



# Esperienze di Cloud computing nelle comunità scientifiche italiane

Luciano Gaido

INFN-Torino & IGI

Workshop congiunto INFN CCR/GARR

Napoli, 14-17 Maggio 2012



## Il Cloud computing



- Da alcuni anni è il paradigma di calcolo che riscuote maggiore successo:
  - sono disponibili vari servizi di tipo commerciale (Amazon EC2, Azure, ....)
  - alcune soluzioni open source sembrano molto interessanti (OpenNebula, Eucaliptus, Openstack, ...)
- L'accesso e l'utilizzo sono più semplici rispetto alla tecnologia grid, ma:
  - c'è il problema dei costi per le soluzioni commerciali
  - e quello della robustezza e scalabilità per le soluzioni open source



## Cloud in italia



- anche in Italia ci sono moltissime esperienze sulle cloud sia nel settore privato (ad es. Telecom Italia, Aruba, Reply, etc.) sia nel mondo dell'università e della ricerca
- grande interesse per l'offerta di servizi cloud per la pubblica amministrazione (vedi anche il PON Smart Cities)
- vedi i numerosi talk di questa sessione e i poster
- la “biodiversità” è un segno del grande interesse per la tecnologia cloud ed è anche in parte dovuta all'atteggiamento empirico caratteristico del mondo della ricerca



## Tutto risolto?



- Alcune questioni importanti sono ancora aperte:
  - scalabilità
  - federazione/interconnessione di cloud diverse (per evitare il lock-in)
  - mancanza di strumenti generali e solidi di contorno (gestione immagini, autenticazione/autorizzazione, portali)
  - lo storage
- Inoltre ci sono esigenze differenti, da parte dei diversi soggetti interessati (utenti, fornitori di risorse e fornitori di tecnologia) → sono distribuite nei 3 tipi di cloud definite dal NIST:
  - Infrastructure as a Service (IaaS)
  - Platform as a Service (PaaS)
  - Software as a Service (SaaS)
- Ma esistono altre varianti, ad es. HPC as a Service (HaaS)



## IGI e le cloud



- A dicembre 2011 IGI ha organizzato un workshop sulle cloud in Italia per capire quali erano le attività interessanti in questo settore e definire una piattaforma di riferimento:
  - Da integrare nell'infrastruttura grid, non per la sua sostituzione
- Presentazioni di (grazie per le prossime slide!):
  - Giovanni Ponti (ENEA)
  - G.B. Barone d D. Bottalico (Università di Napoli)
  - Mehdi Sheikhalishahi (Università della Calabria)
  - Davide Salomoni e Giuseppe Andronico (INFN)
- Esperienze molto diverse tra loro sia per scopi, soluzioni tecnologiche e scala.



## ENEA: Conclusioni e Sviluppi futuri



**OpenNebula è una valida scelta open source per il cloud computing**

### ***Vantaggi:***

- Buon supporto per gli hypervisor
- Ottima GUI di management (Sunstone)
- Facilmente customizzabile (anche a basso livello)
- Molto popolare in ambienti di ricerca
- Progetto molto attivo

### Ottima integrazione in ENEA-GRID! *Future Work*

- Ampliare la gamma degli hypervisor
- Definire politiche di utilizzo delle risorse virtuali basate su tipologie di utenti e sui gruppi
- Installare altre piattaforme di cloud (e.g., Eucalyptus) e confrontare le performance

