

XtreemOS



*Enabling Linux
for the Grid*

Clouds: A New Playground for XtreemOS Grid Operating System

**Christine Morin, INRIA
XtreemOS scientific coordinator**

CCGSC 2008 - Flat Rock, NC, September 16, 2008



Information Society

*XtreemOS IP project
is funded by the European Commission under contract IST-FP6-033576*





- **Cloud & Grid computing**
- **Overview of XtreemOS Grid OS**
- **XtreemOS in Cloud computing scenarios**
- **Conclusion**



- **A paradigm for the on-demand provision of virtualized resources as a service**
 - Commercial clouds
 - Amazon (EC2, S3) & others
 - Private clouds
 - Intra-institution data centers





- **Virtualized services**
- **Service-oriented**
- **Deployment**
- **Scalability / Elasticity**
- **Security**
- **High availability / Self-management**
- **SLA-driven**



Reminder: What is a Grid?



"A fully distributed, **dynamically reconfigurable**, **scalable** and **autonomous** infrastructure to provide **location independent**, pervasive, **reliable**, **secure** and **efficient** access to a coordinated set of services encapsulating and **virtualizing resources** (computing power, storage, instruments, data, etc.) in order to generate knowledge"



- **Grid technologies highly relevant in the Cloud computing era**
- **Even more, some Grid platforms provide a cloud computing-like service (without claiming it)**
 - Example: **Aladdin-G5K**



The Aladdin-G5K Example

- Grid '5000 experimental Grid platform in France (since 2005)
 - Clusters in 9 sites, 5000 cores
- Users reserve a set of resources through a resource management system (OAR) and they can deploy their own software stack on the allocated machines (Kdeploy)
 - Remote reboots of real machines
 - No use of virtual machines but cloud-like service offered to the community of computer scientists developing Grid software

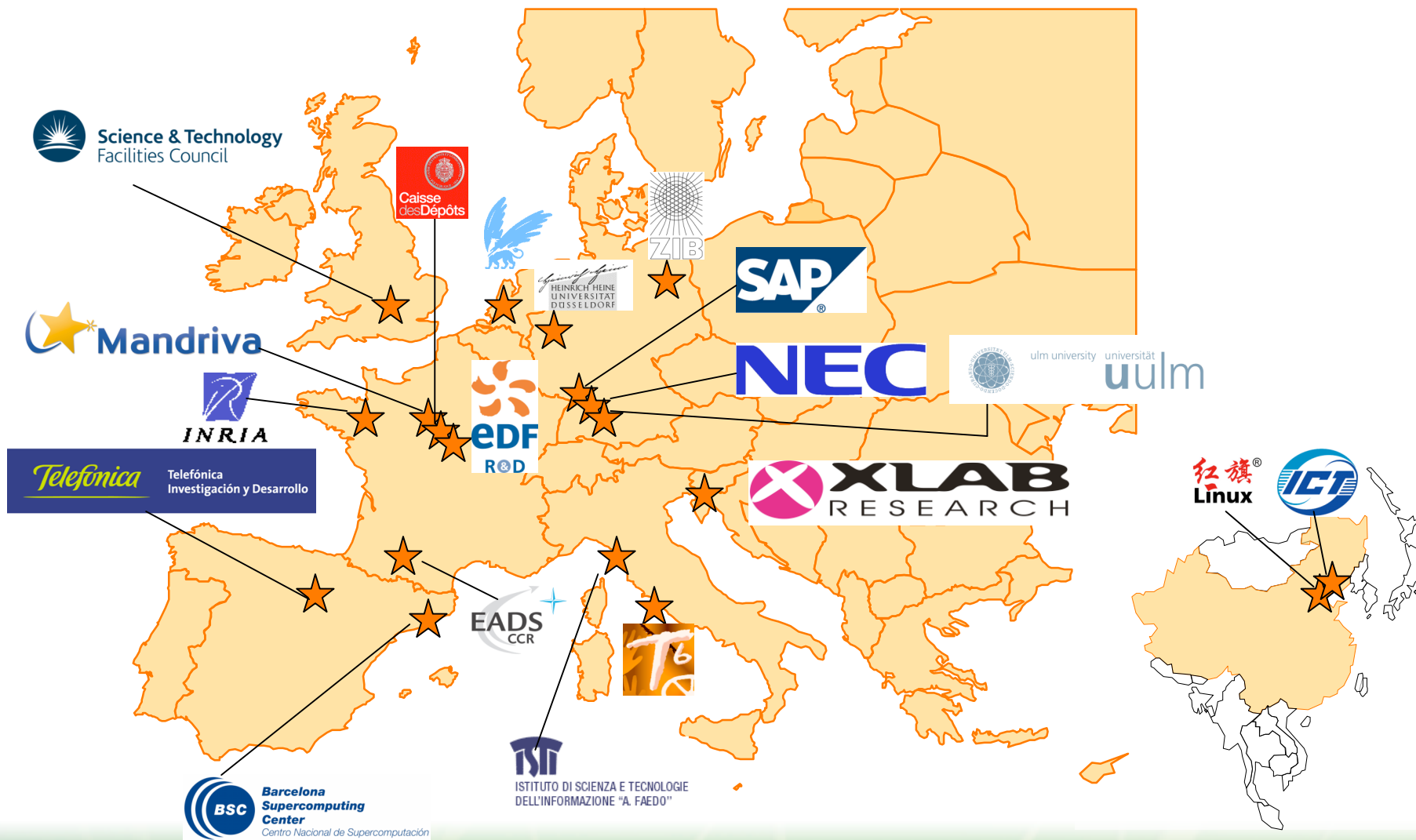




- **Cloud & Grid computing**
- **Overview of XtreemOS Grid OS**
- **XtreemOS in Cloud computing scenarios**
- **Conclusion**



- **4-year R&D project started in June 2006 in the FP6 framework**
- **30 M€ budget, 14.2 M€ EC grant**
- **19 academic & industrial partners from Europe & China**





