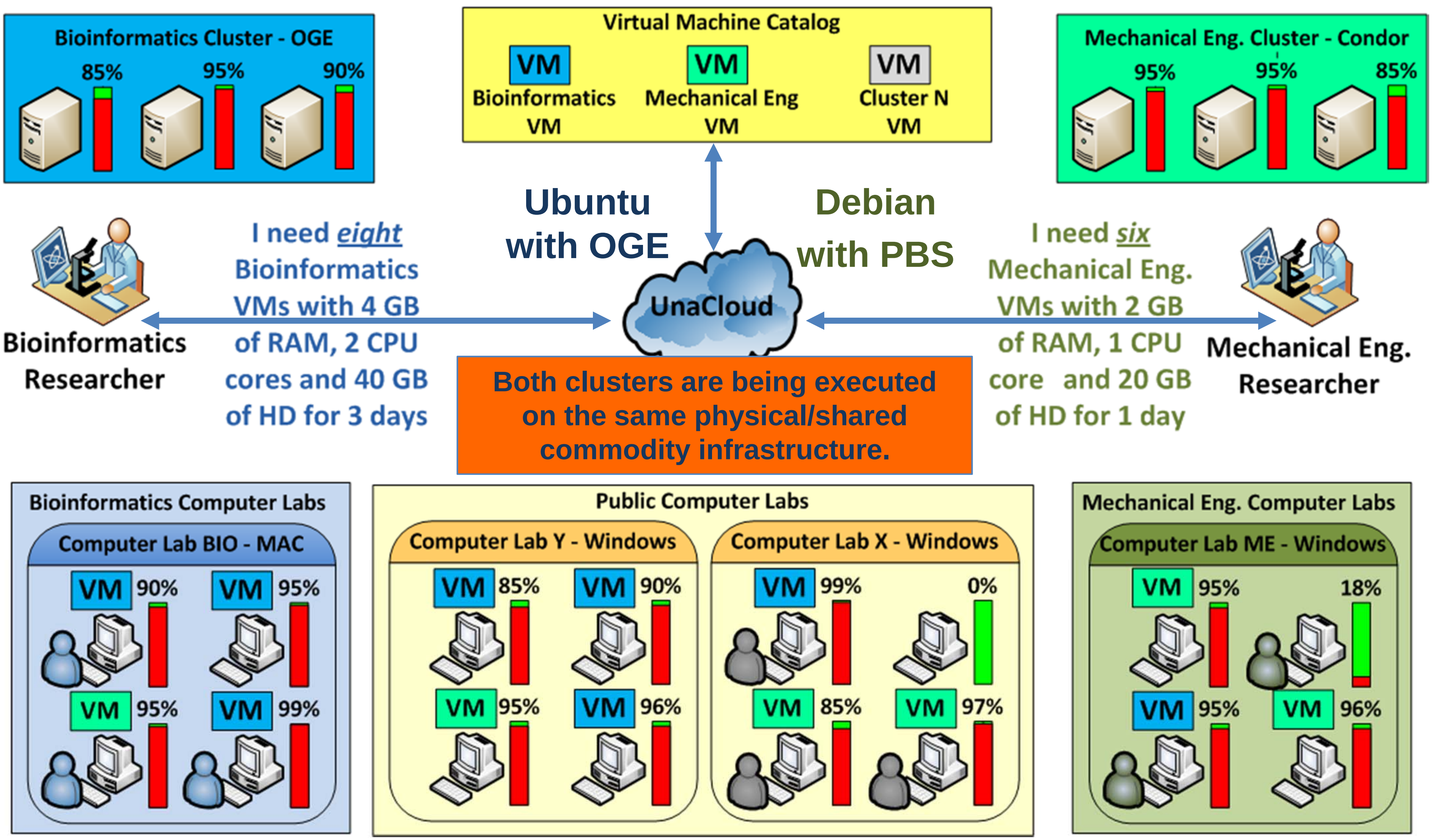
THE DESIRED SOLUTION



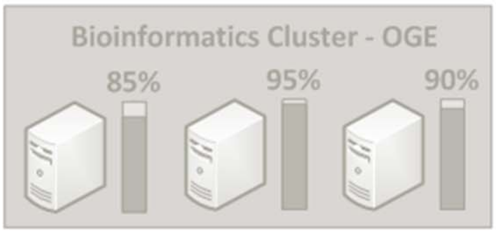
THE DESIRED SOLUTION

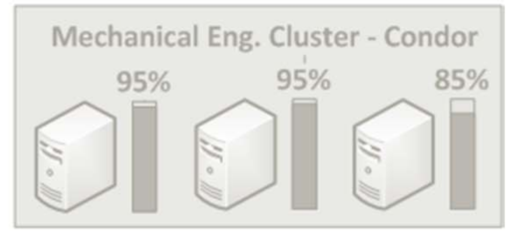


THE DESIRED SOLUTION



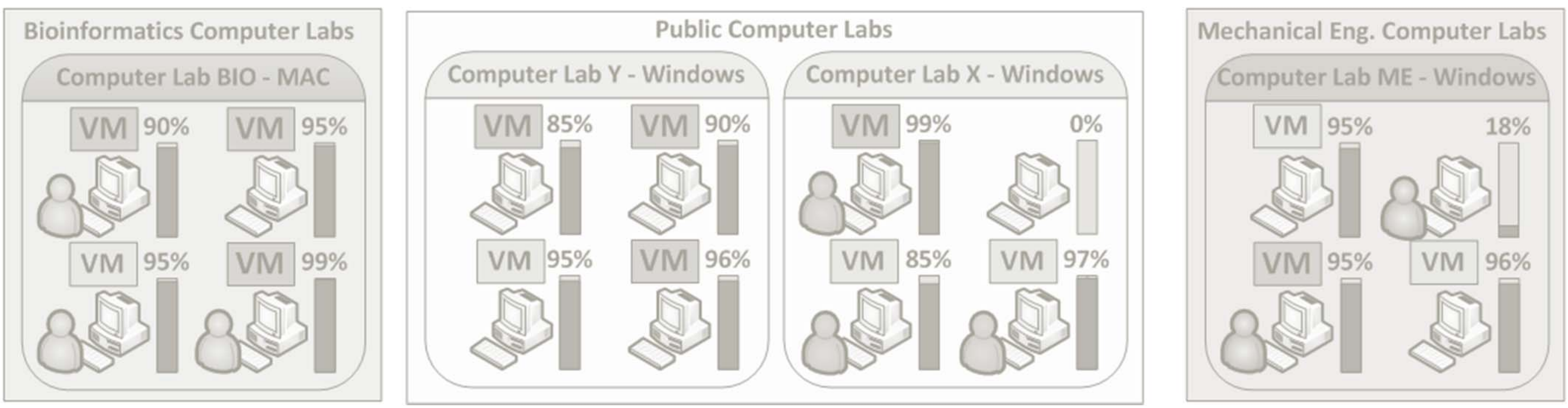
THE DESIRED SOLUTION





UNACLOUD

Research groups can deploy on-demand Computing Services, sharing the same commodity infrastructure. This is achieved using an Opportunistic Infrastructure as a Service Strategy.