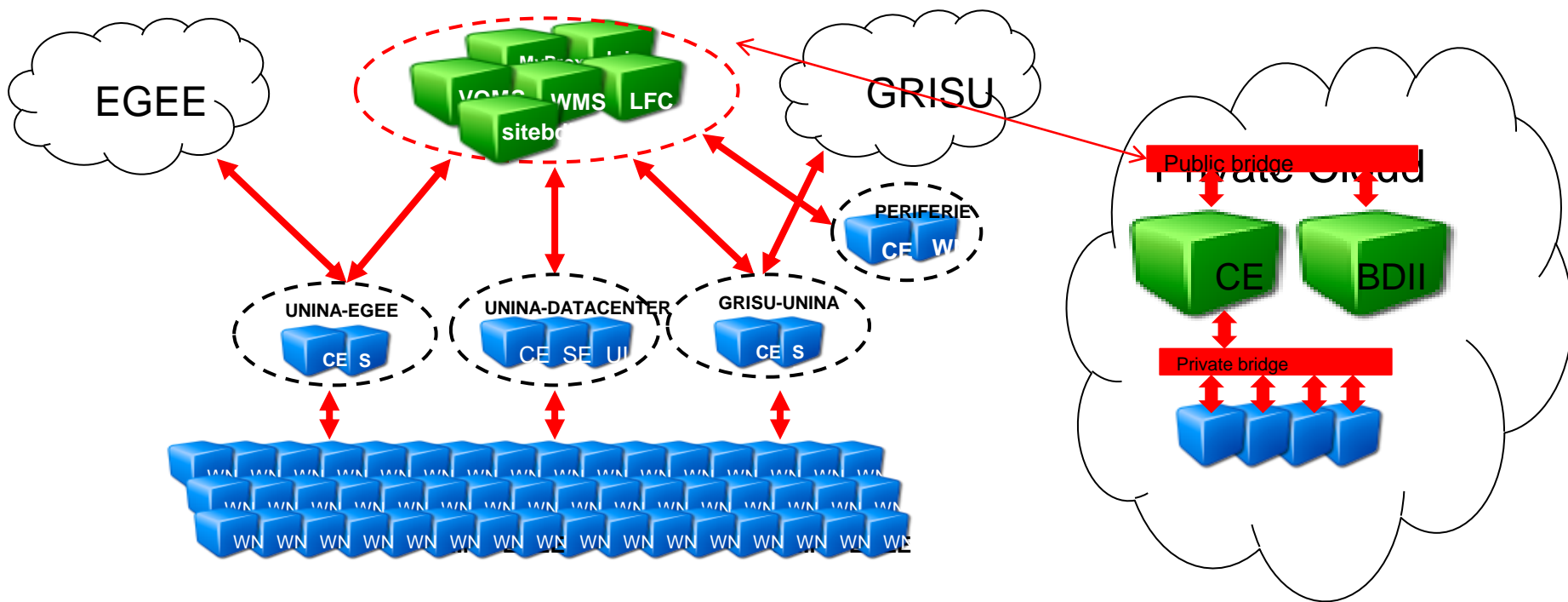
**→ vedi present. G. Ponti**



# Computing on demand in UNINA



Una soluzione per l'integrazione grid-cloud



Ogni nuovo sito, allocato on demand grazie a risorse cloud, rappresenta l'insieme di risorse di calcolo e storage tagliato sulla base delle esigenze espresse dalla comunità di utenti che ne fa richiesta



## Computing on demand in UNINA



### Grid as a Service (GaaS):

- Scelta della piattaforma di Cloud Computing
- Deploy di un sito grid su cloud privato
- Verifica delle prestazioni





## Computing on demand in UNINA



### OpenNebula:

- Piattaforma di Cloud Computing open source che implementa il paradigma di IAAS
- Sistema modulare e customizzabile
- Supporto delle più popolari piattaforme di virtualizzazione (KVM, XEN, VMWare)
- interfacce di gestione (CLI, API, Sunstone GUI, AWS, OCCl)



## Computing on demand in UNINA



### Deploy di un sito grid su OpenNebula:

- Template SL5 paravirtualizzato
- File di contestualizzazione per ruolo (OpenNebula context file)
- File di configurazione di sito (yaim site-def.conf file)

Tutti i ruoli del sito “on demand” sono definiti, installati e configurati a deploy-time

→ Stesso approccio di StratusLab, realizzato un prototipo



## UNINA: Conclusioni



- Quanto in precedenza descritto si pone oltre che l'obiettivo di consolidare le DCI esistenti anche quello di integrare il preesistente con i più moderni paradigmi quali quelli del cloud computing
- A valle della nostra esperienza:
  - abbiamo definito alcune best practice per la realizzazione di un'infrastruttura di calcolo basata sul paradigma del Cloud e sulla virtualizzazione (con particolare attenzione al tema dell'HPC)
  - ci siamo interrogati su quello che deve essere il giusto compromesso fra le esigenze di riduzione dei costi di gestione, le esigenze di flessibilità e la soddisfazione complessiva dell'utente
  - ci siamo convinti che l'adozione di tale tecnologie non può prescindere dalla definizione di opportuni SLA fra i provider e l'utente



## UNICAL: Our cloud infrastructure



- Desktop cloud (Java+VirtualBox virtu.)
- Private IaaS cloud
  - OpenNebula 3.0
  - AWS, OCCl cloud interfaces
  - Xen 4.0 as virtualization backend
  - Consolidation on top of multicore CPU
  - Haizea as advanced scheduling backend
  - Cloud federation (between DEIS dept. and Center of Excellence)
  - Operation research group (vehicle routing problems, transportation, etc.)