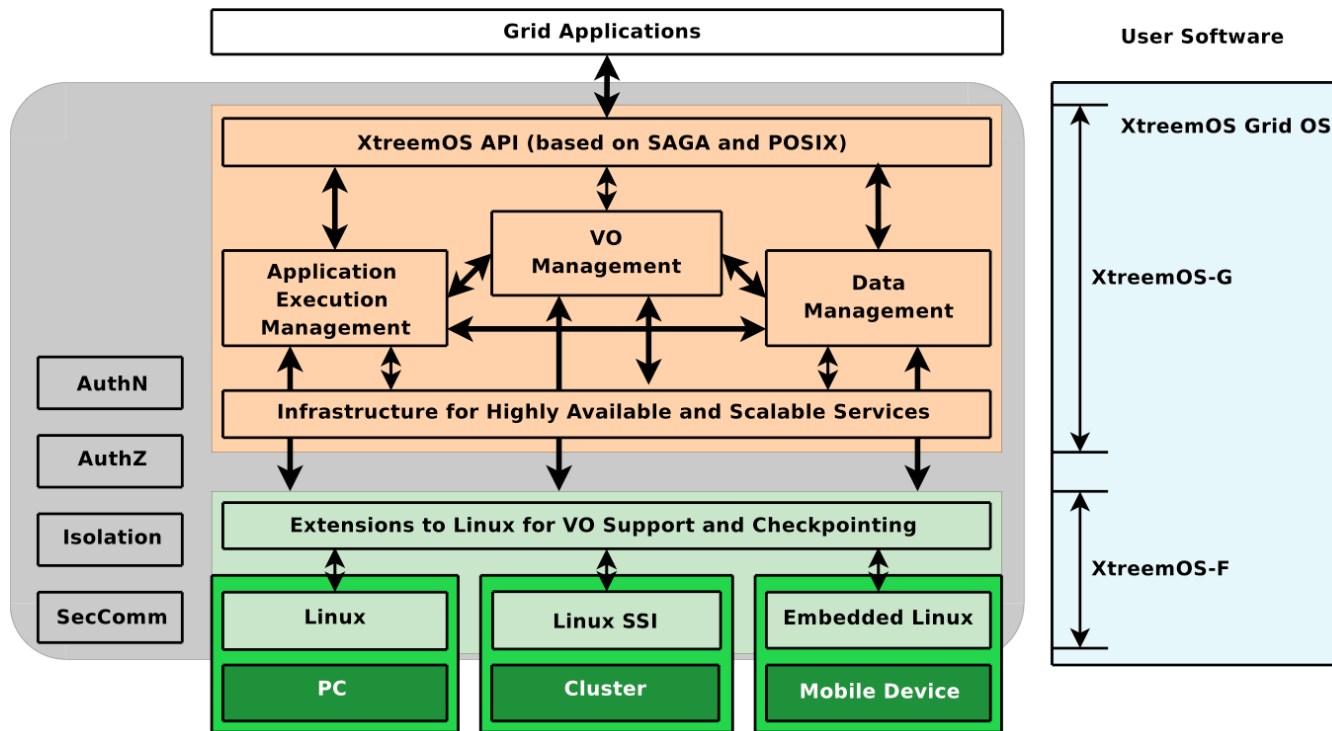
- **Design & implementation of an open source Linux-based Grid Operating System with native VO support**
- **Grid Operating System**

A **comprehensive** set of **cooperating** system **services** providing a **stable** interface for a **large-scale wide-area dynamic distributed infrastructure**

 - Abstraction
 - Sharing of heterogeneous resources in multiple administrative domains
- **Two fundamental properties: transparency & scalability**
 - Bring the Grid to standard users
 - Scale with the number of entities and adapt to evolving system composition



XtreemOS Architecture Overview





- **Scalability of management of dynamic VOs**
 - **VO-centric security architecture**
 - Dynamic mapping between Grid VO users & Linux entities with no modification to Linux kernel
 - No centralized Grid wide data base
 - No grid map file needed
 - User management does not necessitate any resource reconfiguration
- **Flexible administration of VOs**
 - Multiple VO models supported
 - Hierarchical policy management (VO, resource, user)
 - Accountability of data access and service execution
- **Interoperability with third party security infrastructures**
 - Kerberos, LDAP, Shibboleth...
- **Single-Sign-On**





■ Objectives

- Start, monitor, control applications
- Discover, select, allocate resources to applications

■ Features

- No assumption on local node RMS
 - AEM can be used without any batch system
- Job “self-scheduling”
 - No global job scheduler
- Resource discovery based on overlay networks
 - Multi-criteria and range of values queries
- Unix-like job control
- Accurate and flexible monitoring of job execution
- Checkpointing service for grid jobs





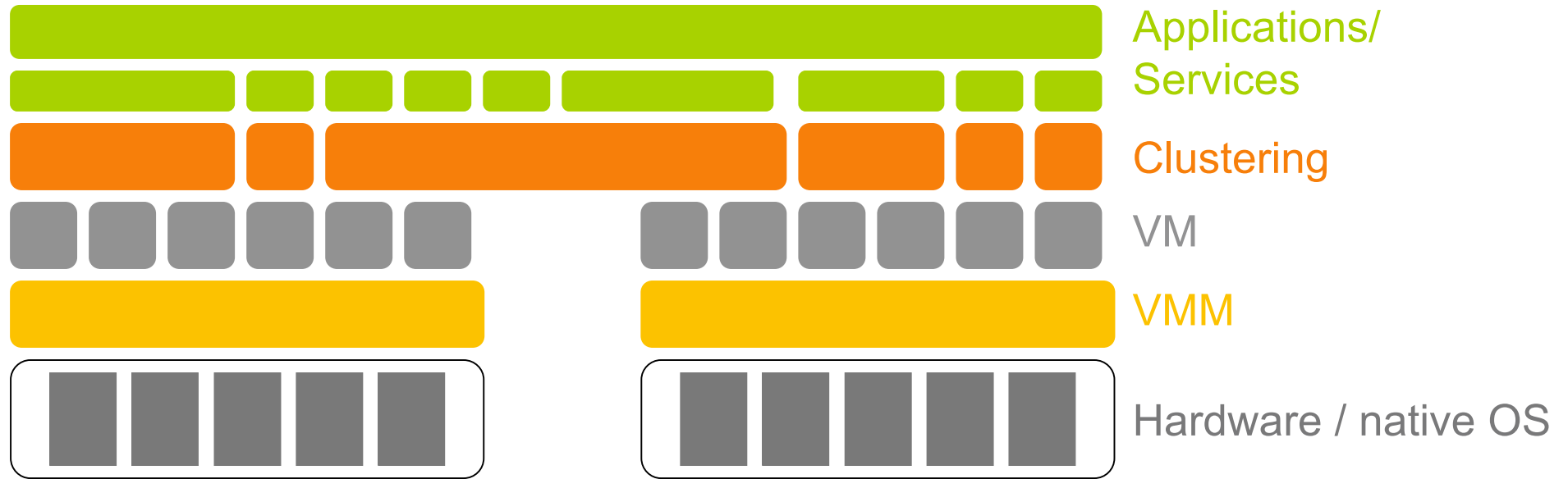
- A Grid file system providing users with a global view of their files
- Posix interface
- Efficient location-independent access to data in a Grid
 - Grid users from multiple VO
 - Data storage in different administrative domains
- Autonomous data management with self-organized replication and distribution
- Consistent data sharing



- **Cloud & Grid computing**
- **Overview of XtreemOS Grid OS**
- **XtreemOS in Cloud computing scenarios**
- **Conclusion**



Cloud Computing Environment



Administrative domain 1

Administrative domain 2 ..n





- Clouds spanning multiple hardware suppliers
- Crossing cloud boundaries
- Supporting cooperation between different institutions using private, commercial clouds & traditional IT infrastructure
- Virtual clusters
- Clouds & mobile devices

