Snooze: an autonomic and energy-efficient management system for virtualized clusters

Eugen Feller, Matthieu Simonin, **Anne-Cécile Orgerie** and Christine Morin

Myriads, IRISA, Rennes, France

GreenDays@Luxembourg, January 29, 2013

Energy Consumption in Clouds

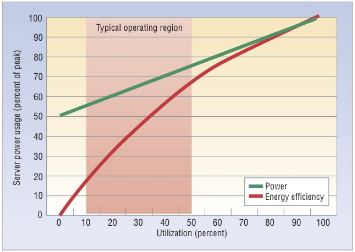
- **Clouds consume huge amounts of energy**
  - Google in 2010: 900 000 servers, ~2 billion kWh
  - Increasing demand in Cloud computing

Energy Consumption in Clouds

- **Clouds consume huge amounts of energy**
  - Google in 2010: 900 000 servers, ~2 billion kWh
  - Increasing demand in Cloud computing
- **Data centers are rarely fully utilized**
  - Highly fluctuating resource demands
  - Low utilization on average

Energy Consumption in Clouds

- **Clouds consume huge amounts of energy**
  - Google in 2010: 900 000 servers, ~2 billion kWh
  - Increasing demand in Cloud computing
- **Data centers are rarely fully utilized**
  - Highly fluctuating resource demands
  - Low utilization on average
- **Servers lack power proportionality**
  - High idle power consumption
  - Energy efficiency significantly drops under light loads



Energy Saving Approaches

- **Slow down the individual server components (e.g. CPU, memory)**
  - Becomes less attractive on modern hardware (Le Sueur et al. (2010))

Energy Saving Approaches

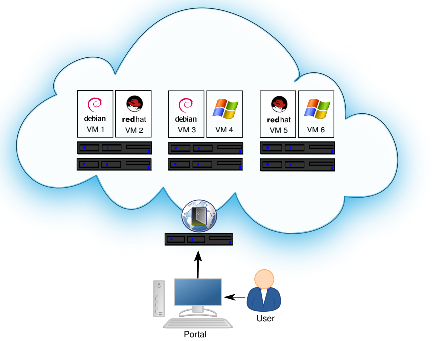
- **Slow down the individual server components (e.g. CPU, memory)**
  - Becomes less attractive on modern hardware (Le Sueur et al. (2010))
- **Transition parts of the server components into a sleep state**
  - Not always easy, minor energy savings

Energy Saving Approaches

- **Slow down the individual server components (e.g. CPU, memory)**
  - Becomes less attractive on modern hardware (Le Sueur et al. (2010))
- **Transition parts of the server components into a sleep state**
  - Not always easy, minor energy savings
- **Transition entire servers into a sleep state**
  - Entering sleep states can yield significant energy savings

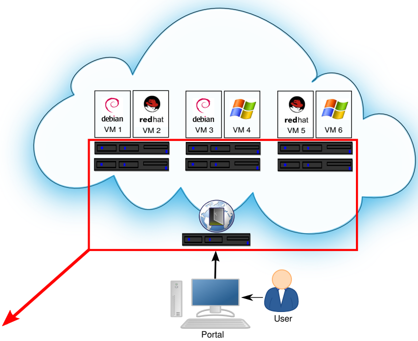
Context: Infrastructure-as-a-Service (IaaS) Clouds

- Provide compute capacity in the form of **Virtual Machines (VMs)**
  - Illusion of a computer running its own operating system
- **Server virtualization**
  - Multiple VMs on a server
  - Live migration



Context: Infrastructure-as-a-Service (IaaS) Clouds

- Provide compute capacity in the form of **Virtual Machines (VMs)**
  - Illusion of a computer running its own operating system
- **Server virtualization**
  - Multiple VMs on a server
  - Live migration