→ l'approccio più "teorico" tra quelli presentati al WS



# UNICAL: Resource Contention Metric for HPC Workloads

$$RC(t) = \sum_{\forall n \in \text{PhyNodes}} RC(t, n) \quad RC(t, n) = \sum_{\forall r \in \text{resTypes}} RC(t, n, r)$$

Let  $JobsSchedOn(n, t)$  be the scheduled jobs on physical host n at time t

$$JS(t, n) = JobsSchedOn(n, t)$$

Let  $j.stresson$  be resources which job j puts stress on

Let  $j.resReq$  be <sup>them</sup> resource requirements of job j in terms of capacity

$$resContFlg(t, n, r) = \begin{cases} 1, & \sum_{\forall j \in JS(t, n)} (r \in j.stresson == True ? 1 : 0) > 1 \\ 0, & \text{Otherwise} \end{cases}$$

$$RC(t, n, r) = \begin{cases} \sum_{\forall j \in JS(t, n) \wedge r \in j.stresson} j.resReq[r], & resContFlg(t, n, r) = 1 \\ 0, & \text{Otherwise} \end{cases}$$



## UNICAL: Live with Standards



Reference platform: interoperability and portability

Federation, hybrid, public and private clouds

- Compute Interfaces: OCCl from OGF
- Data Interfaces: CDMI from SNIA
- VM Image description: OVF from DMTF
- Contextualization: OVF from DMTF
- Management of Cross-site interoperation: such as networking
- Using OCCl for Federation

→ implementazione di un prototipo su piccola scala



## WNoDeS



- L'ultima presentazione riguardava l'esperienza fatta in INFN/IGI con WNoDeS, una soluzione cloud integrata in grid e la collaborazione con l'Università di Messina per realizzare la federazione di cloud (CLEVER)
- Questa soluzione è stata scelta come piattaforma di riferimento per IGI per la sua robustezza, il suo uso in produzione al Tier1 INFN (circa 2000 VM) e in alcuni altri siti e per la sua inclusione nel middleware rilasciato dal progetto European Middleware Initiative (EMI)
- Dettagli nelle prossime slide (courtesy D. Salomoni)



## WNoDeS

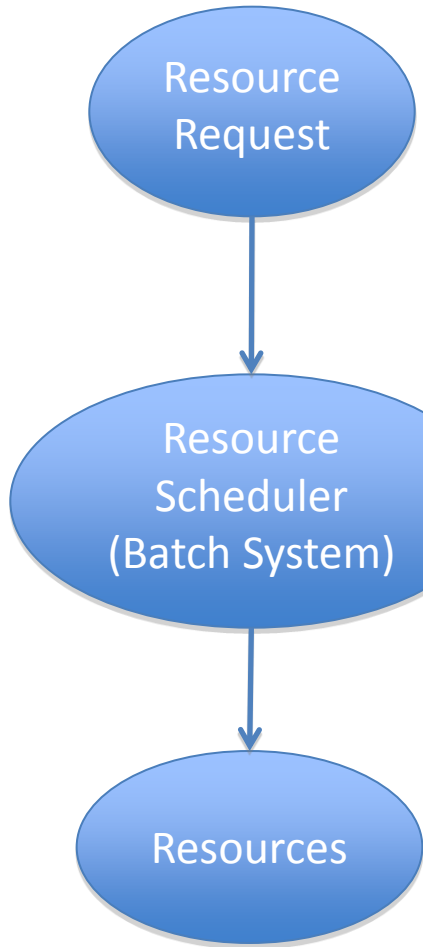


- WNoDeS → **Worker Nodes on Demand Service**
- A software framework created by INFN to **integrate Grid and Cloud provisioning through virtualization**
  - **Key feature**: all resources (presented via Grid, Cloud, or else) are taken from a **common pool** to avoid static partitioning.
  - **Key feature**: resource matchmaking policies are handled by an LRMS.
- **Scalable and reliable** – it is **in production** at several Italian centers, including the INFN Tier-1 since November 2009
  - Currently managing about 2,000 on-demand Virtual Machines (VMs) there.
- **Transparent** for local and Grid Computing Center services
- Leveraging **proven open source software** technologies like Linux KVM, Torque/Maui (Platform LSF also supported), EMI gLite middleware
- **WNoDeS version 2 is part of EMI-2**, to be released on May 18, 2012