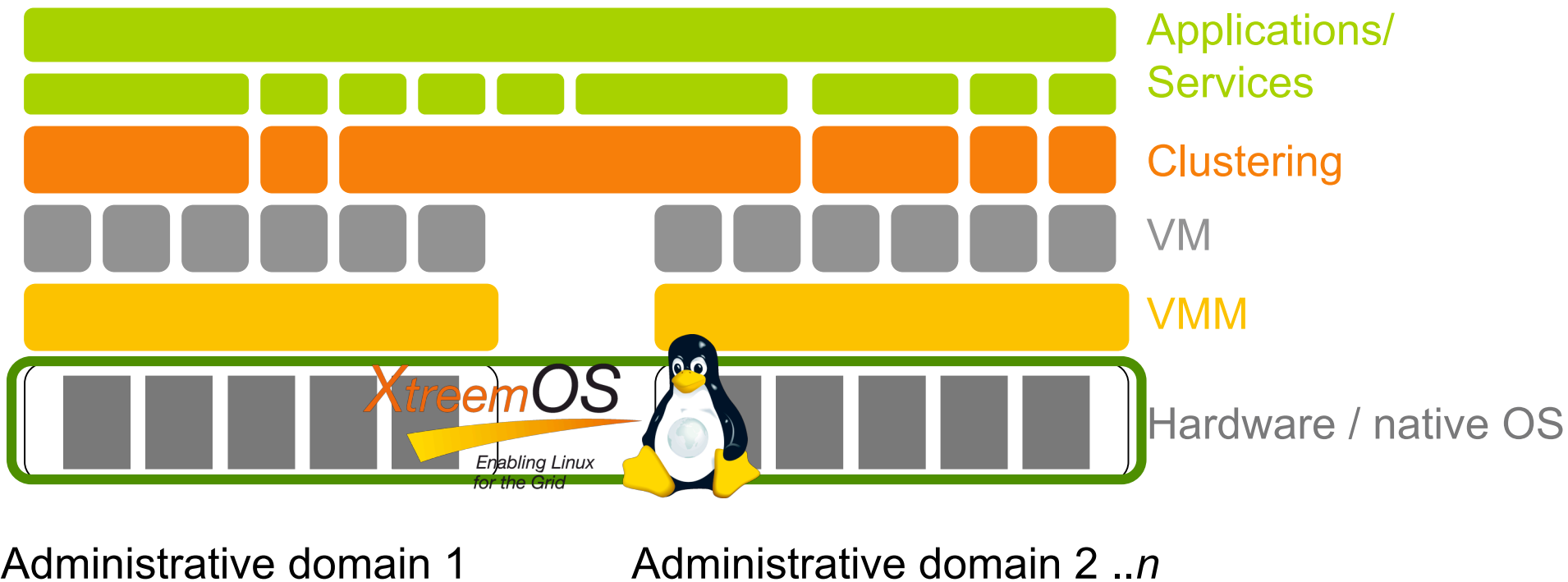


Clouds Spanning Multiple Hardware Suppliers





Clouds Spanning Multiple Hardware Suppliers





- **XtreemOS scenario**

- Hardware suppliers

- Each resource node runs XtreemOS
 - Define local policies for their resource usage

- Cloud service suppliers

- One VO per Cloud supplier where resources from the different hardware suppliers are registered
 - The cloud service supplier is the VO manager & defines its VO level policies

- Cloud customers

- Registered as VO users in a Cloud supplier's VO



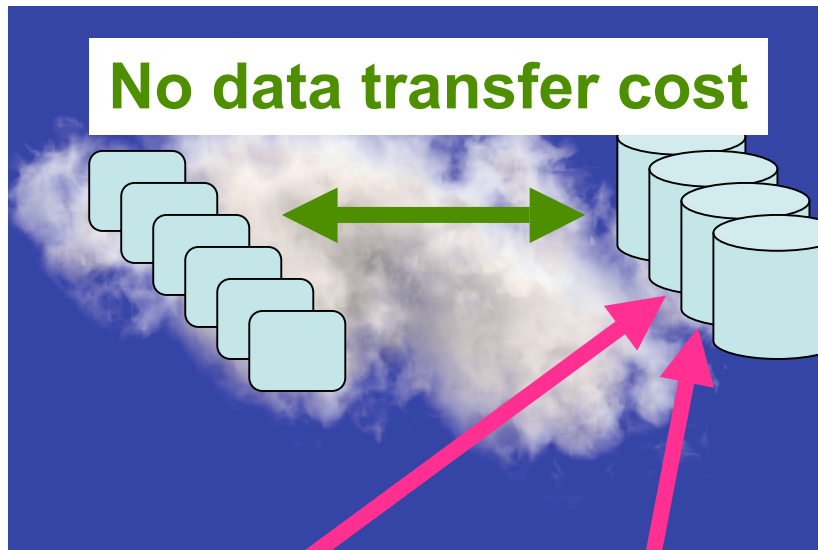
■ XtreemOS Key Features

- VO management
 - Multiple VO
 - Hardware supplier can sell their computing power to different cloud suppliers
 - Dynamic VO
 - VO policies
 - A cloud supplier can define the way it uses the various resources from different hardware suppliers
 - Hardware suppliers keep control on their resources
- Scheduling
 - Scheduling of VMs
 - Scheduling policies well-suited to clouds need to be defined
- XtreemFS Grid file system
 - Management of VM images



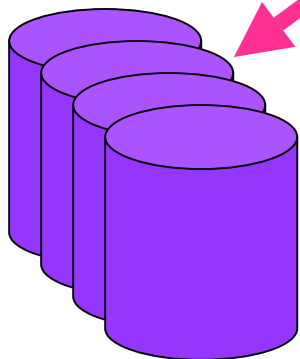


Crossing Cloud Boundaries



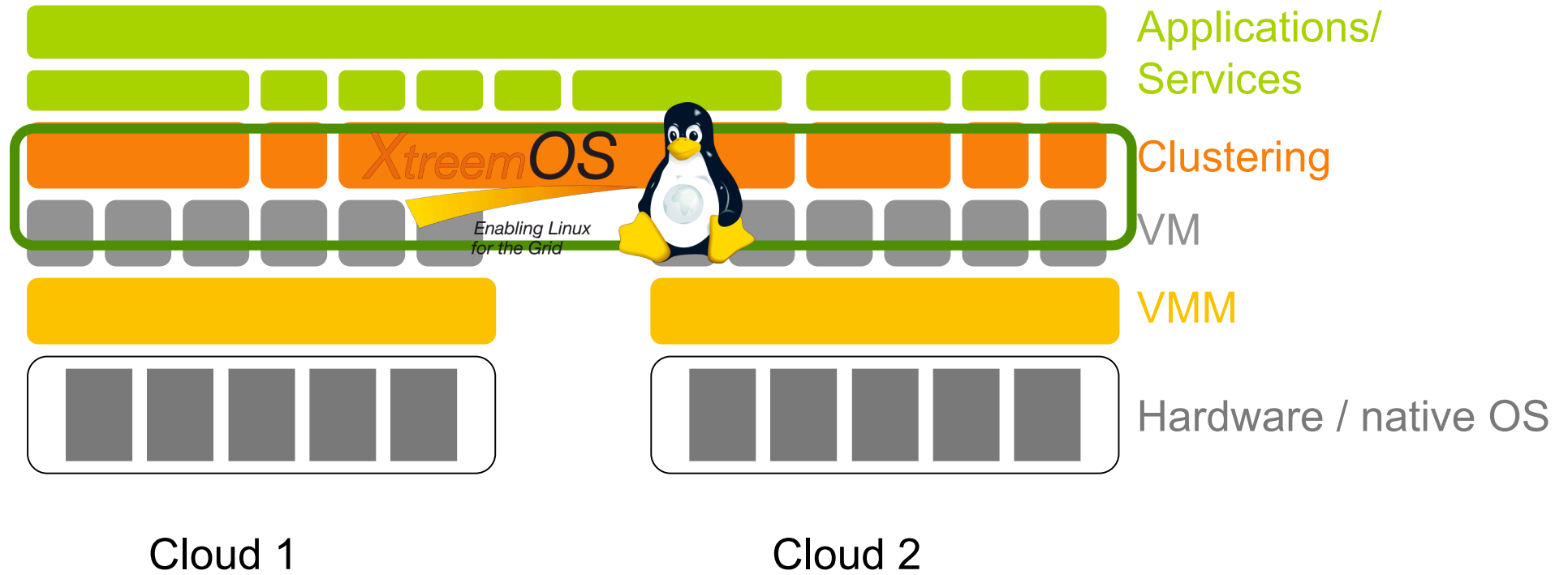
An institution has its own IT infrastructure and uses commercial cloud resources for resource on-demand (peak of activity)