VM management system

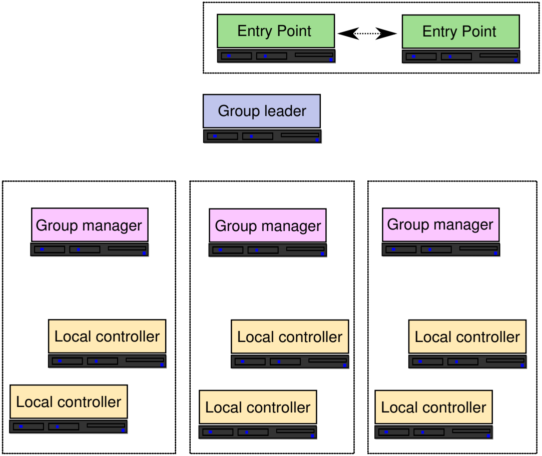
- Controls the servers
- Accepts user requests
- Places VMs on the servers

Snooze in one slide

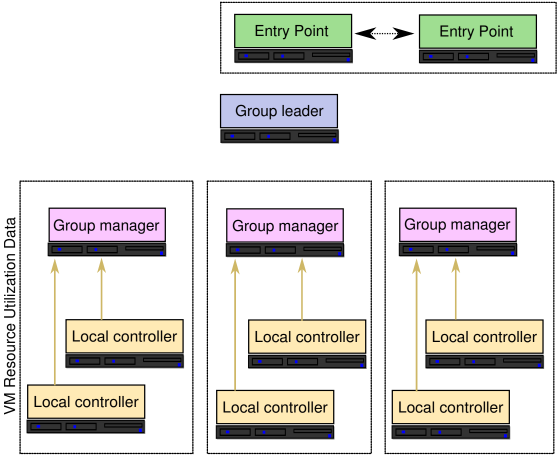
- Open-source, scalable, autonomic and energy-efficient virtual machine management framework for private clouds
- Developed within the French ANR EcoGrappe project (Energy Consumption Management in Clusters) by Eugen Feller during his PhD
- Currently supported by an Inria ADT project with one full-time engineer for 2 years (Matthieu Simonin)
- Released under the GPL v2 licence
- Tested on Grid'5000 and on an EDF cluster (40 nodes)