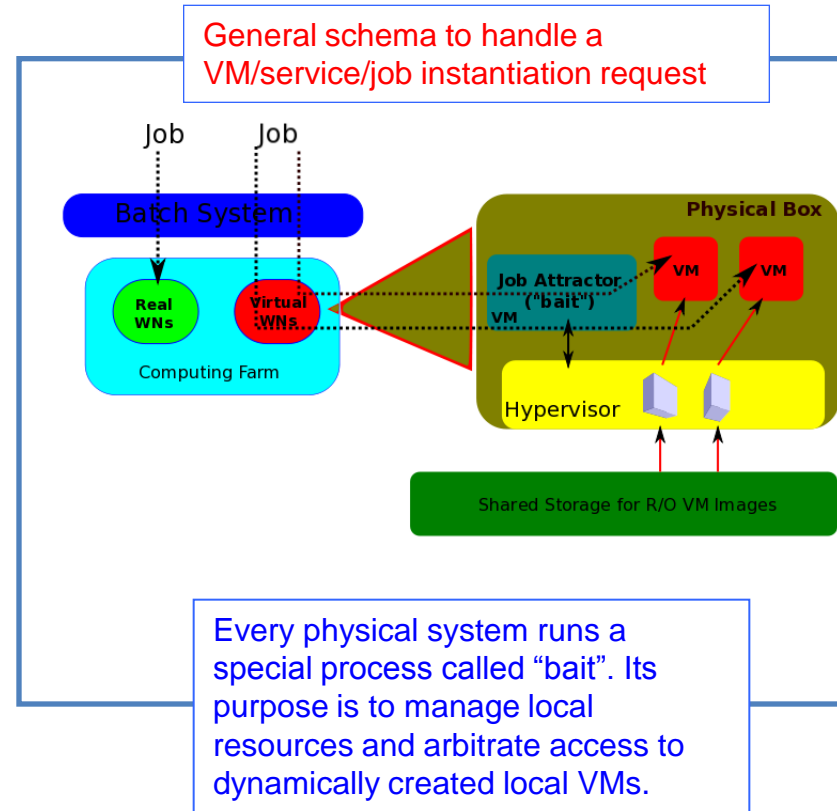
# WNoDeS, Architectural Overview



It can be a **local job**, a **Grid job**, a **VM instantiation request**, a **request for a Cloud service**. All these requests get transparently translated into "jobs"...

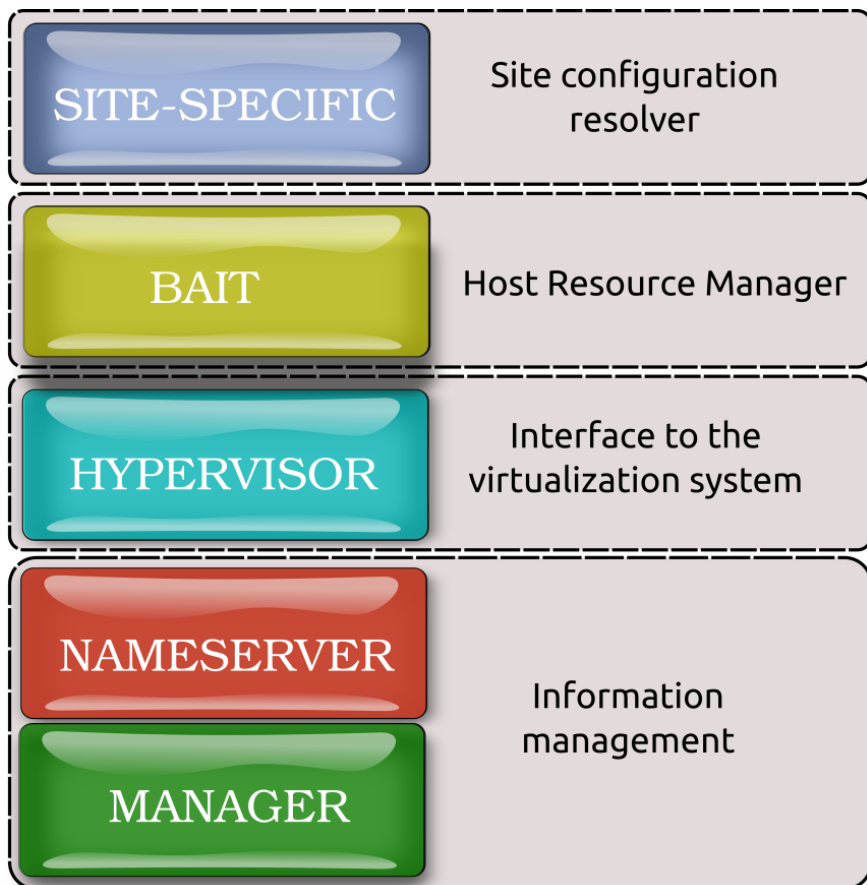
... that are handled **by an LRMS** (a batch system). LRMS-based policies allow flexible scheduling and scalable access to **resources**...