Data transfers costs





Crossing Cloud Boundaries





■ XtreemOS scenario

- XtreemOS runs in Cloud VMs
- XtreemOS runs on the institution own computers
- A VO is managed by the institution and contains its own resources and members as well as Cloud VMs

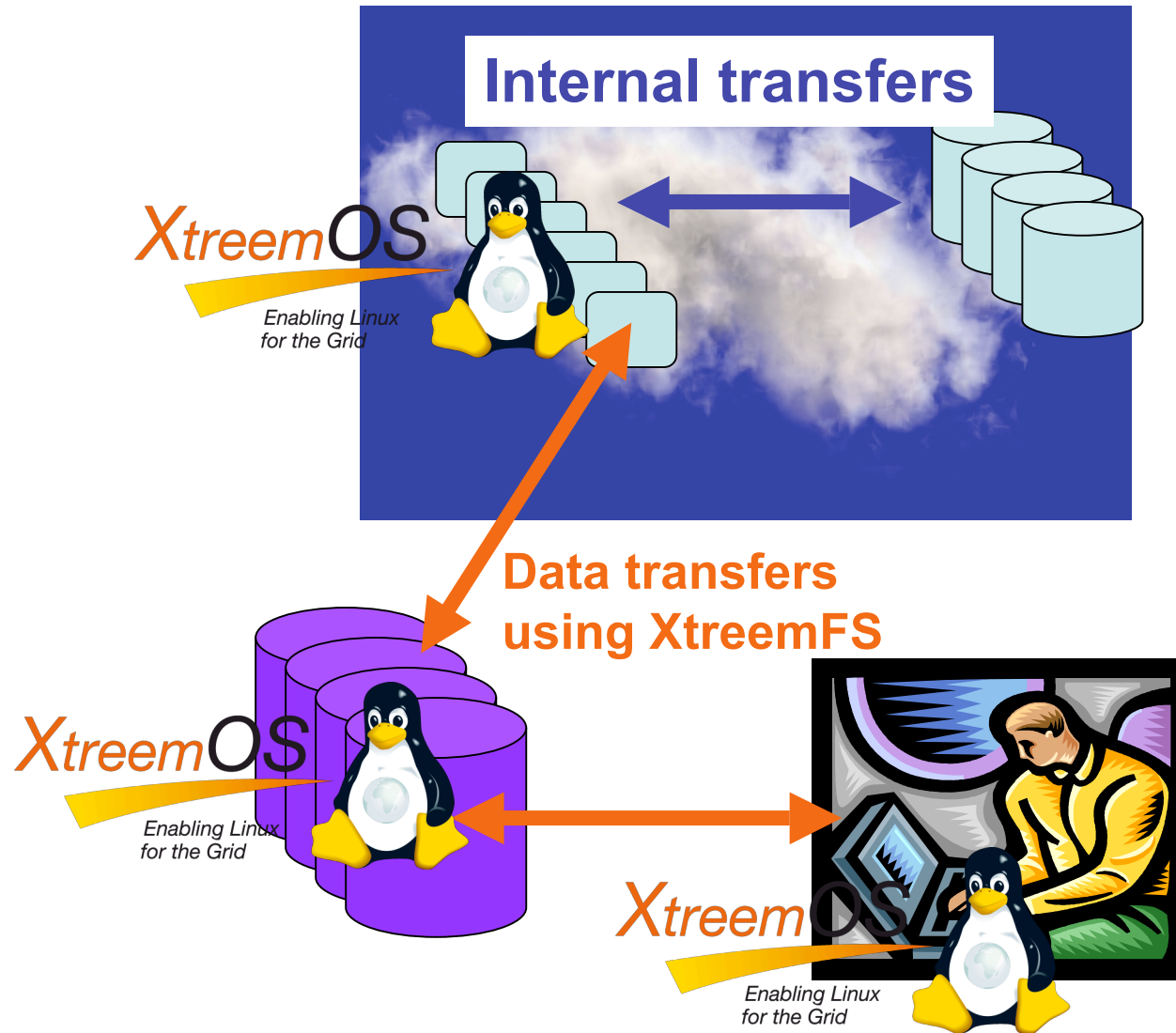
■ Key XtreemOS features

- XtreemFS Grid file system
- Dynamic mapping of VO users on on-fly created user (Linux) accounts
- Scalable dynamic VO management
 - Cloud VMs dynamically added to the VO





Crossing Cloud Boundaries





- **This scenario extends the previous one**
 - Considering multiple commercial clouds used by the different institutions
 - Considering multiple VOs
 - One VO per institution
 - One VO per established collaboration (collaborating institutions belonging to the collaboration VO)

- **Virtual Grid Concept**

XtreemOS

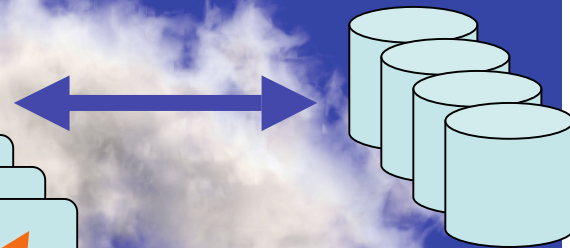
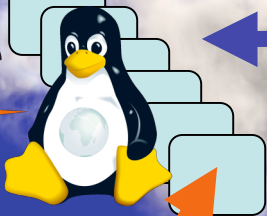
Enabling Linux for the Grid



Supporting cooperation between different institutions using private, commercial clouds & traditional IT infrastructure