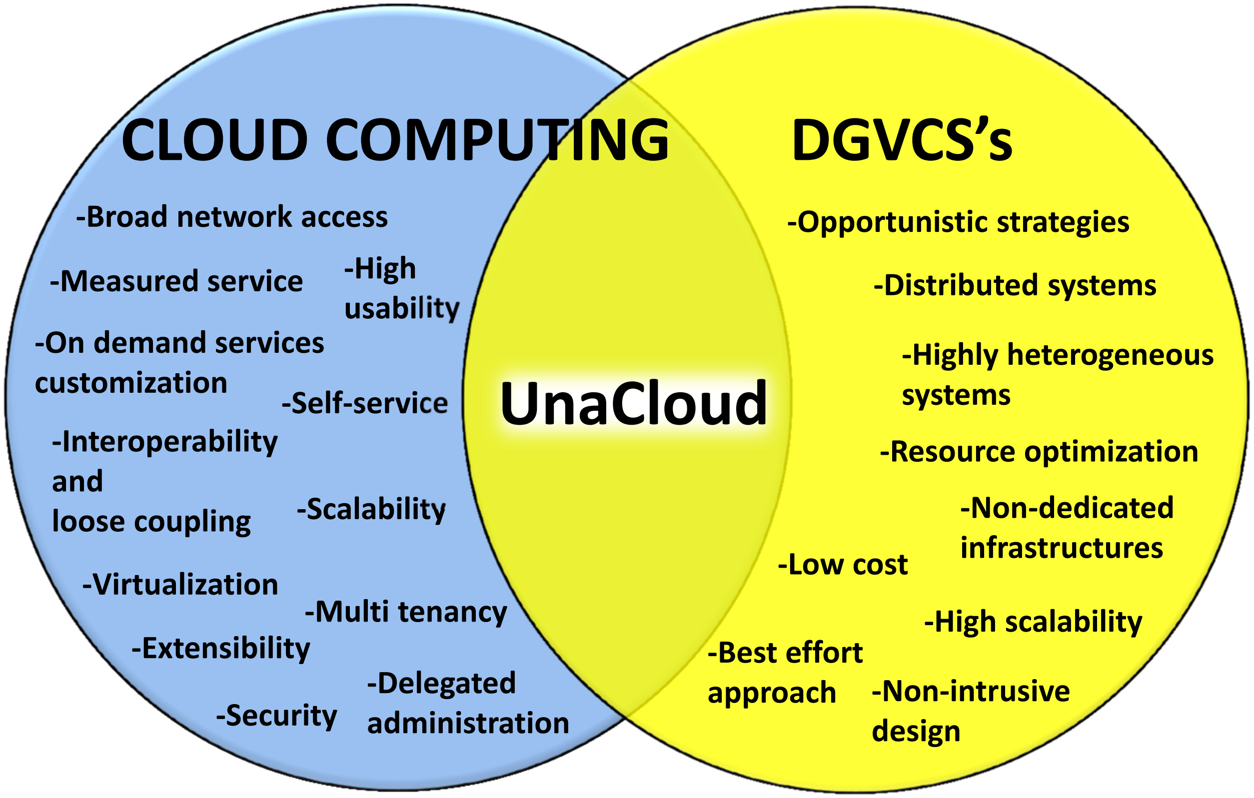


SOLUTION REQUIREMENTS

Our designs have specific requirements:

- **On demand computing services customization and infrastructure assignment to meet specific requirements related to e-Science applications (native executing environments).**
- **High usability user interfaces based on a self-service model to intuitive creation, deployment and management of large computing infrastructures to support e-Science projects.**
- **Strategies to allow the use of idle computing resources in a non-intrusive manner to provide large scale computing capabilities at low cost.**
- **Measured service to record and report the infrastructure user-level consumption.**
- **Delegated administration of the underlying infrastructure complexity used to provide customized computing services to researchers, professors and students.**
- **Solutions based on ubiquitous, interoperability, loose coupling, extensibility and information security principles.**

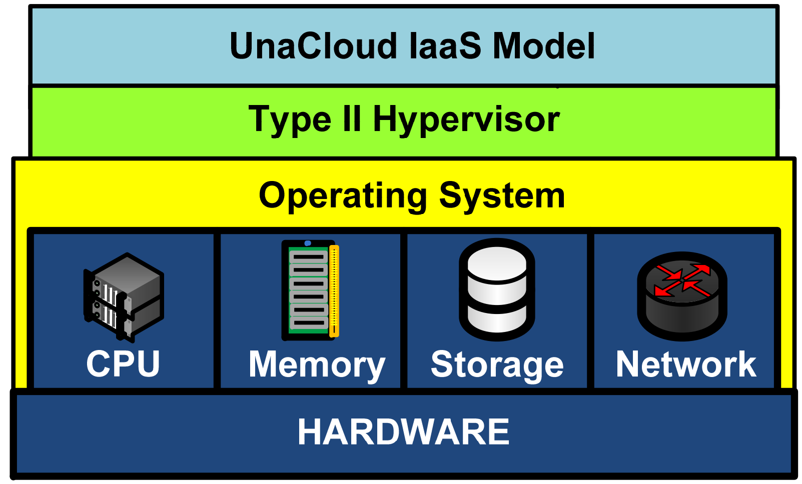
DGVCS's / CLOUD COMPUTING STRATEGY



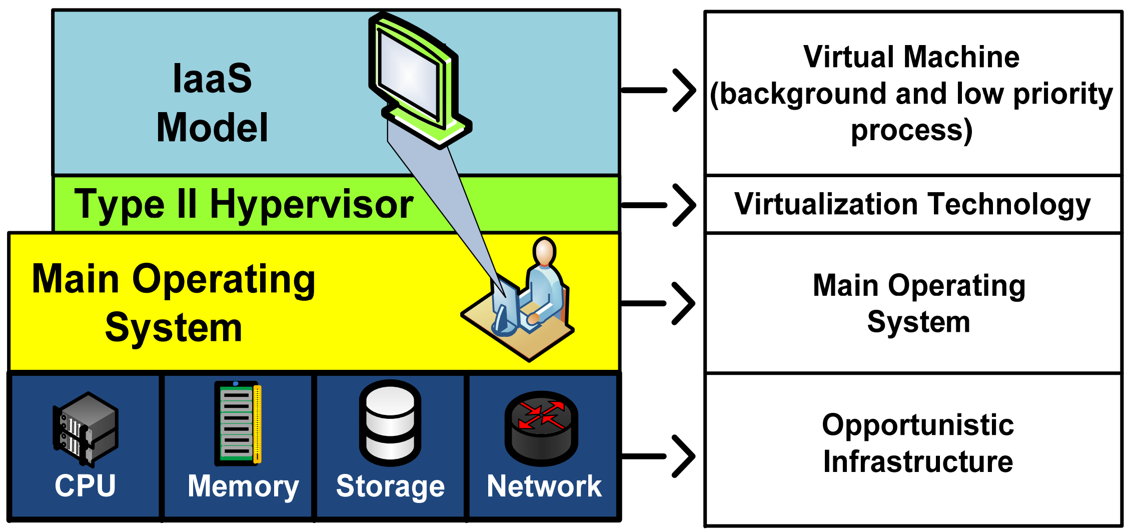
In this work we analyze the prospect and performance of using an opportunistic underlying infrastructure to support a Cloud Computing IaaS model. This is the main motivation and contribution of this research work.

UNACLOUD ARCHITECTURE

- Virtualization technologies allow UnaCloud to access the following advantages: resource optimization, execution environments isolation, on demand deployment and high portability to build a cloud computing IaaS model over type II hypervisor services.

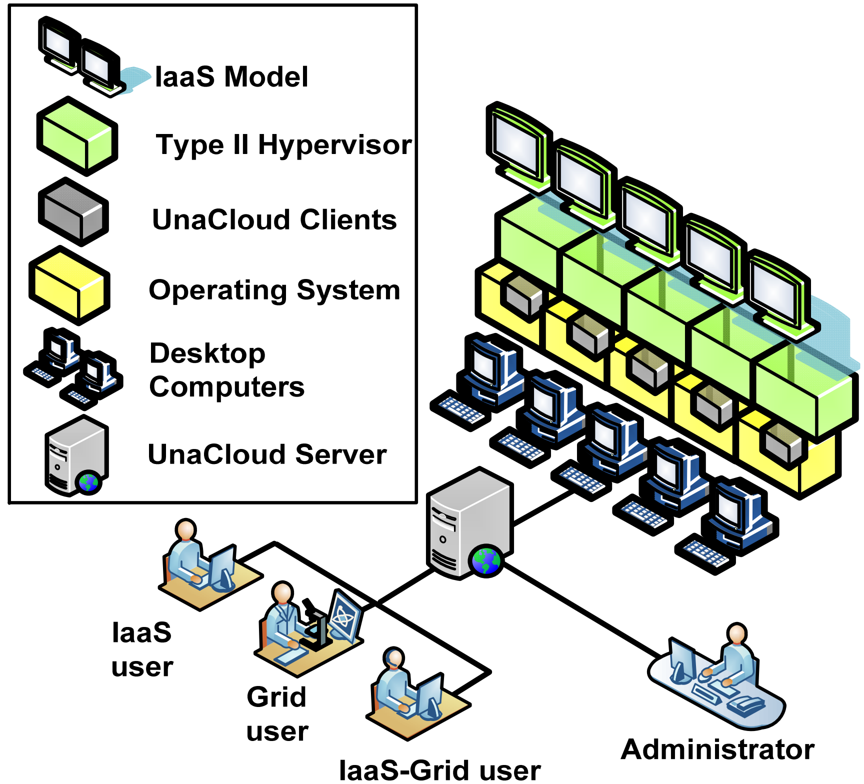


- VMs are executed in background and as low priority processes, allowing that UnaCloud can provide the following advantages: virtual and physical environments isolation and non-intrusive idle resources use.