... that can be based on a **mix of systems**: some capable of KVM-based virtualization, some traditional, non-virtual resources.





## What's in WNoDeS 2



- WNoDeS 2 will be released this month **as part of the EMI-2 stack**
- **Batch job execution**
  - Local batch jobs can be run on both real and virtual execution hosts (worker nodes)
  - A virtual worker node is instantiated on-demand to provide computing resources as they are needed
- **Grid integration**
  - Dynamic selection of resources through standard gLite submission tools
  - Resource selection based on a Glue 1.x attribute
  - Authentication based on VOMS
- **Support for LSF and PBS**
- **New feature: mixed mode**

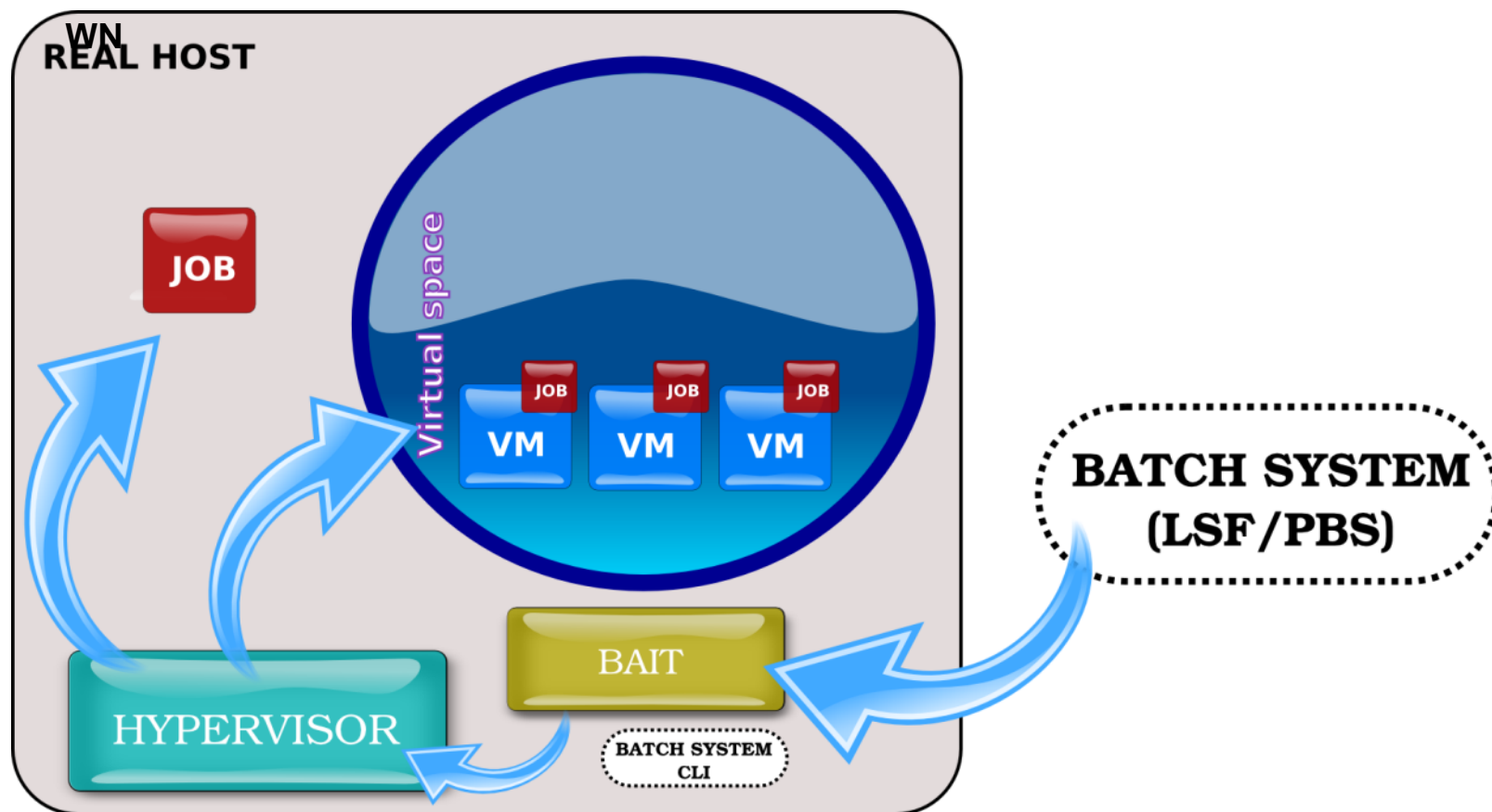


## WNoDeS Mixed Mode



- **What**
  - A WNoDeS configuration option allowing the use of physical resources as both traditional batch nodes **and** as hypervisors for the instantiation of virtual machines – on the same hardware, at the same time.
  - VMs can be used to also run batch jobs, or to provide cloud services.
- **Why**
  - Some tasks are not suitable to being executed on virtual nodes – for example, jobs requiring GPGPU resources, or jobs with high I/O requirements → run them on physical nodes.
  - On the same physical nodes, one can also offer virtualized services for those users requiring them → no need to set aside nodes for virtualized services.
- **Where**
  - Mixed mode is included in the WNoDeS version released with EMI-2 and can be administratively turned on or off.

# WNoDeS, Mixed Mode on





## Mixed Mode Pros and Cons



- **Pros**

- Progressively install WNoDeS in a farm without first having to decide which nodes will support virtualization and which not.