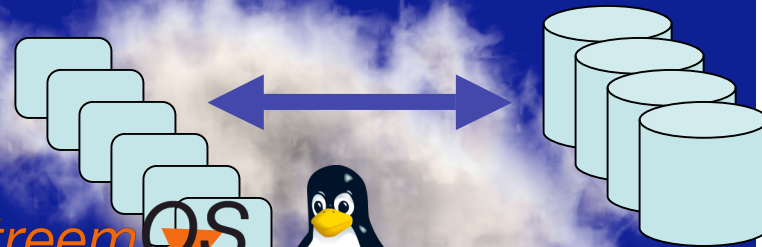
XtreemOS

Enabling Linux for the Grid



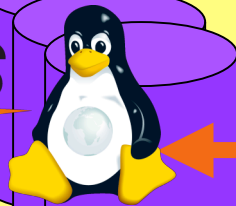
XtreemOS

Enabling Linux for the Grid



XtreemOS

Enabling Linux for the Grid



XtreemOS

Enabling Linux for the Grid

XtreemOS

Enabling Linux for the Grid



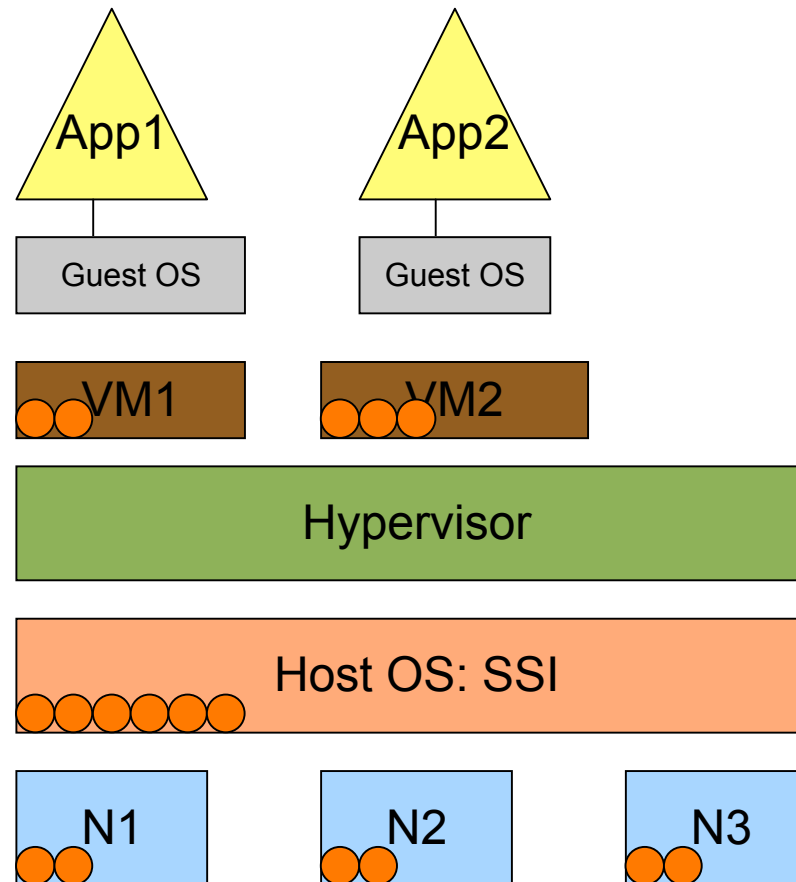


- **What do we gain by combining Linux SSI technology with VMs?**
 - VHPC'08 joint paper with G. Vallée & S. Scott (ORNL)
- **Linux SSI**
 - Illusion of a powerful PC
 - Aggregation of cluster nodes resources (eg.memories)
- **Virtualization Technologies**
 - Portability
 - Flexible hardware resource management
 - Suspend/restart, migration
 - Isolation
- **Scenarios**
 - VMs on top of a LinuxSSI cluster
 - LinuxSSI on virtual clusters



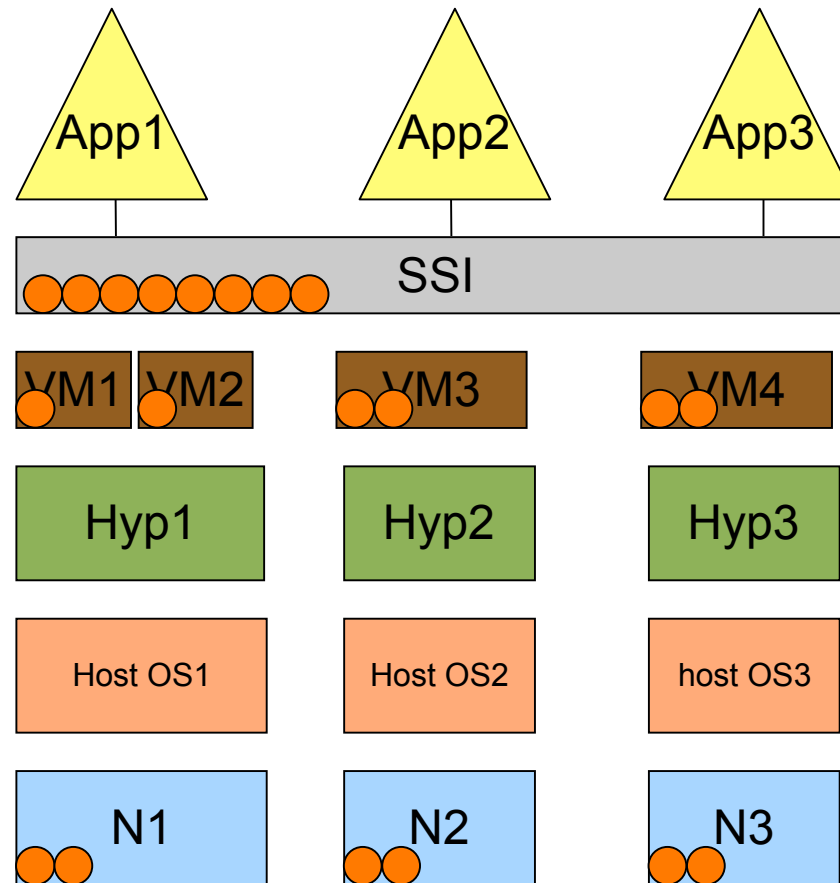


A VM can get more resources than offered by a single physical node (for instance, memory)