UNACLOUD ARCHITECTURE

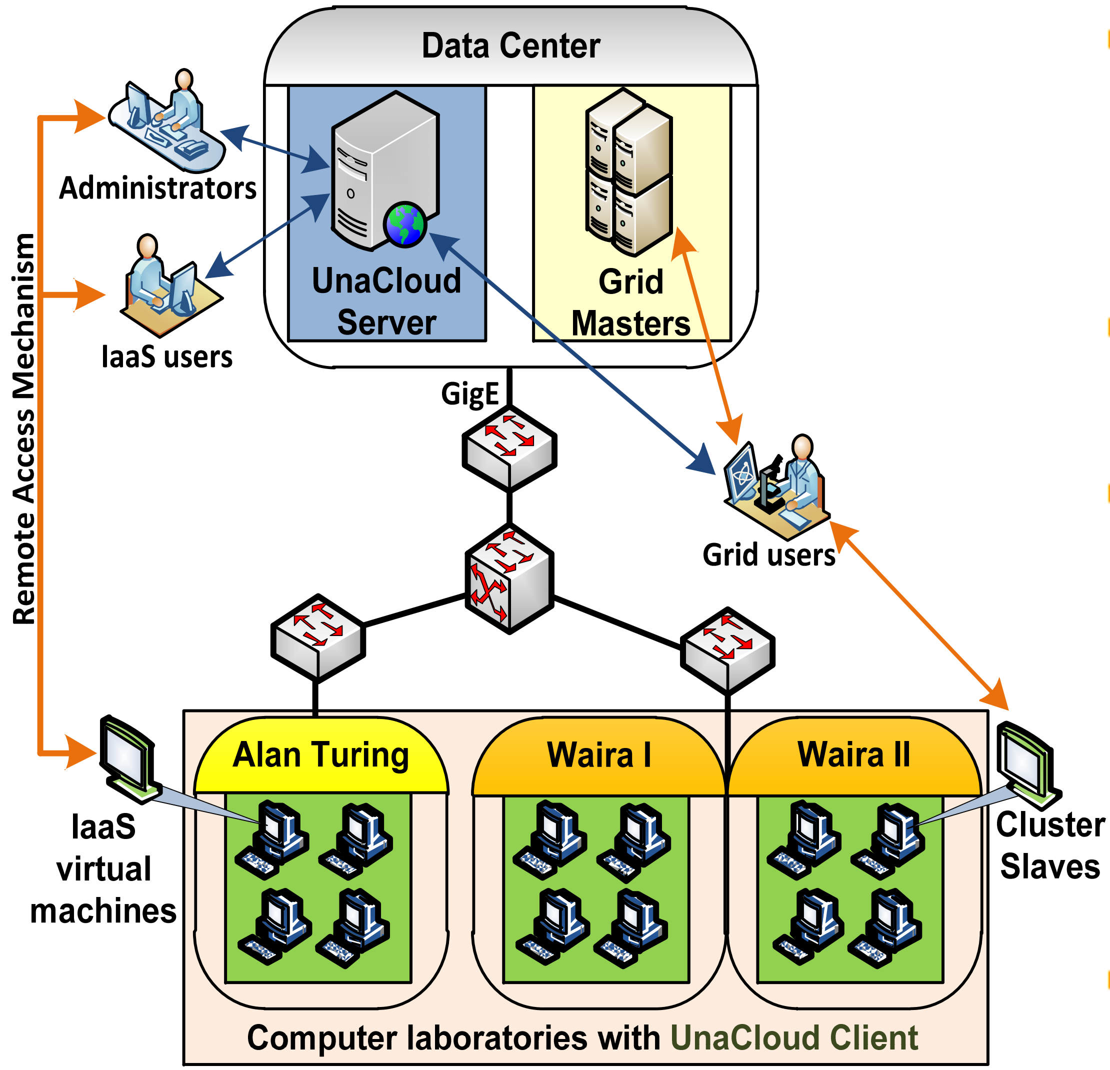
END USER	ROLE
IaaS	Non-located IaaS customization and/or deployment
Grid	Located Grid customization and/or deployment
IaaS-Grid	Any of the above roles
Administrator	Administrative and monitoring exclusive services



UnaCloud architecture is based on the integration of a Web information system with an underlying computing and communication infrastructure

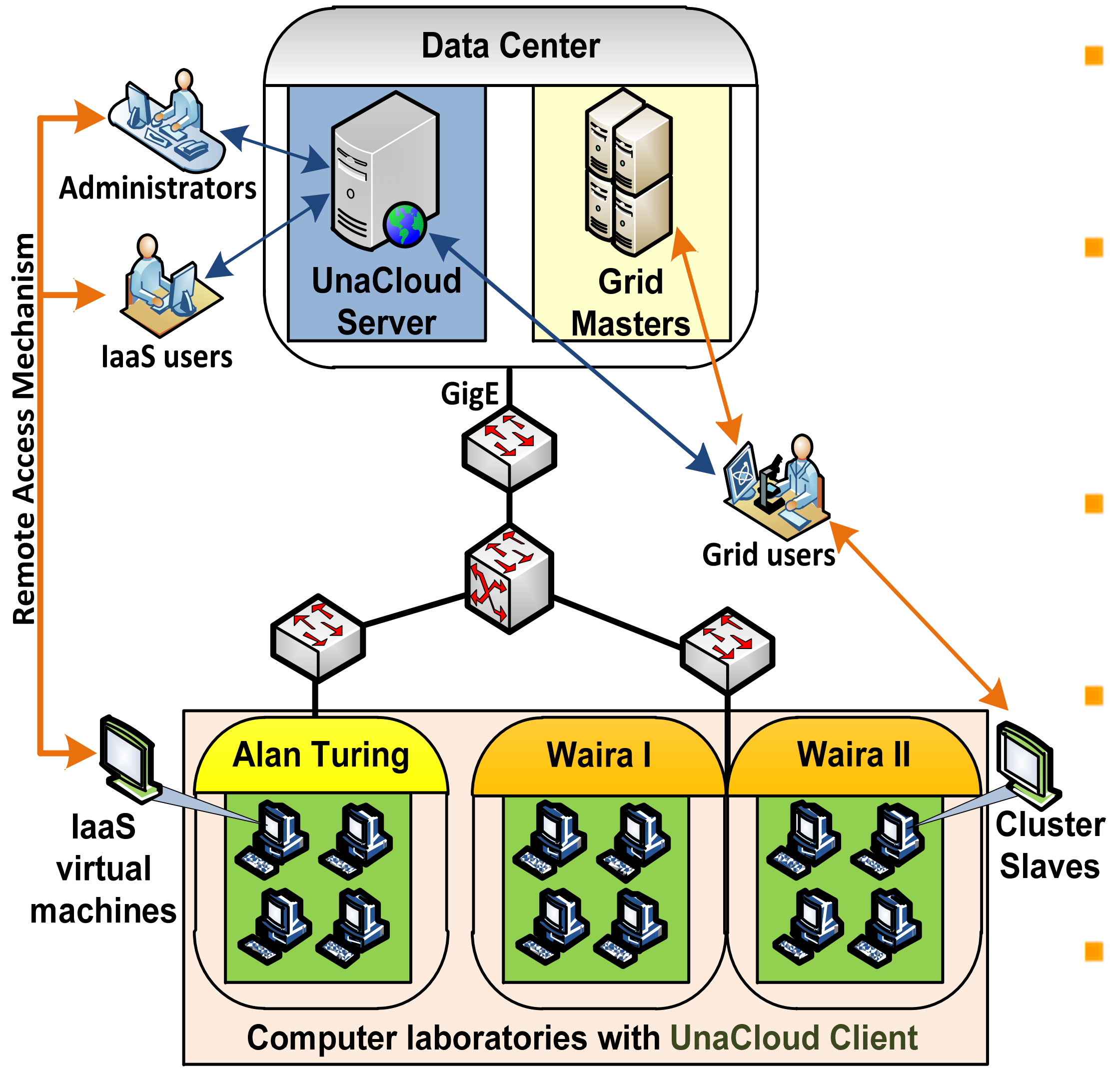
A Web portal (UnaCloud Server) articulated (through UnaCloud Clients) with a non-dedicated infrastructure provides an opportunistic cloud computing IaaS model

UNACLOUD IMPLEMENTATION



- The UnaCloud server was developed using J2EE and MySQL, and it is installed on a dedicated machine.
- The UnaCloud client is a J2SE application.
- The UnaCloud client is deployed in 3 computer labs using 105 desktops computers.
- Desktops have an Intel Quad core processor, 8GB of RAM and Windows 7 pro.
- All nodes are interconnected using a GigE LAN.

UNACLOUD IMPLEMENTATION



- A Web user interface was used, providing a self-service model.
- Deployments on different networks and administrative domains were performed.
- Secure mechanisms were used in message passing.
- Monitoring reports and statistics were generated for UnaCloud users.
- VMware workstation was used as type II hypervisor.