- Add support e.g. for Cloud computing, interactive usage on custom VMs etc. in a traditional farm.
- Direct jobs to VMs or to real hardware using LRMS policies and a simple pre-exec/prologue script (a template is supplied with the WNoDeS distribution). One can differentiate real vs. virtual requests/jobs e.g. based on queues, users, requirements, Grid VOs, etc.

- **Cons**

- In a purely virtual farm set up, physical systems are only used as hypervisors, so they can be put e.g. in private address space. With mixed mode, they can also be used (like in a traditional farm) to run jobs and may need public access.
- With mixed mode, a physical system is part of the LRMS cluster and may use up LRMS licenses proportionally to the number of its cores. If the same physical system is then also used to create VMs that become part of the LRMS cluster (e.g. to run batch or grid jobs), these VMs will also use up LRMS licenses and the total number of LRMS licenses used by a physical system may be  $O(2 \cdot \text{cores})$ . This can be a problem with some sites using commercial LRMS.



## Some Upcoming Features (1)



- Several of the following features have been part of WNoDeS 1 (pre-EMI) for a long time. They will be released as EMI updates in the coming months.
- **Interactive usage**
  - We use this at the INFN Tier-1 to support self-instantiation of VMs by local users. These VMs can be used for e.g. analysis tasks, testing purposes, etc.
    - As with other WNoDeS services, resources can be taken from the general purpose Tier-1 farm (no service partitioning, unless one specifically configures it.)
- Cloud computing: **OCCI compliance**
  - Will be compliant to OCCI 1.1 (validated e.g. through doyou speakocci) and accessible via a CLI.
- Cloud computing: **Web interface**