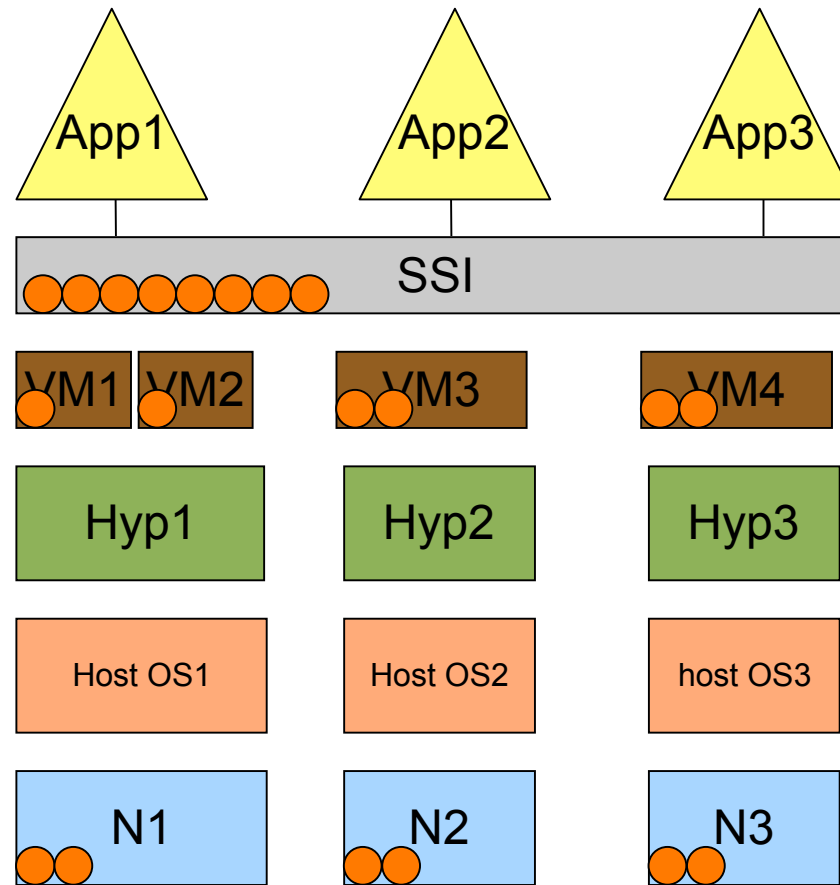
SSI upon Type II-Virtualization





Flexibility in hardware resource management

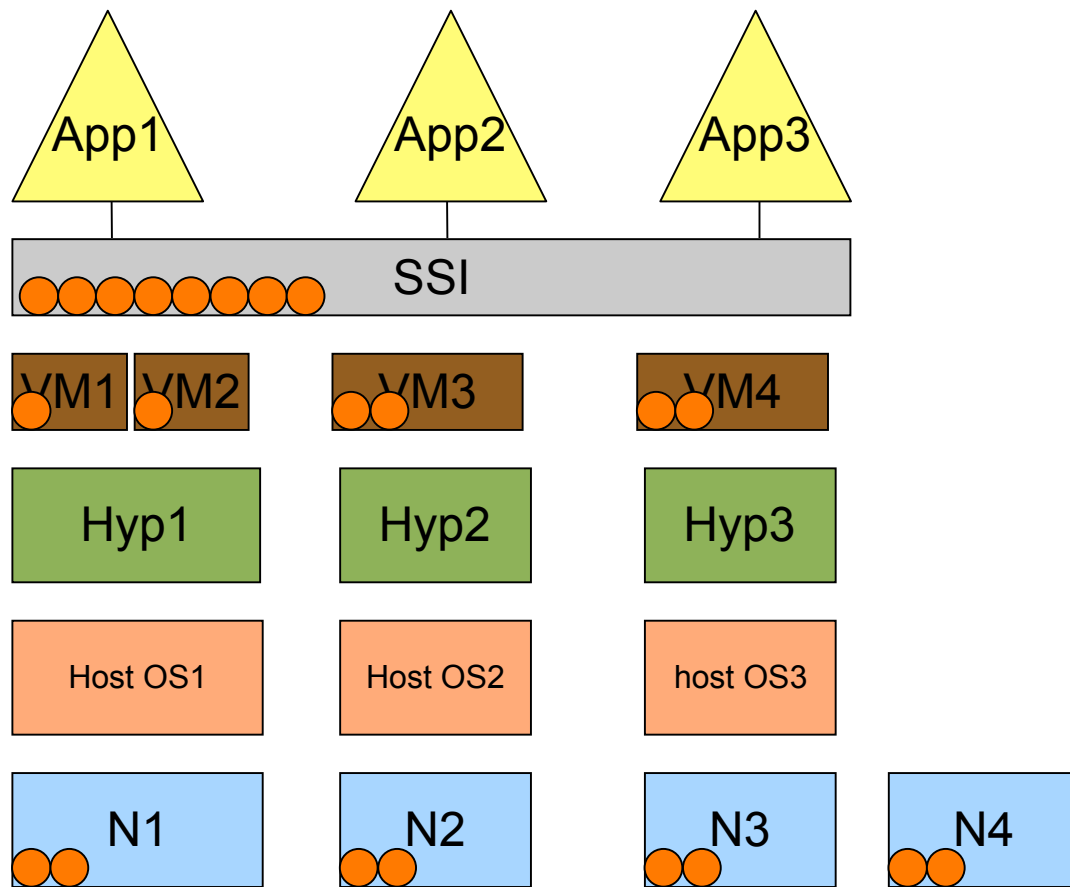
A solution to simply manage node addition/eviction in a SSI cluster





Flexibility in hardware resource management

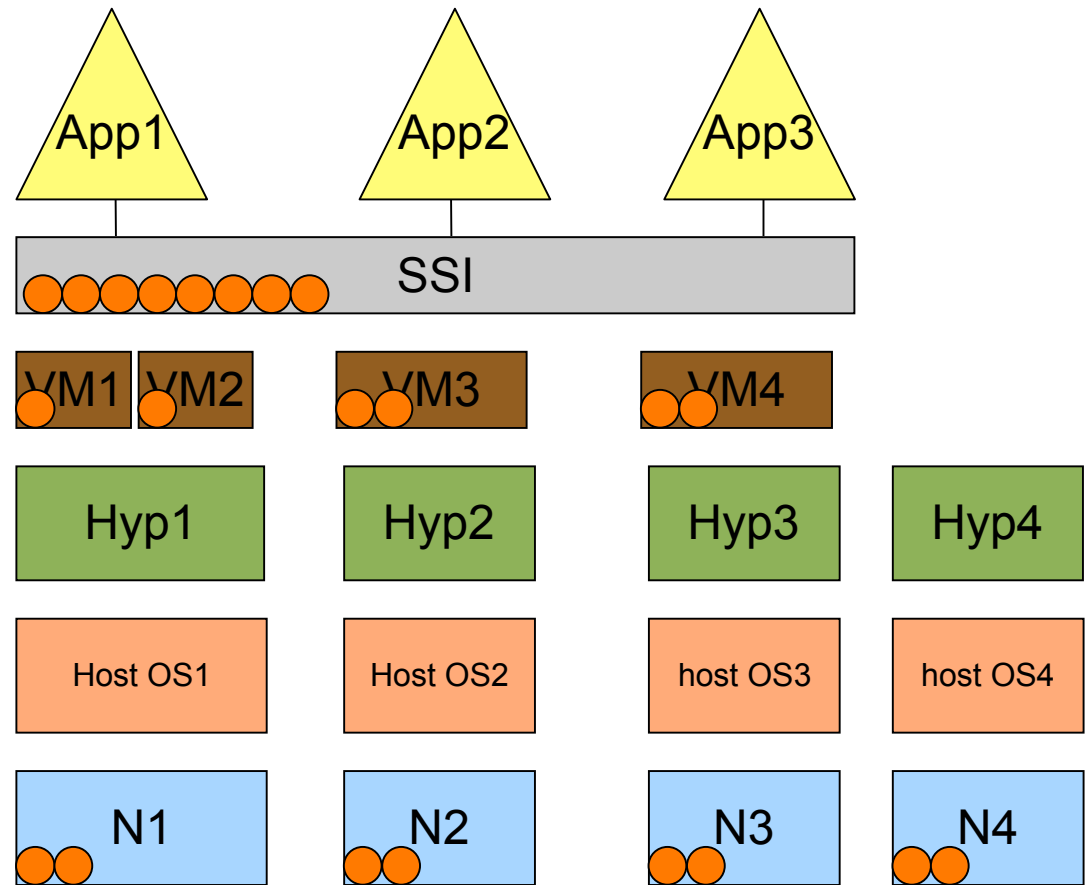
A solution to simply manage node addition/eviction in a SSI cluster