http://snooze.inria.fr

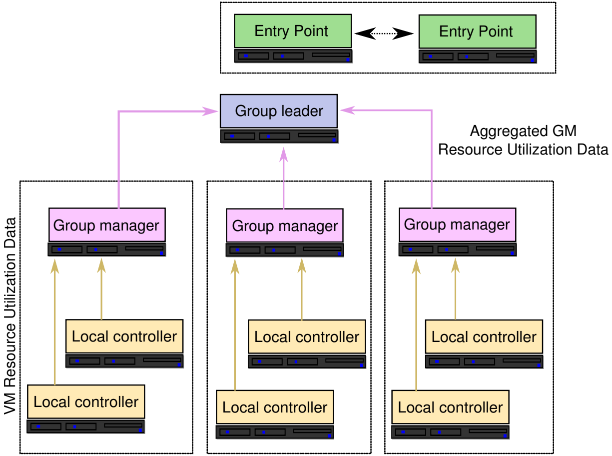
System Architecture



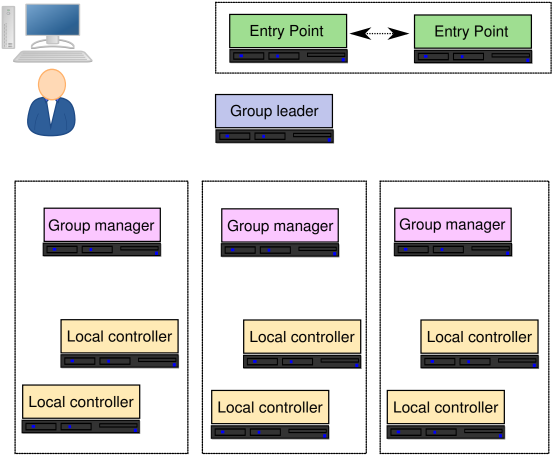
System Architecture



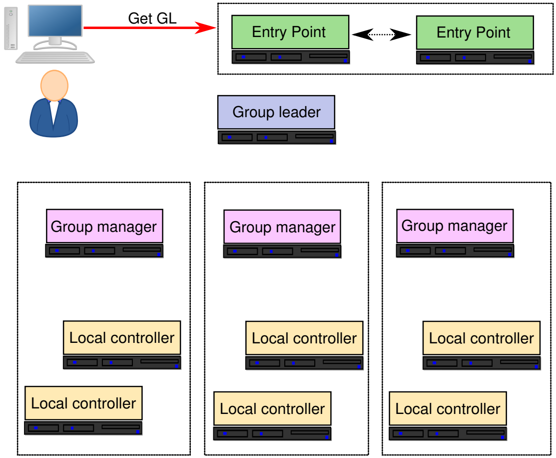
System Architecture



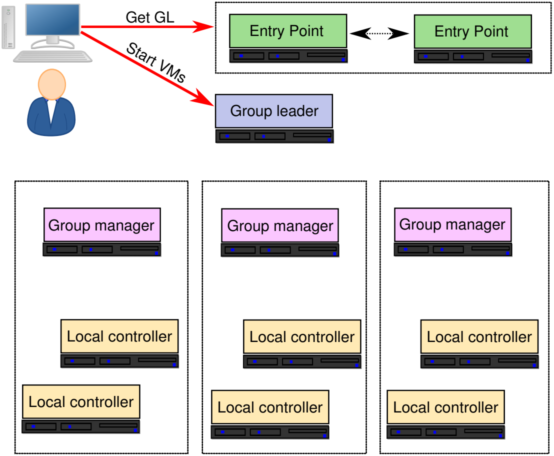
VM Submission Example



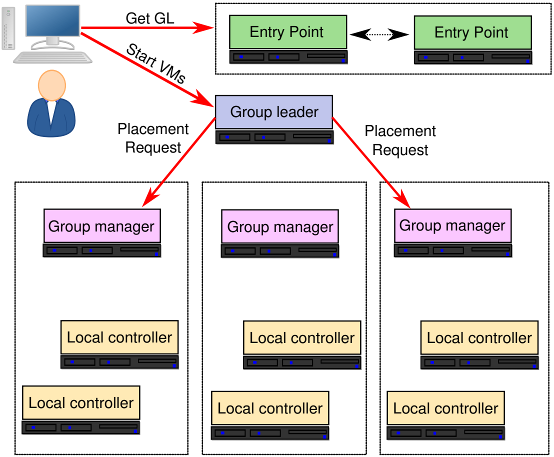
VM Submission Example



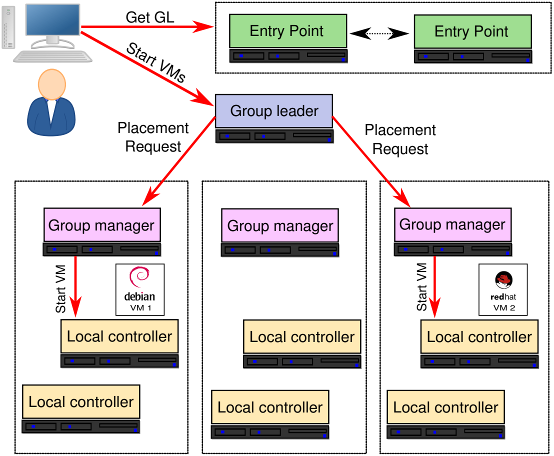
VM Submission Example



VM Submission Example



VM Submission Example

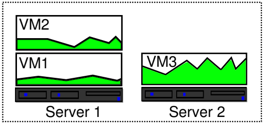


Snooze hierarchy in a nutshell

Components	Roles	Functionnalities
Entry point	user interface	availability
Group Leader	deal with client requests assign GM to LCs dispatch VMs among GM	load-balancing
Group Manager	place VMs overload/underload mitigation VM consolidation (migration plan)	scalability self-optimization **energy-efficiency**
Local Controller	monitors VMs enforce VM states (start, migration, ...) power off resources overload/underload detection	self-configuration fault-tolerance
Virtual Machines	run applications	user-friendly

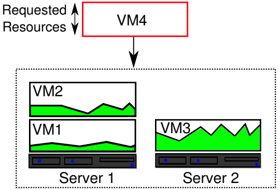
Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement



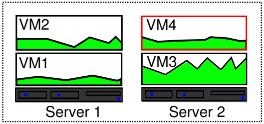
Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement



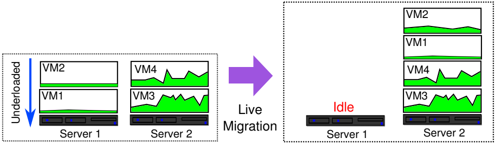
Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement



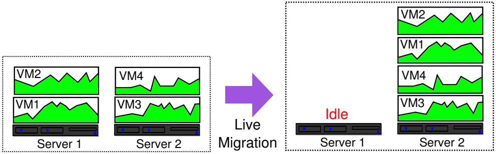
Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement
  - Server underload detection and mitigation



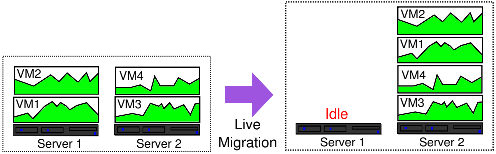
Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement
  - Server underload detection and mitigation
  - Periodic VM consolidation



Idle Time Creation

- **Three methods**
  - Energy-efficient VM placement
  - Server underload detection and mitigation
  - Periodic VM consolidation



Self-optimization for energy efficiency

Mechanisms and Algorithms for Energy Efficiency

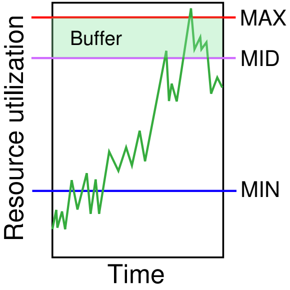
- **How to favour idle times**
  - Energy-efficient VM placement
  - Underload server detection and mitigation
  - Periodic VM consolidation
- **Server overload detection and mitigation**
- **Power management**
  - Automatic detection and power cycling of idle servers
  - Server wakeup when not enough resources are available

Underload and Overload Mechanism

- **How to deal with underload and overload situations?**
  - Detection of server underload/overload situations
  - Relocation of VMs from underloaded/overloaded servers

Local controllers periodically estimate their resource utilization based on locally aggregated VM resource utilization data

- Multi-dimensional
  - CPU
  - RAM
  - Network Rx
  - Network Tx



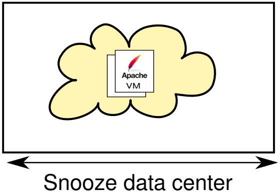
Underload Relocation Algorithm

Triggered by the GM in the event of server underload

- Key ideas
  - Move VMs from underloaded LC to LCs with enough spare capacity
  - **All-or-nothing approach:** Either migrate all VMs or none
- Description
  - Sort VMs from underloaded LC in decreasing order of estimated utilization
  - Sort destination LCs in decreasing order of estimated utilization
  - Attempt to assign the VMs to the destination LCs starting from the first one
  - If some VM could not be assigned abort the algorithm
  - . . . otherwise perform live migrations

Integrated Energy Management Evaluation

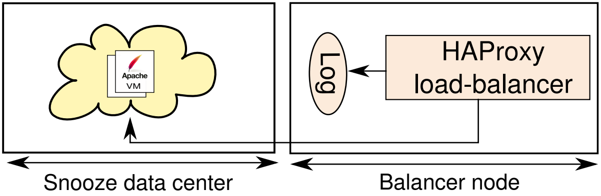
Evaluation with an elastic web service



Deployed on 34 power-metered servers of the Grid'5000 testbed

E. Feller, C. Rohr, D. Margery, and C. Morin. Energy Management in IaaS Clouds: A Holistic Approach. In the *5th IEEE International Conference on Cloud Computing (CLOUD)*, May 2012.