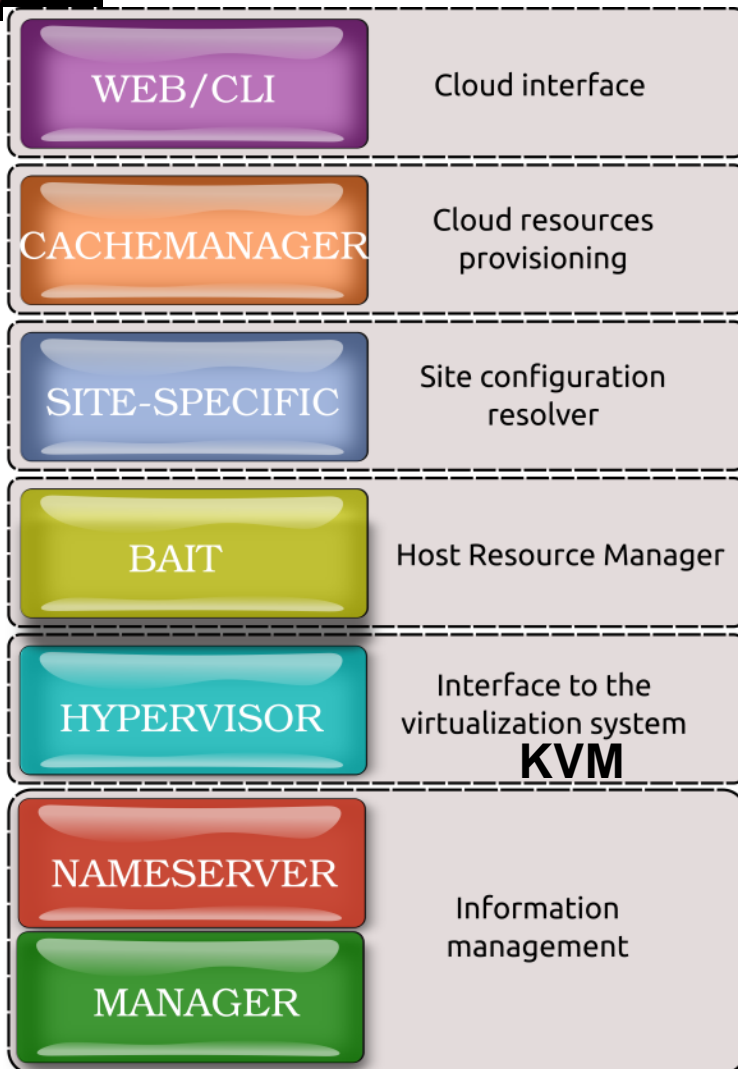
## Some Upcoming Features (2)



- **Dynamic virtual networks** (→ vedi talk di M. Caberletti)
  - Won't require using 802.1q to partition networks and will allow dynamic instantiation of private VLANs (either local or across multiple sites) and address assignment for VM isolation – a much needed feature in cloud environments.
- **Integration of multiple authentication methods**
  - The current WNoDeS cloud Web application and OCCl interface use X.509 + VOMS; this will be extended to support federated access.
- **Storage volume support**
  - Will allow dynamic instantiation and connection to VMs of persistent storage volumes.