Flexibility in hardware resource management

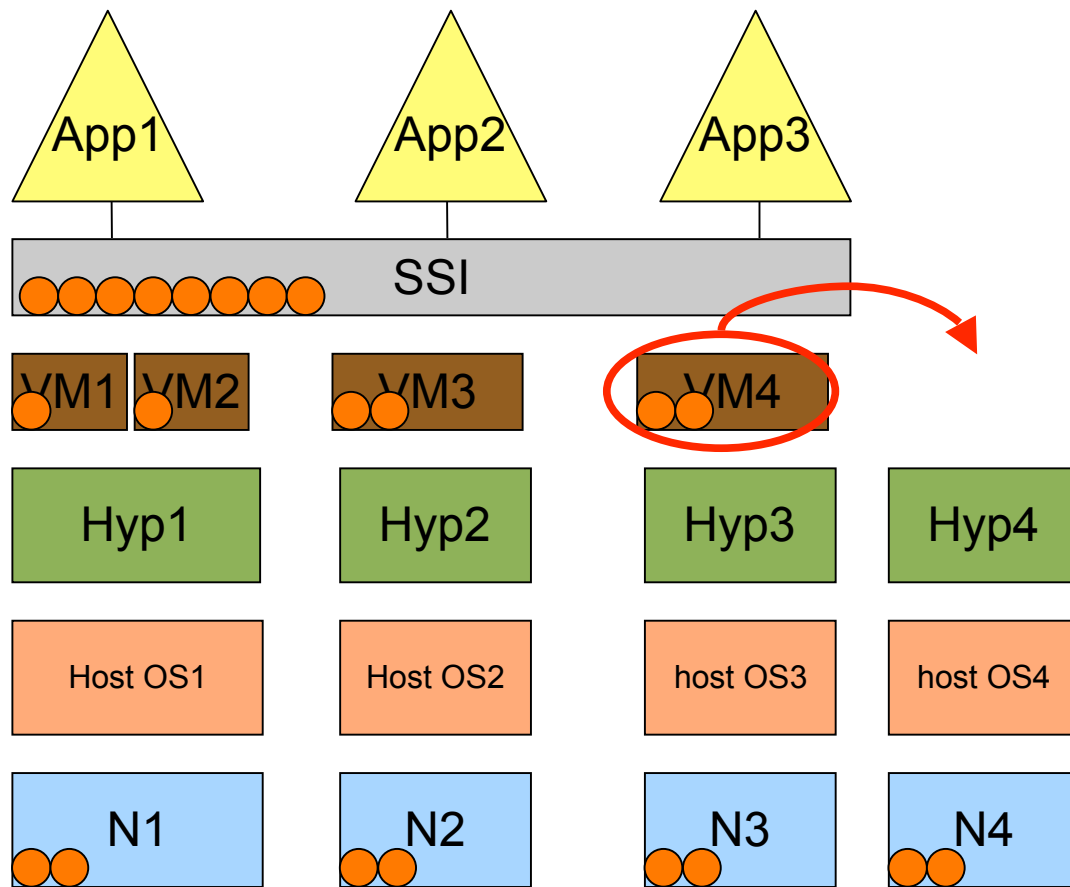
A solution to simply manage node addition/eviction in a SSI cluster





Flexibility in hardware resource management

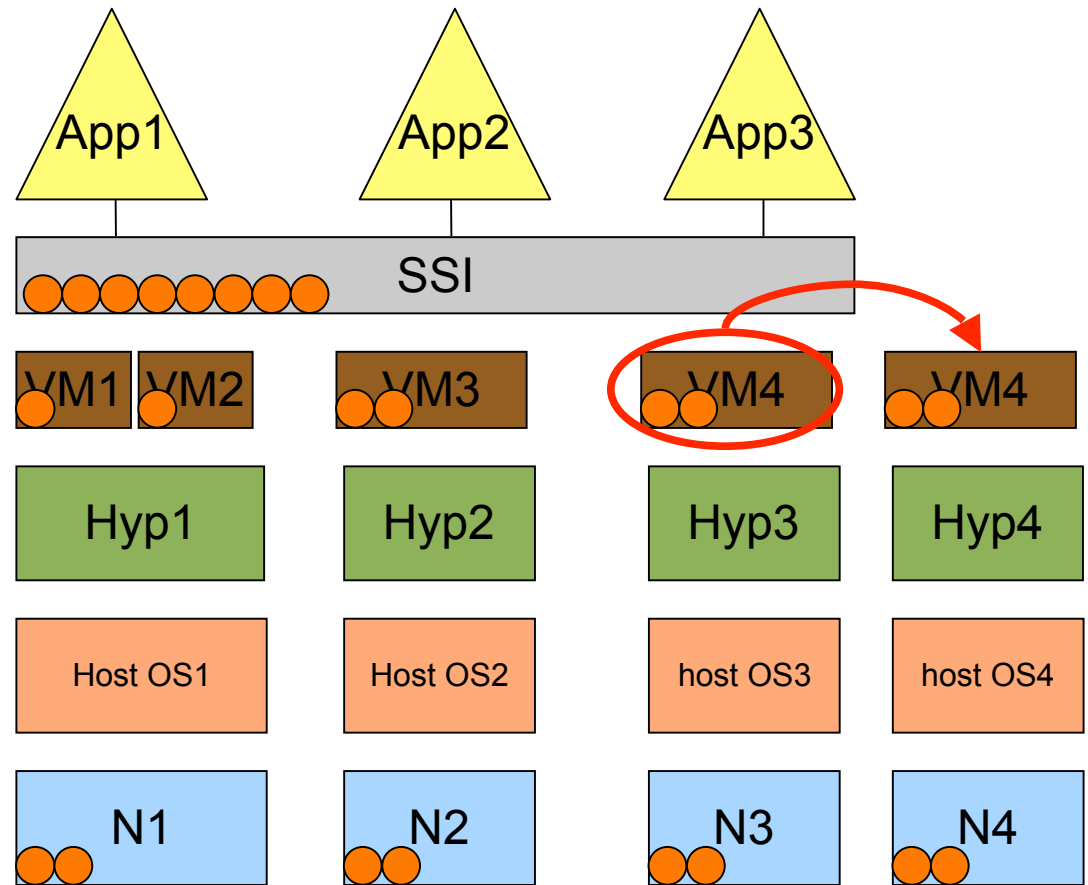
A solution to simply manage node addition/eviction in a SSI cluster