Evaluation with an elastic web service

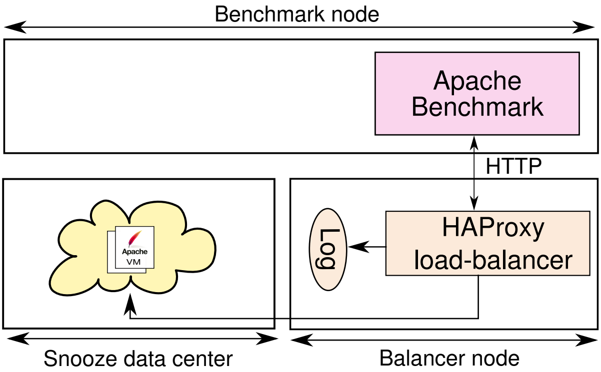


Deployed on 34 power-metered servers of the Grid'5000 testbed

E. Feller, C. Rohr, D. Margery, and C. Morin. Energy Management in IaaS Clouds: A Holistic Approach. In the *5th IEEE International Conference on Cloud Computing (CLOUD)*, May 2012.

Integrated Energy Management Evaluation

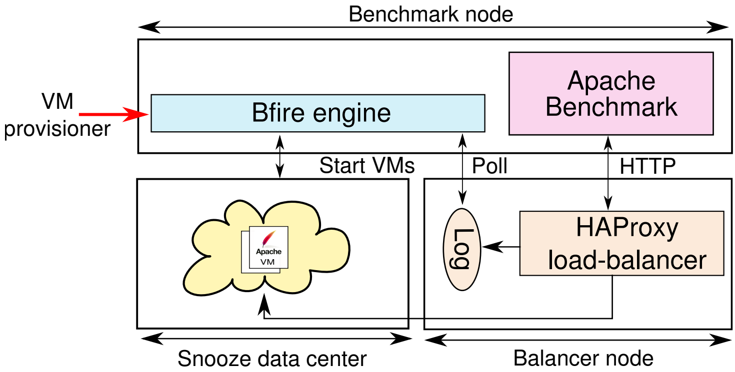
Evaluation with an elastic web service



Deployed on 34 power-metered servers of the Grid'5000 testbed

E. Feller, C. Rohr, D. Margery, and C. Morin. Energy Management in IaaS Clouds: A Holistic Approach. In the *5th IEEE International Conference on Cloud Computing (CLOUD)*, May 2012.

Evaluation with an elastic web service

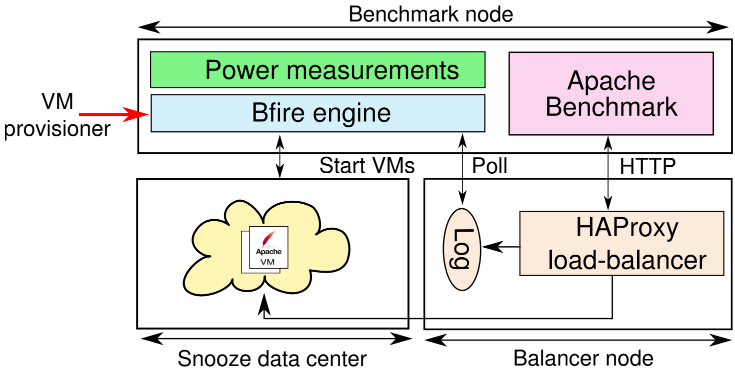


Deployed on 34 power-metered servers of the Grid'5000 testbed

E. Feller, C. Rohr, D. Margery, and C. Morin. Energy Management in IaaS Clouds: A Holistic Approach. In the *5th IEEE International Conference on Cloud Computing (CLOUD)*, May 2012.

Integrated Energy Management Evaluation

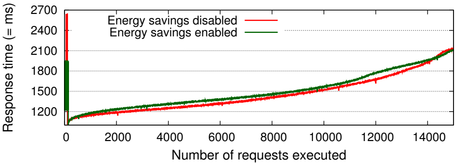
Evaluation with an elastic web service



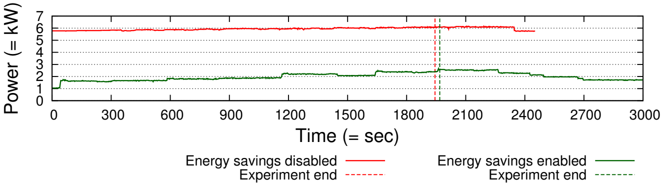
Deployed on 34 power-metered servers of the Grid'5000 testbed

E. Feller, C. Rohr, D. Margery, and C. Morin. Energy Management in IaaS Clouds: A Holistic Approach. In the *5th IEEE International Conference on Cloud Computing (CLOUD)*, May 2012.