# WNoDeS, Cloud Support



## EMI2 Updates

## EMI2





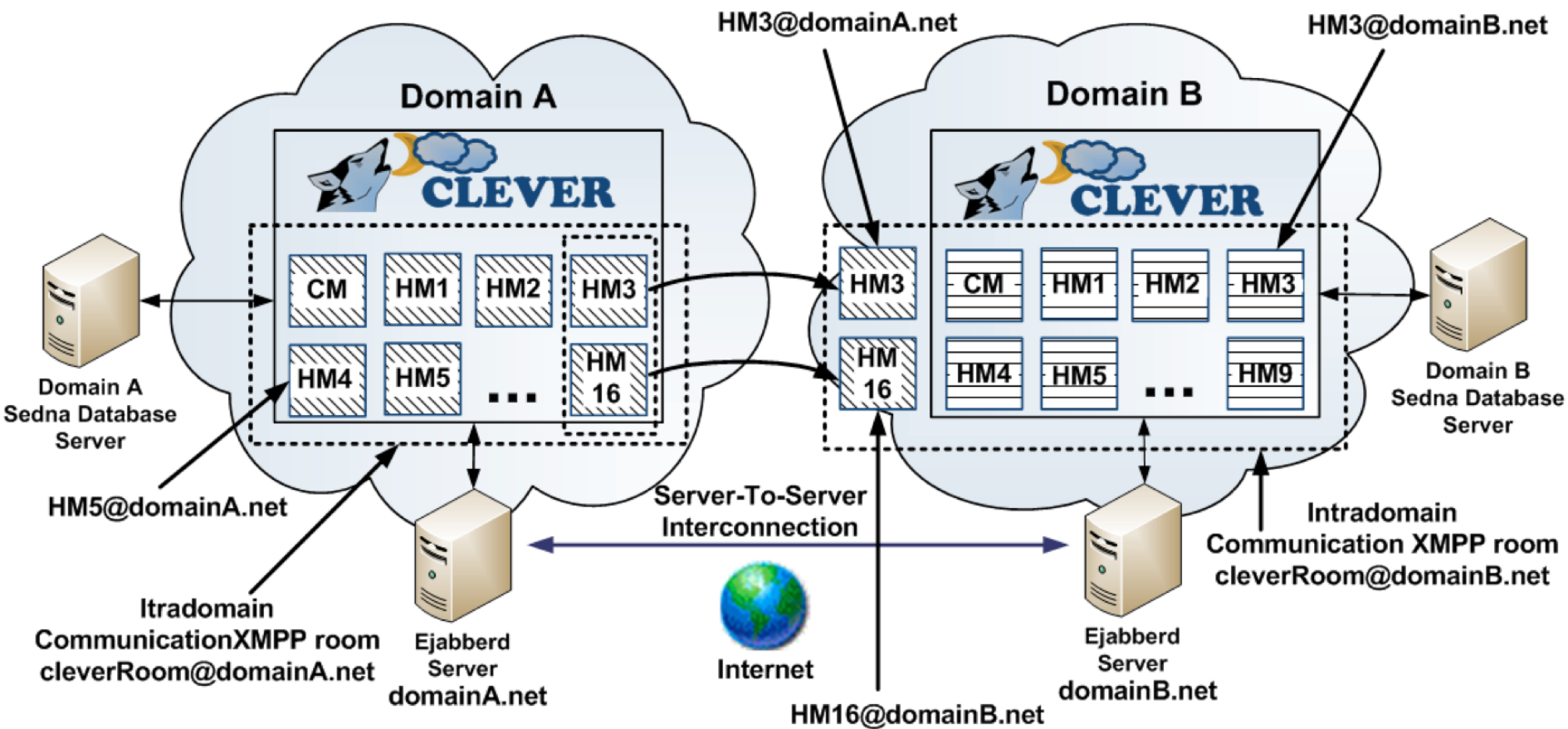
## A Cloud-Enabled Virtual EnviRonment



- A lightweight Cloud environment, jointly developed by University of Messina (A. Puliafito, M. Villari et al.) and INFN Catania (G. Andronico et al.)
- Main goals:
  - Simplify access management of private/hybrid clouds
  - Provide interfaces to interact with interconnected clouds, deploy VMs and perform load balancing through VM brokerage



# CLEVER Architecture





## WNoDeS / CLEVER Integration



- WNoDeS provides a proven, scalable foundation for VM instantiation and Grid/Cloud integration
- CLEVER provides scalable and resilient mechanisms to implement inter-cloud connectivity
- The technical solution for the integration of WNoDeS and CLEVER is under definition
- A collaboration has recently started



- More info:
  - <http://web.infn.it/wnodes>
  - [wnodes-support@lists.infn.it](mailto:wnodes-support@lists.infn.it)
  - The WNoDeS software will be available and supported through the usual EMI channels after EMI-2 is released.



## Conclusioni (I)



- Il paradigma cloud riscuote molto successo e moltissimi soggetti stanno sperimentando soluzioni per scopi più o meno ambiziosi, **ma servono soluzioni open e scalabili**
- Nel mondo della ricerca ci sono varie attività e iniziative che aspirano a sviluppare una soluzioni aperte, robuste e semplici
- Non si sta (ancora) delineando una soluzione 'vincente', e forse non ci sarà



## Conclusioni (II)



- Gli aspetti più importanti sono:
  - **soluzioni integrate** con i servizi esistenti (grid, batch system, etc.)
  - cloud storage: si stanno definendo i servizi necessari ma serve una soluzione che **integri** questi servizi e **sia appetibile sia per i service provider che per gli utenti**
  - necessaria la **semplificazione dell'accesso** alle risorse (ad es. autenticazione)
  - ultimo ma non meno importante, serve una soluzione per la **federazione delle cloud**



Grazie per l'attenzione

Domande?