Flexibility in hardware resource management

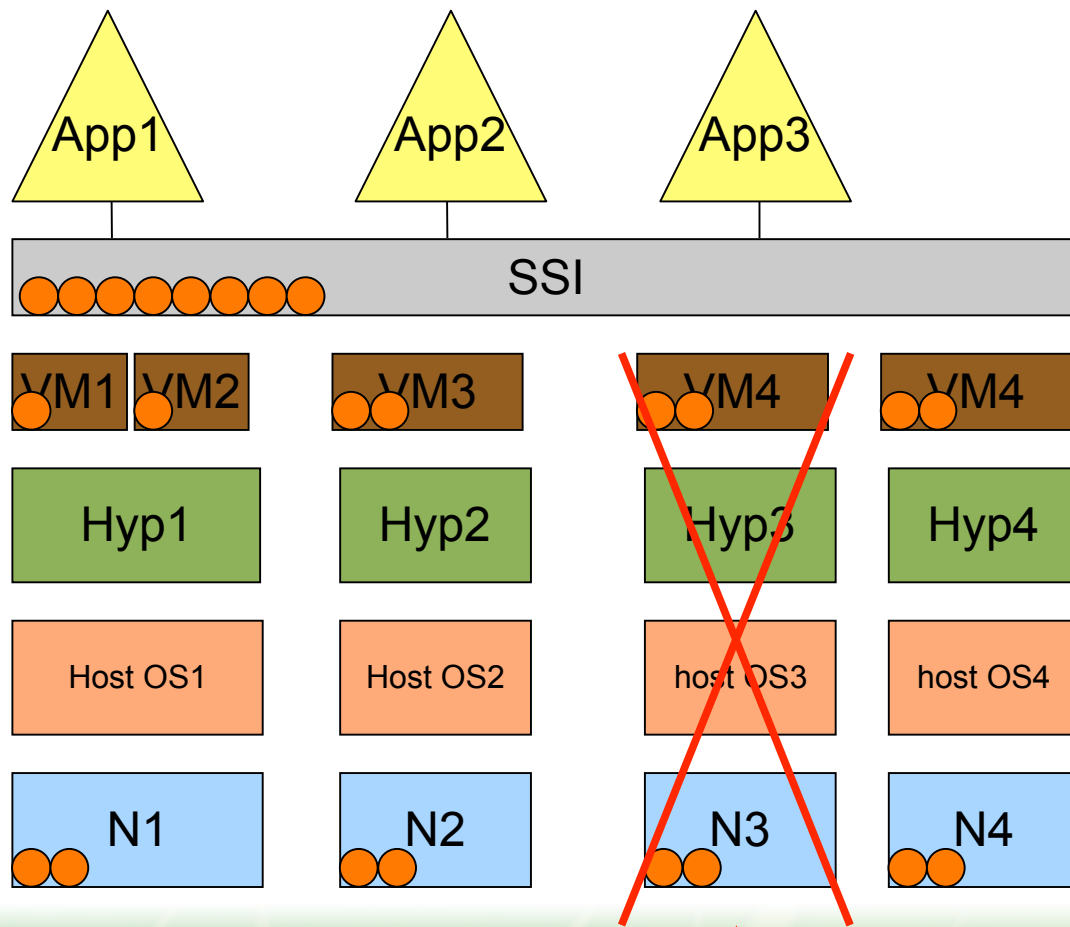
A solution to simply manage node addition/eviction in a SSI cluster





Flexibility in hardware resource management

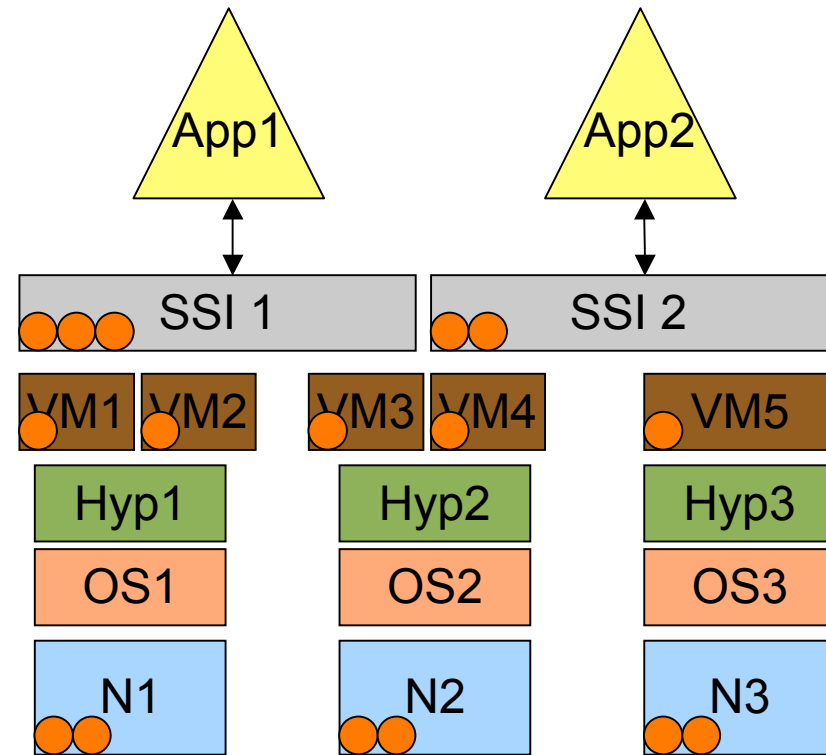
A solution to simply manage node addition/eviction in a SSI cluster





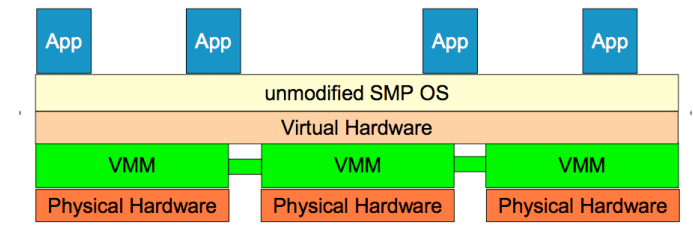
Combining Virtualization and SSI Isolation of Two Distinct SSIs

Running several isolated SSIs on the same cluster nodes





- Aggregation of hardware resources (LinuxSSI) combined with flexibility in resource management, isolation, portability (hypervisor & VM)
- **Concept of distributed hypervisor**
 - ScaleMP intrinsically not scalable
 - Running an SMP OS on a distributed system
 - Linux SSI approach attractive
 - Distributed system well-suited to the distributed nature of a cluster
 - Resource aggregation to provide virtual resources
- Migration of virtual clusters in clouds
 - Resource management policies



Credit: Erich Focht, NEC



- Node failure management in LinuxSSI
- Revisit LinuxSSI implementation with the integration of virtualization mechanisms in Linux kernel
 - Containers
- Performance
 - Multi-level scheduling, VM scheduling



- **Cloud & Grid computing**
- **Overview of XtreemOS Grid OS**
- **XtreemOS in Cloud computing scenarios**
- **Conclusion**



- **XtreemOS key features for cloud computing**
 - Virtual organization management & security
 - XtreamFS distributed file system spanning multiple administrative domains
 - Virtual clusters based on the SSI technology
- **Open research issues**
 - Standardized cloud API
 - VM management at large in scientific clouds
 - Virtual Grids & security
 - ...
- **XtreemOS: a sound system for experimenting scientific cloud computing scenarios**
 - XtreamOS = Linux + a set of packages (Mandriva, Redflag Linux)
 - **We are open to collaborations!**





- **First public release of XtreemOS software**
 - **October 1st, 2008**
 - Mandriva & RedFlag Linux distributions
 - <http://www.xtreemos.eu> & sf.net

- **Demonstrations**
 - Internet of Services 2008, Brussels, September 22-23, 2008
 - OGF-24, Singapore, September 15-19, 2008
 - SC '08, Austin, November 16-20, 2008 (XtreemOS booth #3019)



Thank you for your Attention

<http://www.xtreemos.eu>

Christine.Morin@inria.fr