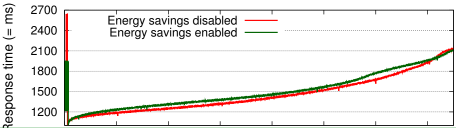
- **Apache Benchmark Performance**



- **Data Center Power Consumption**

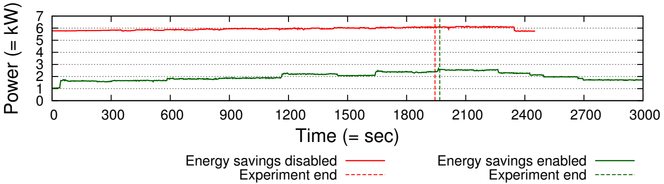


- **Apache Benchmark Performance**



Limited performance degradation
Up to 67% energy savings for the evaluated application

- **Data Center Power Consumption**



Summary

- Self-configuring and healing hierarchical architecture
- Integrated energy management approach
  - VM placement and consolidation, server underload/overload mitigation, power management
  - Four-dimensional aggregation-based underload/overload mitigation
  - First implementation of the Sercon algorithm in a real system
- A robust prototype
- Experimentally validated on the Grid'5000 testbed

Short-term Perspectives

- Further evaluate the Snooze system
  - Larger-scale experiments
  - Real-world workloads
  - Hierarchy energy overheads
- Exploit Snooze to experimentally compare state of the art VM management algorithms
- Further increase the Snooze hierarchy autonomy and energy-efficiency
  - Re-balance the hierarchy dynamically
  - Rethink hierarchical organization
  - Power-cycle idle GMs

Long-term Perspectives

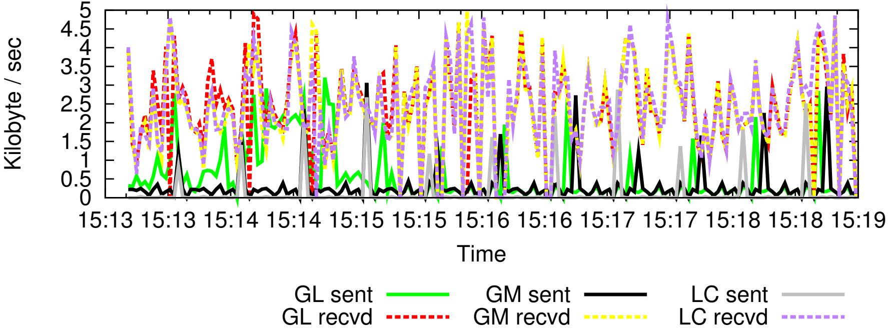
- Metrics for better capturing aggregated resource utilization data
- Improving consolidation
  - Co-location and anti-colocation constraints
  - Consider VM resource demand complementarities
  - Data center network topology aware consolidation
  - Consolidation interval predictions
- Thermal management

Questions?

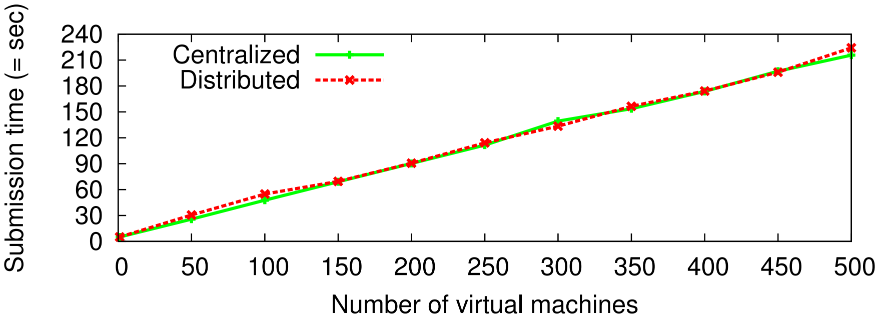
→ CURRENTLY HIRING **POST-DOCS** ON ENERGY-EFFICIENCY ISSUES
http://www.irisa.fr/myriads/open-positions/postdoc/