UNACLOUD IMPLEMENTATION

IAAS CUSTOMIZATION	Customizable Virtual Clusters (CVCs) through 5 settings: software, hardware, quantity, location (optional) and execution time.
IAAS DEPLOYMENT	On demand CVC deployment and provision of necessary data to secure remote access.
IAAS ADMINISTRATION	VM operations such as: start, stop, restart, change execution time and monitoring.
IAAS TRACEABILITY	IaaS model traceability at user level with basic reports and statistics.
PHYSICAL INFRASTRUCTURE ADMINISTRATION	Physical machine operations such as: turn off, restart, logout and near-real time monitoring.

UNACLOUD IMPLEMENTATION

IAAS CUSTOMIZATION

- Customizable Virtual Clusters (CVCs) through 5 settings: software, hardware, quantity, location (optional) and execution time.

UNACLOUD IMPLEMENTATION

IAAS CUSTOMIZATION

- Customizable Virtual Clusters (CVCs) through 5 settings: software, hardware, quantity, location (optional) and execution time.

UNACLOUD IMPLEMENTATION

Deployable image: SGE64Bits

Total available instances: 57

Instances to execute: 2

Total execution time: 2 days

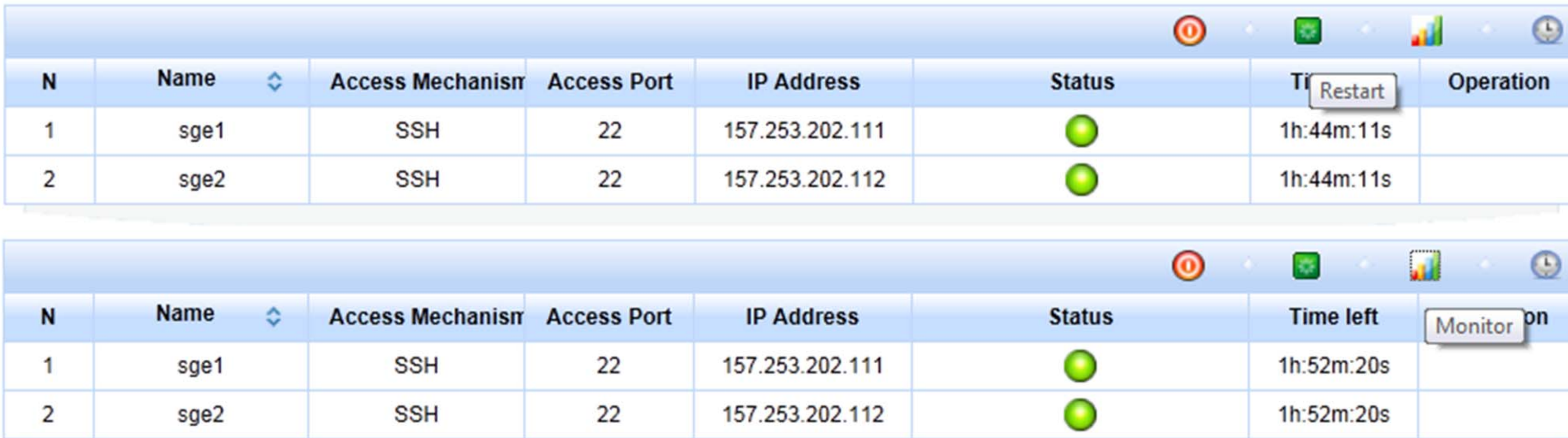
N	Name	Access Mechanism	Access Port	IP Address	Status	Time left	Operation
1	sge1	SSH	22	157.253.202.111		1h:59m:59s	
2	sge2	SSH	22	157.253.202.112	Deploying...	1h:59m:59s	

N	Name	Access Mechanism	Access Port	IP Address	Status	Time left	Operation
1	sge1	SSH	22	157.253.202.111		1h:58m:29s	
2	sge2	SSH	22	157.253.202.112	Machine started	1h:58m:29s	

IAAS DEPLOYMENT

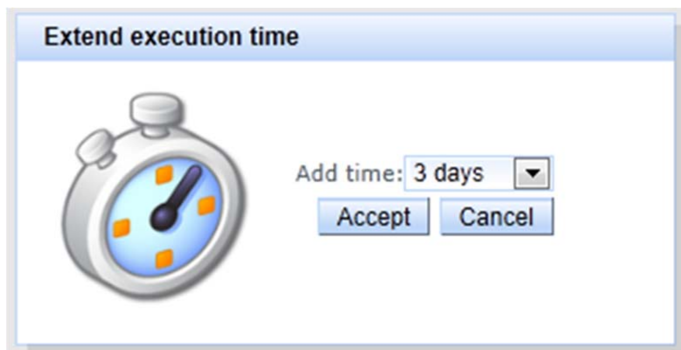
- On demand CVC deployment and provision of necessary data to secure remote access.

UNACLOUD IMPLEMENTATION



N	Name	Access Mechanism	Access Port	IP Address	Status	Time left	Operation
1	sge1	SSH	22	157.253.202.111		1h:44m:11s	
2	sge2	SSH	22	157.253.202.112		1h:44m:11s	

N	Name	Access Mechanism	Access Port	IP Address	Status	Time left	Operation
1	sge1	SSH	22	157.253.202.111		1h:52m:20s	
2	sge2	SSH	22	157.253.202.112		1h:52m:20s	



IAAS ADMINISTRATION

- VM operations such as: start, stop, restart, change execution time and monitoring.

UNACLOUD IMPLEMENTATION

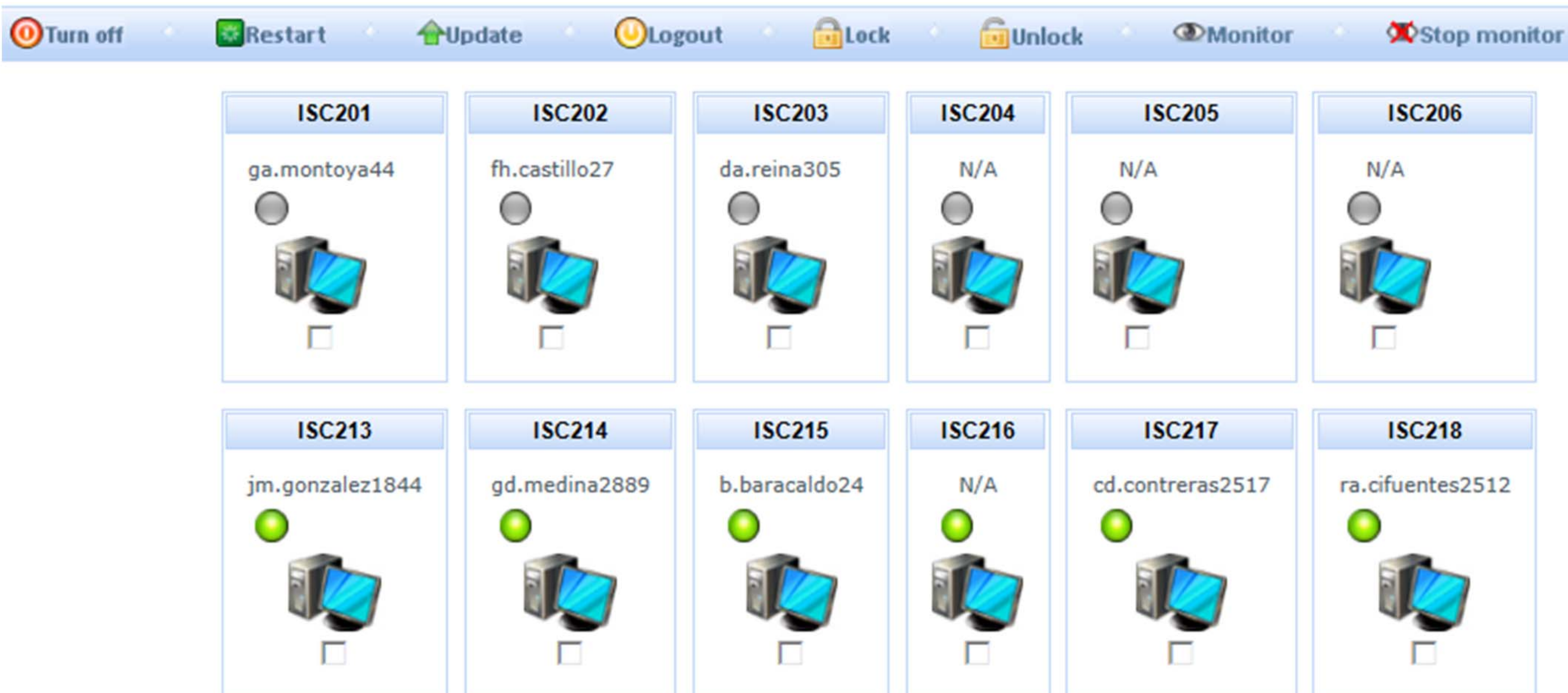
IaaS Traceability

User Name:	mj.villamizar24	**Virtual Machine Name:**	nodo17
Physical Machine Name:	ISC416	**Laboratory Name:**	Alan Turing
Template Name:	Gromacs	**Template Type:**	Cloud
Operating System Version:	Ubuntu 9	**Operating System:**	Linux
Start Time:	Thu Jun 02 11:21:22 COT 2011	**Stop Time:**	Thu Jun 02 13:21:22 COT 2011
Status:	Off	**Hard Disk:**	20
CPU Cores:	1	**RAM Memory:**	2048
IP Address:	157.253.201.206	**MAC Address:**	28:bd:68:38:2e:30
Hypervisor:	VMware Workstation 6.5	**Security Schema:**	SGE