Backup slides

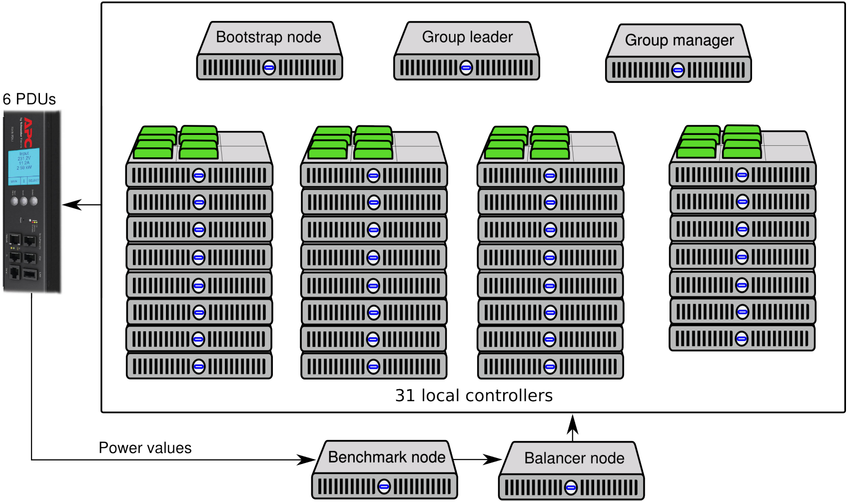
Snooze Heartbeat Overhead



Snooze Submission Time



Energy Management Data Center



Energy Management Parameters

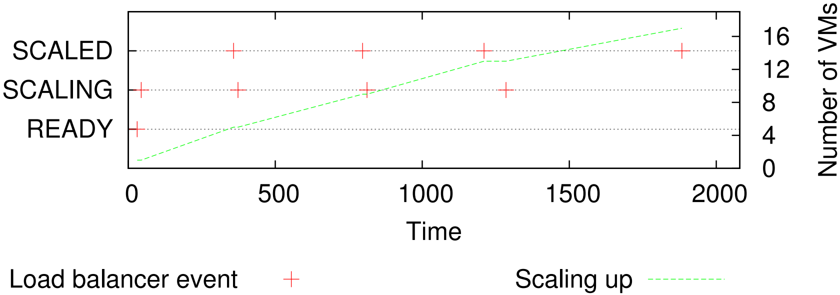
Resource	MIN, MID, MAX
CPU,	0.2, 0.9, 1
Memory	0.2, 0.9, 1
Network	0.2, 0.9, 1

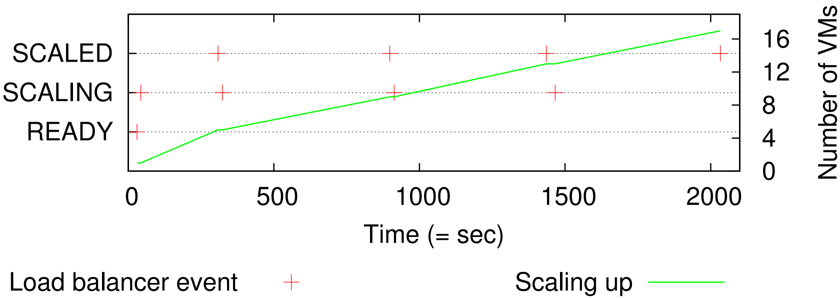
Parameter	Value
Packing density	0.9
Monitoring backlog	15
Resource estimators	average
Consolidation interval	10 min

Policy	Algorithm
Dispatching	RoundRobin
Placement	FirstFit
Overload	Greedy
Underload	Greedy
Consolidation	Sercon