IAAS TRACEABILITY

- **IaaS model traceability at user level with basic reports and statistics.**

UNACLOUD IMPLEMENTATION



PHYSICAL INFRASTRUCTURE ADMINISTRATION

- Physical machine operations such as: turn off, restart, logout and near-real time monitoring.

UNACLOUD TESTING AND RESULTS

	IAAS USERS
VIRTUAL MACHINES	70 academic VMs
DEPLOYMENT MODE	Non-located deployment
DEPLOYMENT TIME	13 seconds

	GRID USERS	
VIRTUAL MACHINES	70 bioinformatics VMs	35 chemical engineering VMs
DEPLOYMENT MODE	Located deployment	
DEPLOYMENT TIME	15 seconds	7 seconds

- **The average time to load the guest operating systems is 4-5 minutes.**

UNACLOUD TESTING AND RESULTS

APPLICATION NAME	INFRASTRUCTURE USED	CPU NUMBER	JOB NUMBER	TIME BY JOB (SEC)	EXECUTION TIME (DAYS)
BSGrid Model A	PC	2	150000	35	30,38
	Chemical Eng. CVC	70	150000	85	2,11
BSGrid Model B	PC	2	150000	63	54,69
	Chemical Eng. CVC	70	150000	111	2,75
HMMER	PC	2	4200	11700	284,40
	Biological Science CVC	140	4200	12900	4,50

- Performance degradation perceived by owner users (students or administrative personal) is less than 3%.
- The maximum overload of grid jobs executed on UnaCloud virtual machines is of 17%.
- To avoid resource competition among virtual machines only one virtual machine is executed on each desktop.

UNACLOUD OPPORTUNITIES AND LIMITATIONS

REQUIRED FEATURES	UNACLOUD	
USABILITY	High usability Web user interfaces, which operation is almost intuitive, requiring basic IT knowledge	✓
SELF-SERVICE	Unilaterally computing service provision	✓
BROAD NETWORK ACCESS	Web portal available over Intranet and Internet	✓
ON DEMAND SERVICES CUSTOMIZATION	On demand computing services customization, even to meet large scale computational requirements	✓
HARDWARE MULTI TENANCY	Opportunistic use of idle computing resources	✓
VIRTUALIZATION	On demand VM deployment through virtualization	✓
SCALABILITY	Horizontal scaling model based on private clouds	✓
INTEROPERABILITY AND LOOSE COUPLING	Web and service oriented architecture	✓
EXTENSIBILITY	Use of open source tools, broadly diffused	✓
DELEGATED ADMINISTRATION	Services to support common administration tasks	✓
SECURITY	Authentication, authorization, confidentiality and non-repudiation mechanisms to secure the deployments	✓
MEASURED SERVICE	Infrastructure traceability at user level with basic reports	✓
QOS AND SLAs	Best effort approach due to the infrastructure availability	✗

CONCLUSIONS

- UnaCloud validates the convergence of cloud computing and DGVCS's, offering promising opportunities to meet customized computational requirements thought the use of an open source, low cost, extensible, interoperable, efficient, scalable, secure and opportunistic IaaS model.

- UnaCloud provides a multipurpose cloud computing experimental platform to deploy Customizable Virtual Clusters that support new specific computational requirements of academic and research projects.

- UnaCloud represents an economically attractive solution for building and deploying large scale computing infrastructures.

- UnaCloud cloud computing features are promising to reduce the development cycle and the generation of results time of projects depending on the agile and flexible provisioning and sharing of low cost computing resources.

FUTURE WORK

RELATED TO CLOUD COMPUTING	
SERVICE MODELS:	PaaS and SaaS
DEPLOYMENT MODELS:	Public, community and hybrid
INTEROPERABILITY:	Amazon, Eucalyptus, etc.
IAAS MODEL:	On demand networking customization
HYPERVISORS:	Compatibility with type I and other type II hypervisors
STORAGE:	Decoupled and persistent
PAY-PER-USE MODEL:	Based on computing, storage, backup and networking resources consumption
SECURITY:	Stronger information security architecture

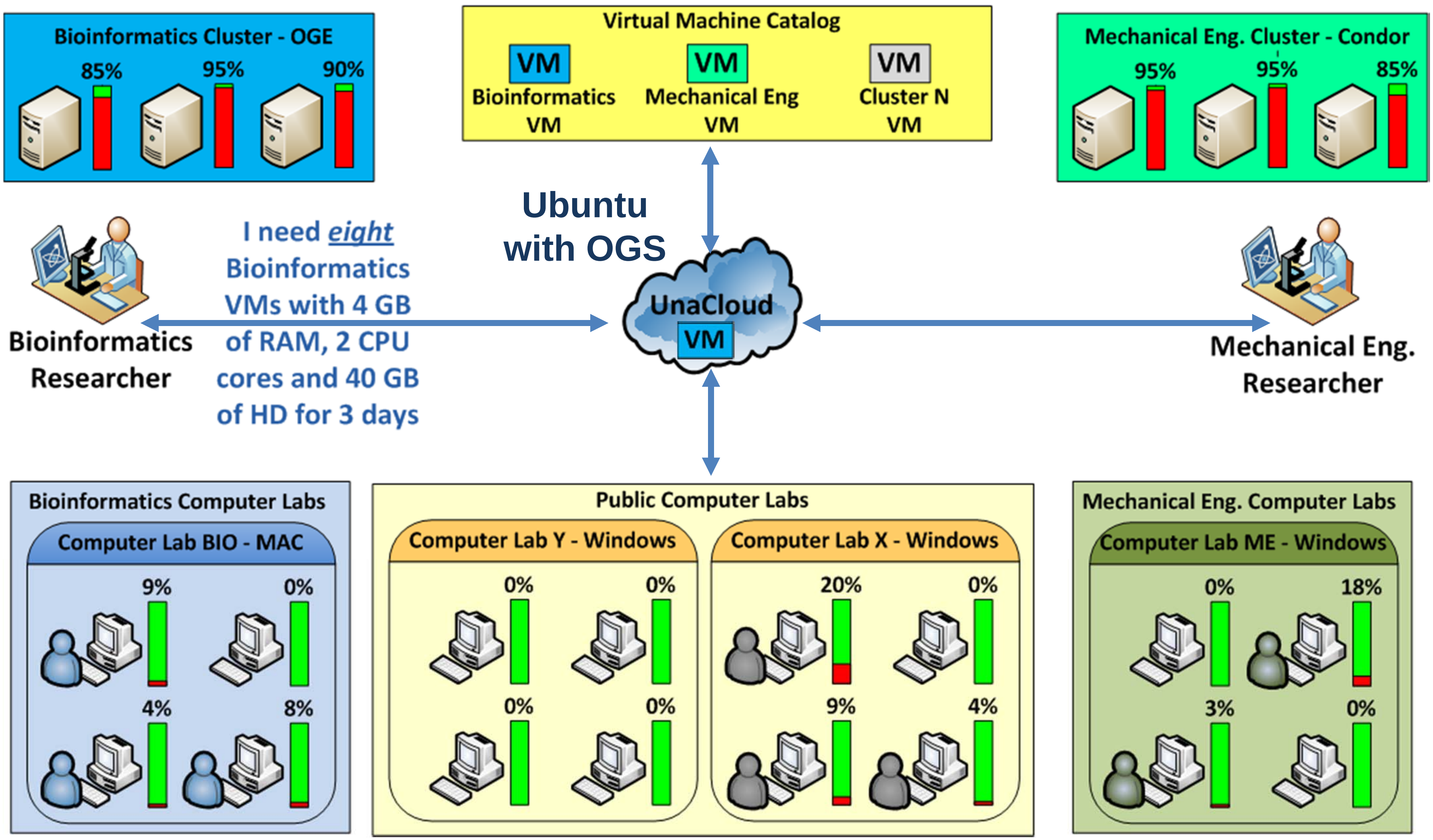
RELATED TO DGVCS's	
QUALITY OF SERVICE (QoS):	Statistic QoS approach
TESTS AND VALIDATIONS:	Comprehensive testing and model validations

RELATED TO ADMINISTRATION	
INTERFACES:	Secure command line interfaces
INITIAL DEPLOYMENT:	First deployment and VM copy automation

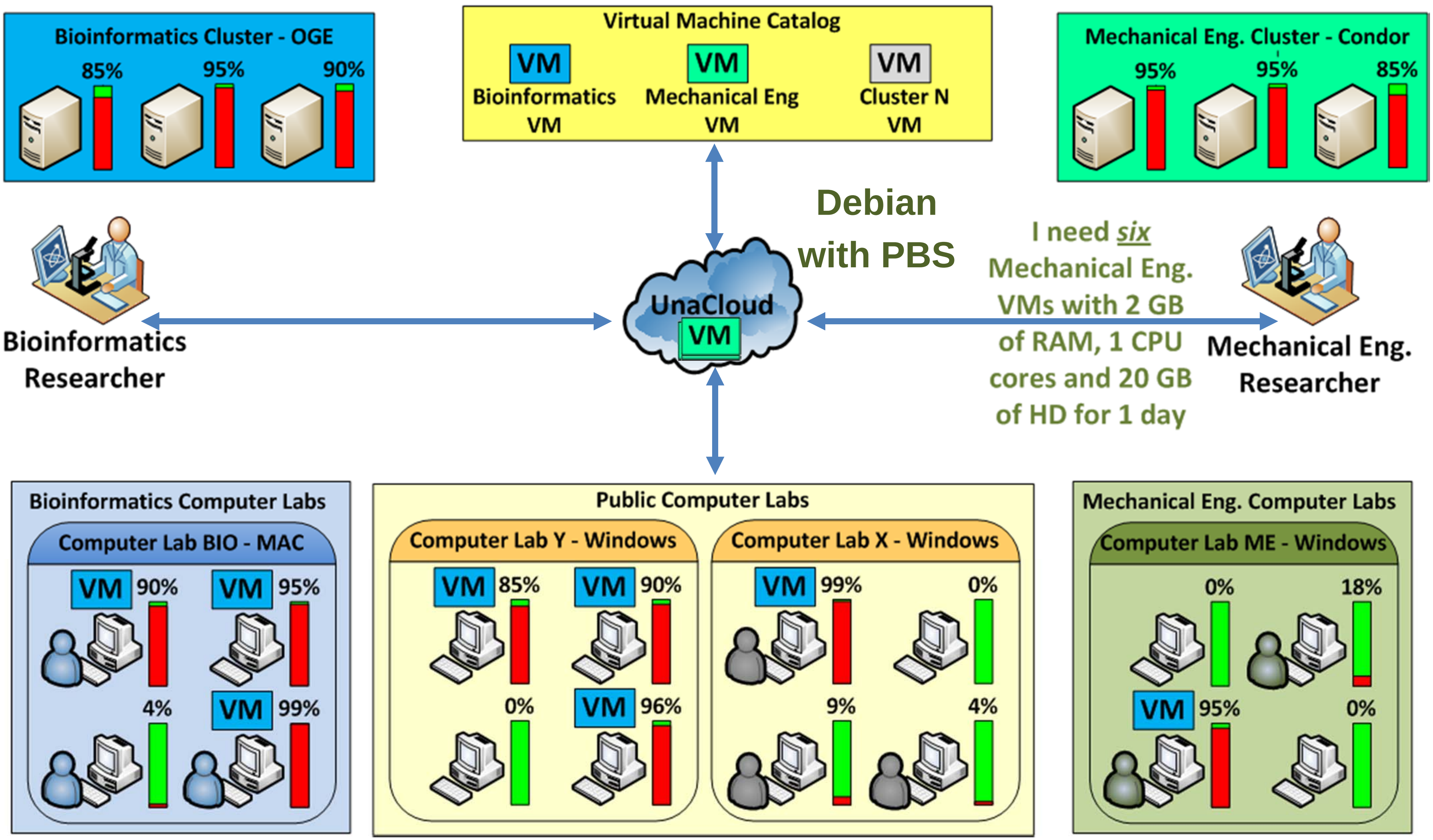
…

THANKS FOR YOUR ATTENTION!