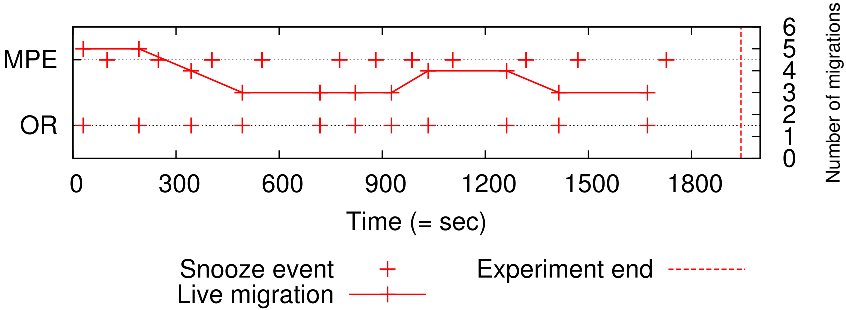
Parameter	Value
Idle time threshold	2 min
Wakeup threshold	3 min
Power saving action	shutdown
Shutdown driver	system
Wakeup driver	IPMI

Bfire Events With Energy Savings Disabled

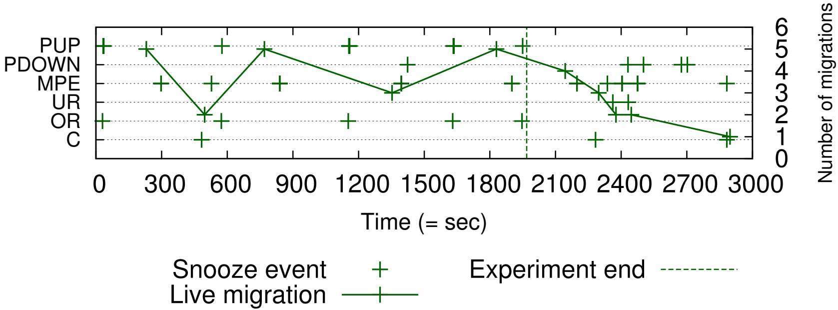



SCALED
SCALING
READY
0
500
1000
1500
2000
Time (= sec)
16
12
8
4
0
Number of VMs
Load balancer event
Scaling up

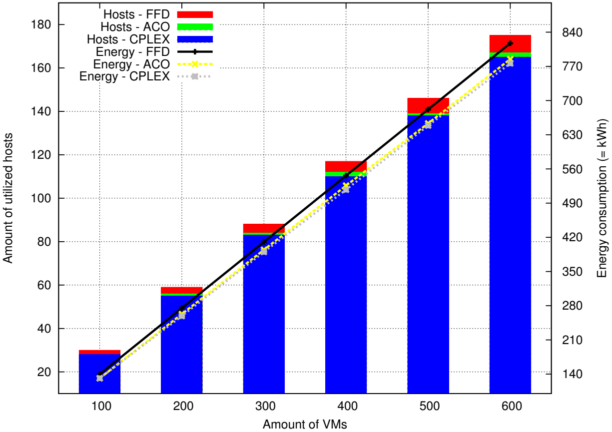
Snooze Events No Energy Savings Disabled



Snooze Events With Energy Savings Enabled



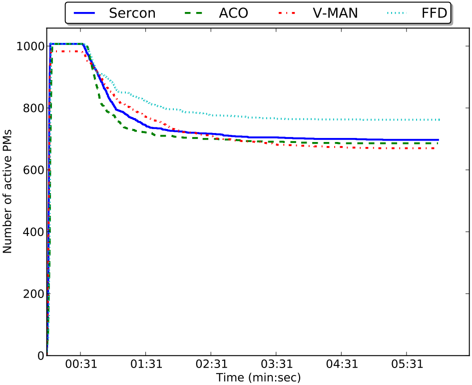
VM Placement Evaluation



Fully Decentralized VM Consolidation - Emulator Parameters

Parameter	Value
Number of PMs and VMs	1008 (resp. 6048)
Experiment duration	360s
Consolidation interval	30s
Shuffling interval	10s
Neighbourhood size	16 PMs
Considered resources	CPU, memory and network
PM total capacity vector	(48, 26, 20)
VM requested capacity vectors	(0.2, 0.5, 0.1), (1, 1, 1), (2, 1, 1), (4, 2, 2), (8, 4, 4), (16, 8, 4)

Fully Decentralized VM Consolidation - Number of Active Servers



Fully Decentralized VM Consolidation - Number of Migrations