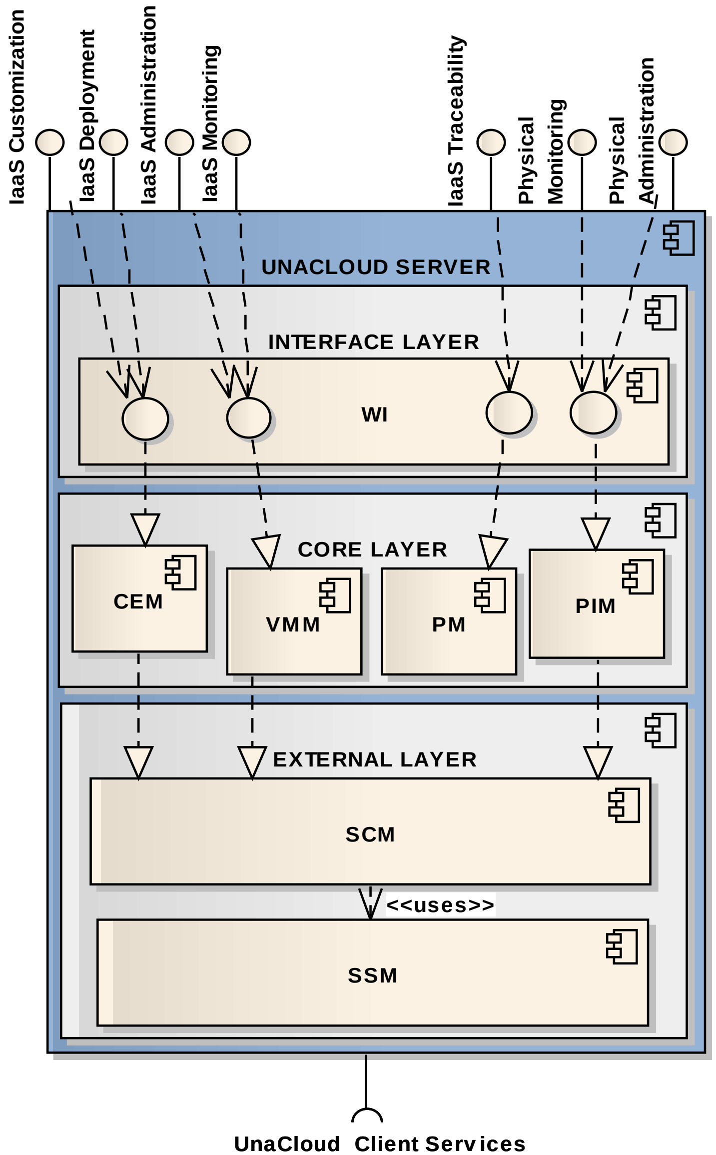
THE DESIRED SOLUTION



THE DESIRED SOLUTION



UNACLOUD SERVER ARCHITECTURE



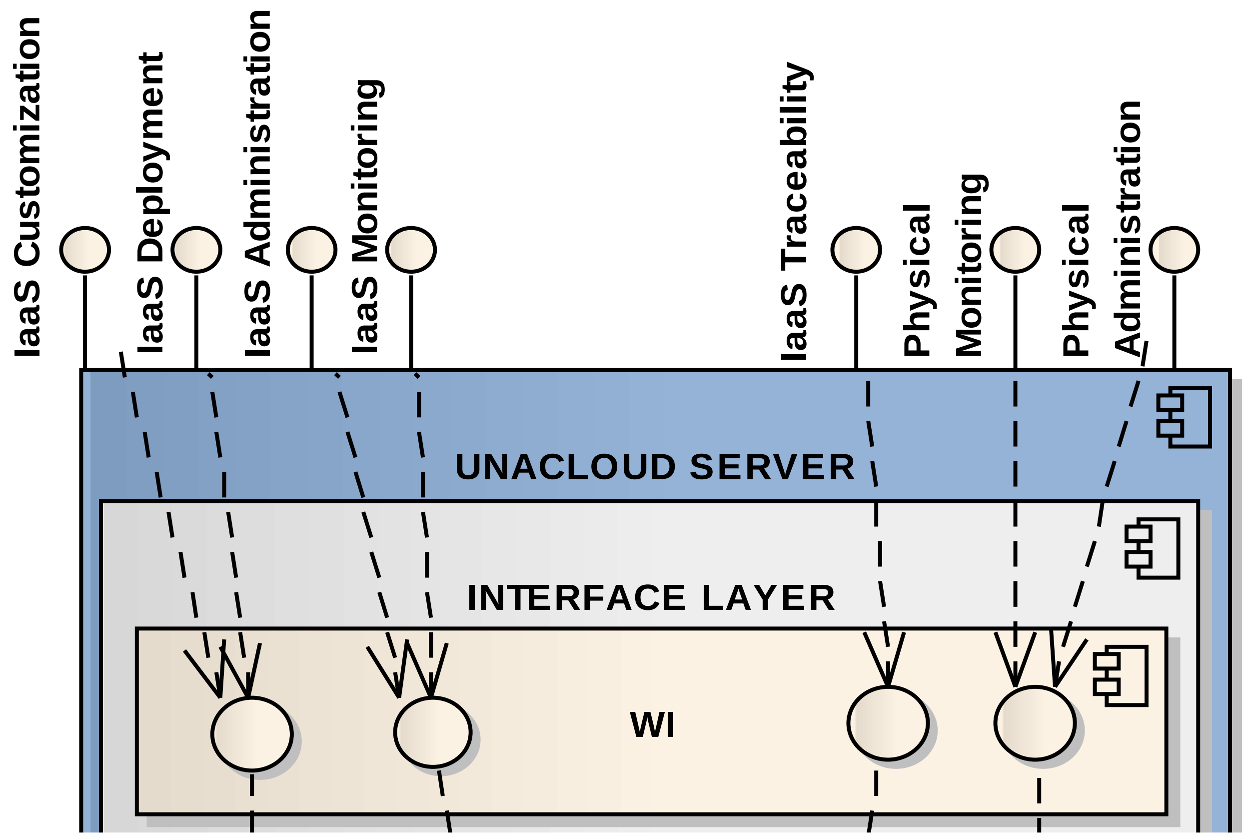
The UnaCloud server architecture is based on three layers:

Layer 1 - Interface Layer

Layer 2 - Core layer

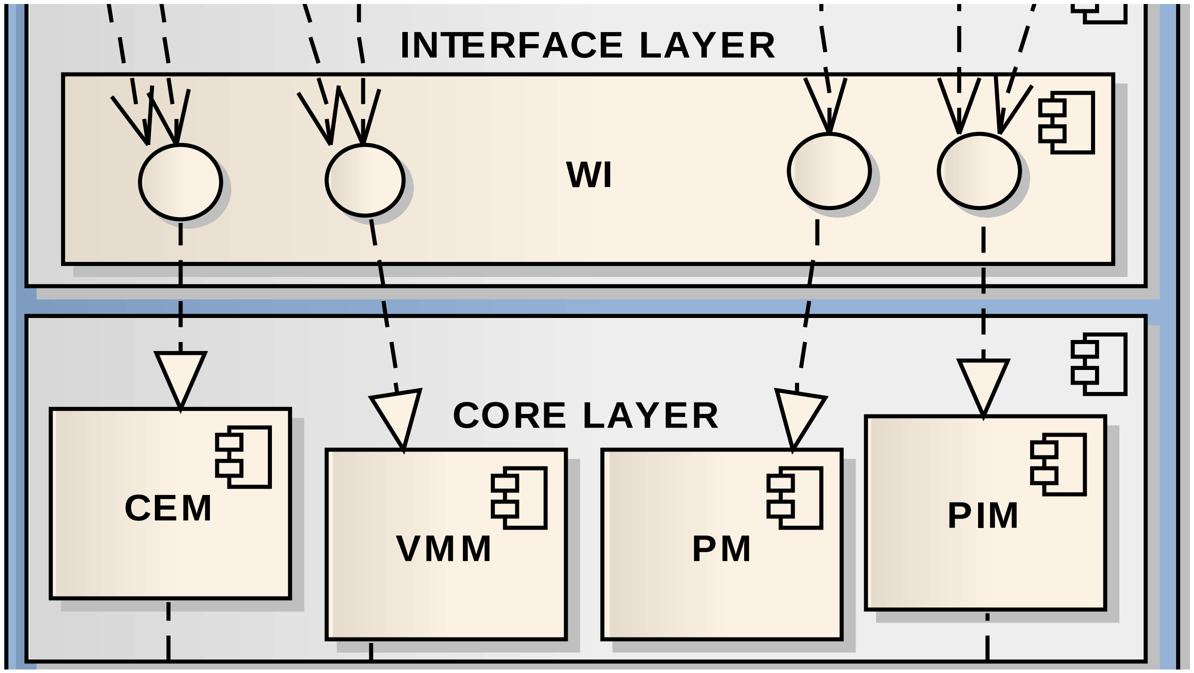
Layer 3 - External layer

UNACLOUD SERVER ARCHITECTURE



Layer 1 - Interface Layer: A Web Interface (WI) used by users (IaaS, Grid and Admin) for accessing and consuming the services through authentication and authorization mechanisms.

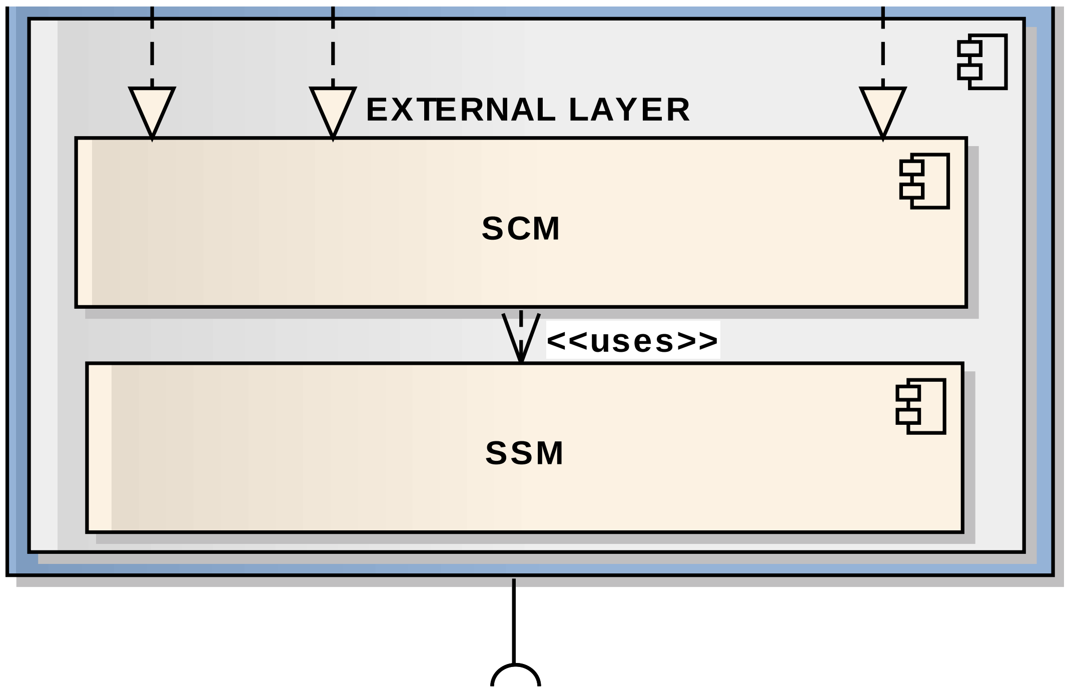
UNACLOUD SERVER ARCHITECTURE



Layer 2 - Core layer: it is responsible for processing all user requirements and delivering the solutions desired. It has four main components:

Customized Environment Manager (CEM): it prepares user orders
Persistence Manager (PM): information database
Virtual Machine Manager (VMM): it prepares hypervisor orders
Physical Infrastructure Manager (PIM): desktops computers of physical labs

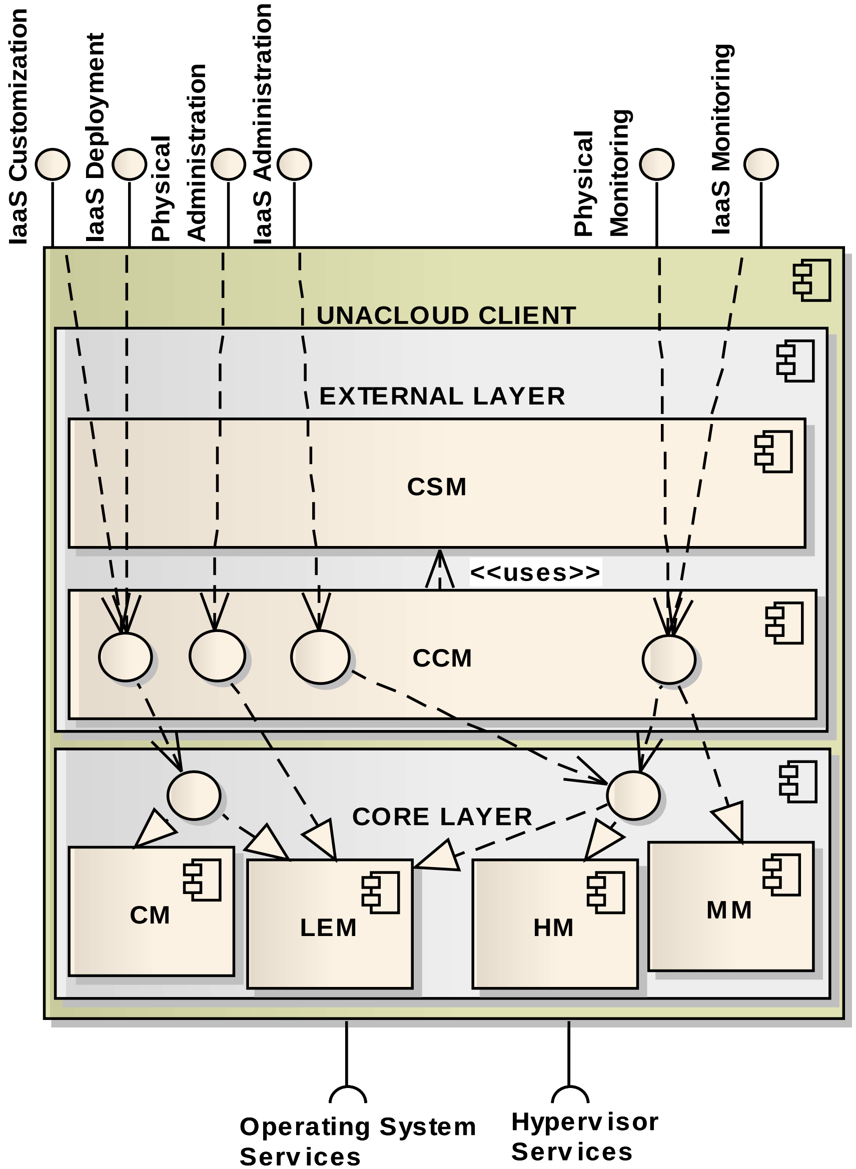
UNACLOUD SERVER ARCHITECTURE



Layer 3 - External layer: it is responsible of providing secure communications with the UnaCloud clients. It has two main components:

Server Communication Manager (SCM): connection, disconnection and message passing between UnaCloud Server and UnaCloud Clients
Server Security Manager (SSM): it provides security in communications

UNACLOUD CLIENT ARCHITECTURE



The Client architecture is based on two layers:

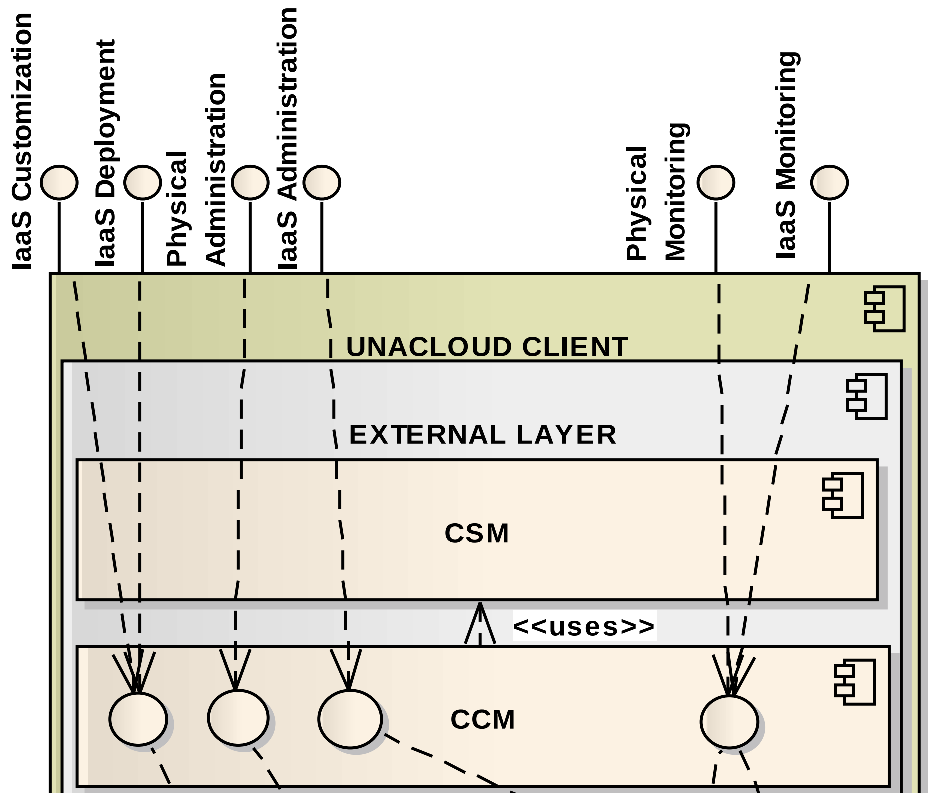
Layer 1 - External Layer

Layer 2 - Core layer

It is a lightweight, highly portable and easy to install client which is installed and run directly on each desktop.

It is responsible for receiving and processing all of UnaCloud Server orders to provide a dynamic and on-demand IaaS model.

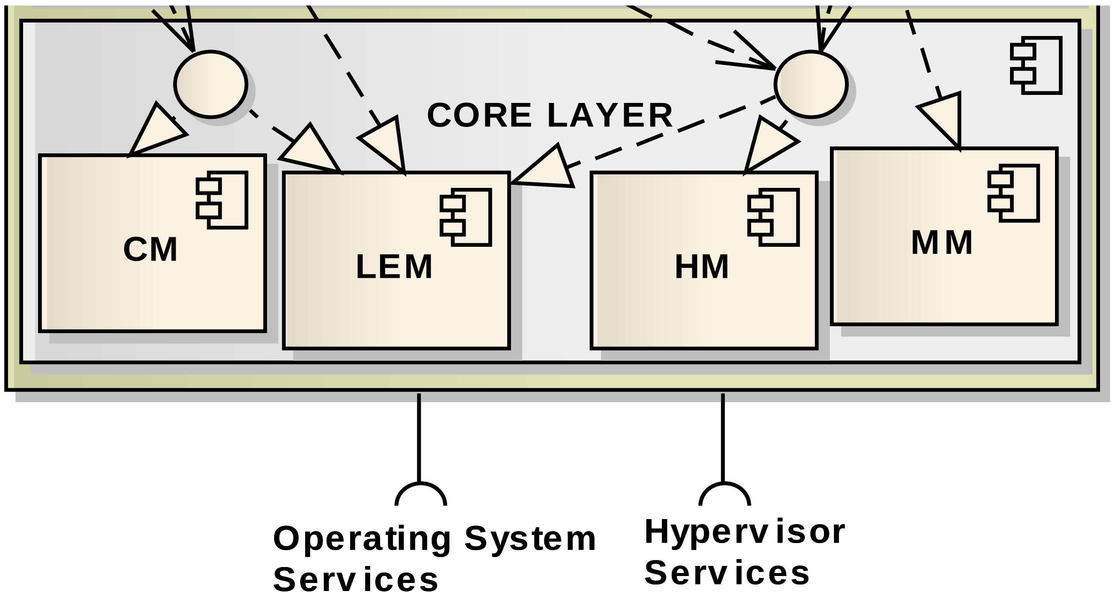
UNACLOUD CLIENT ARCHITECTURE



Layer 1 – External Layer: it is responsible for managing secure communication services on the client side. It has two main components:

Client Communication Manager (CCM): message passing with the server
Client Security Manager (CSM): it provides security in communications

UNACLOUD CLIENT ARCHITECTURE



Layer 2 - Core layer: it is responsible for attending and meeting UnaCloud Server orders through local operating system and hypervisor invocations. It has for main components:

Context Manager (CM): adapt the virtual machine execution context
Local Executor Manager (LEM): command execution on the local operating system
Hypervisor Manager (HM): commands to invocate the local hypervisor
Monitoring Manager (MM): monitoring the physical and virtual machine