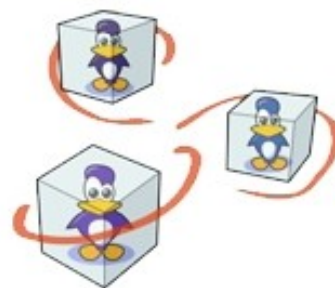


Is Virtualization Killing SSI Research?

Jérôme Gallard
Kerrighed Summit
Paris – February 2008

Supervisor : Christine Morin
Co-supervisor: Adrien Lèbre



Kerrighed
Linux clusters made easy



Context

- Virtualization / SSI

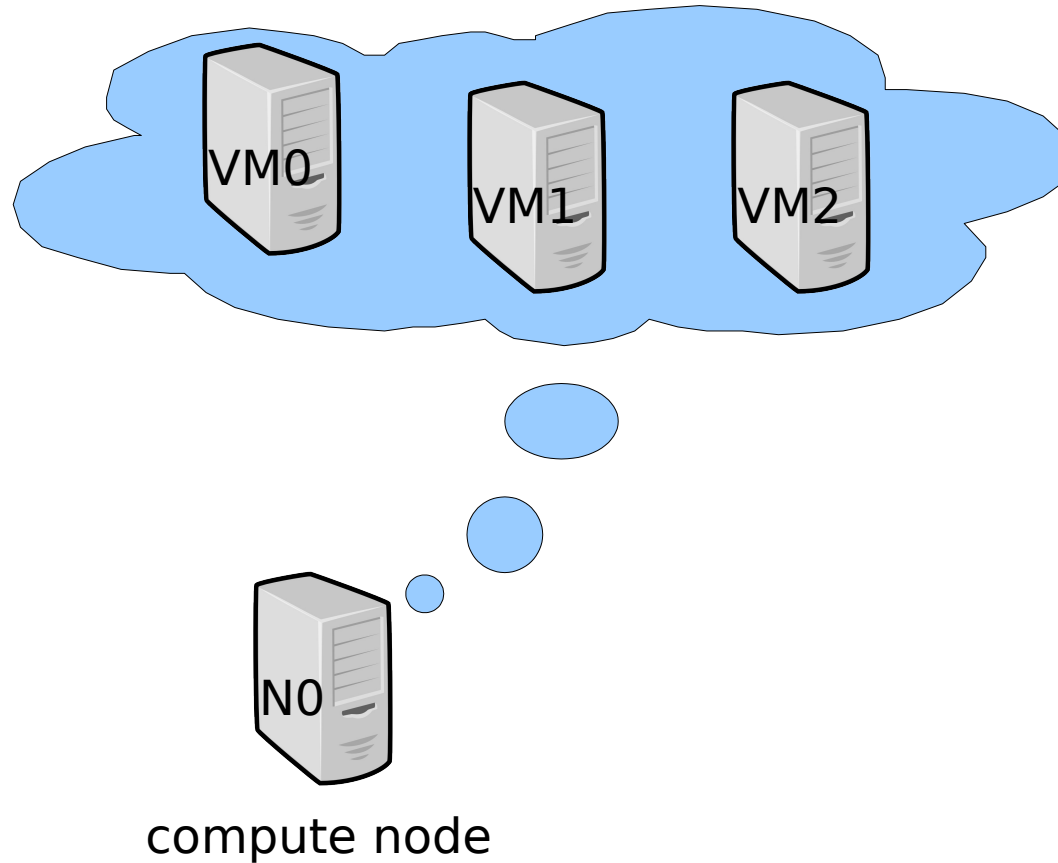
→ Combining Virtualization and SSI

Conclusion

- Lessons learnt
- Perspectives

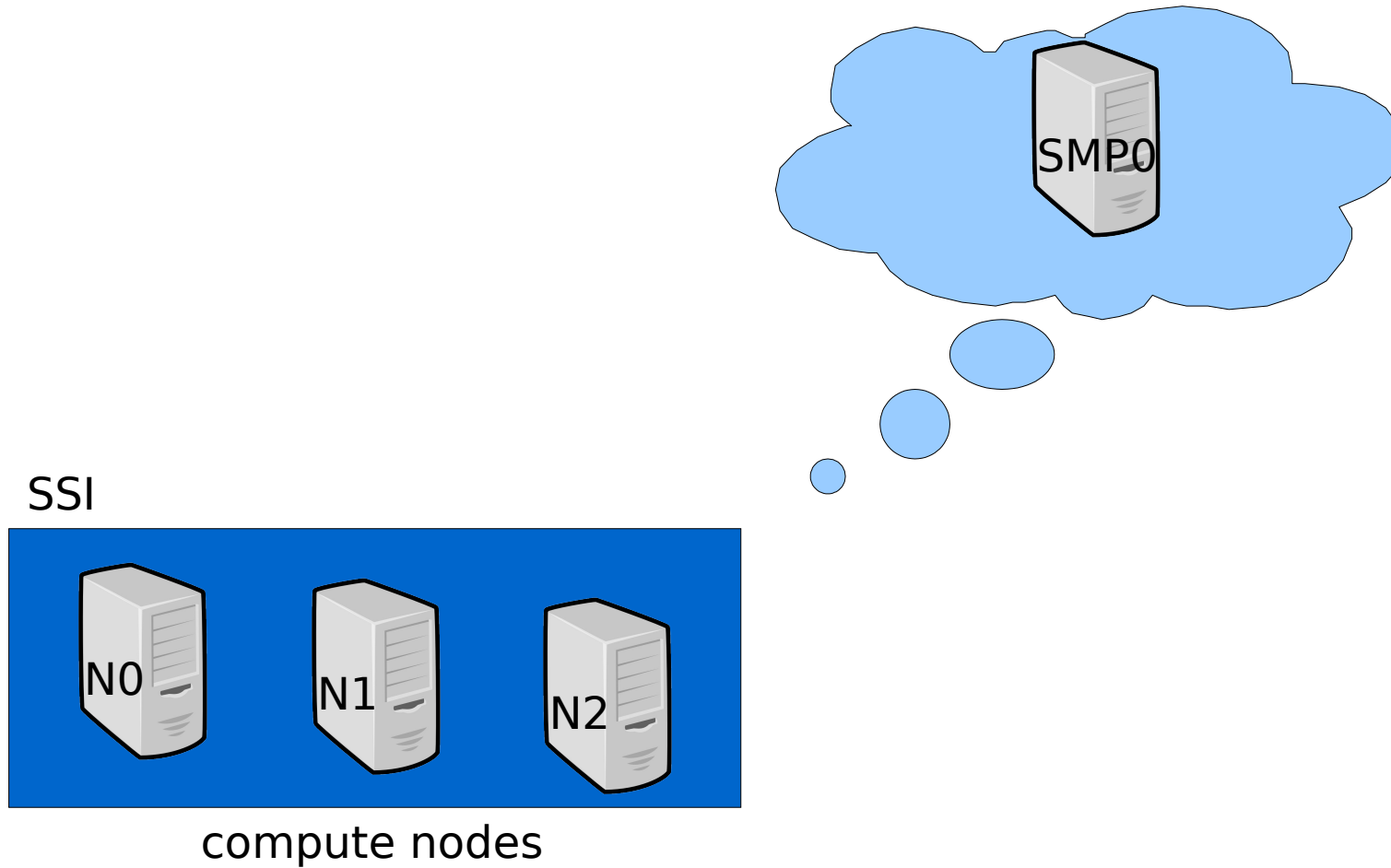
Virtualization Technologies

virtualized machines running on the compute node



SSI Technologies

virtualized an SMP machine running on compute nodes



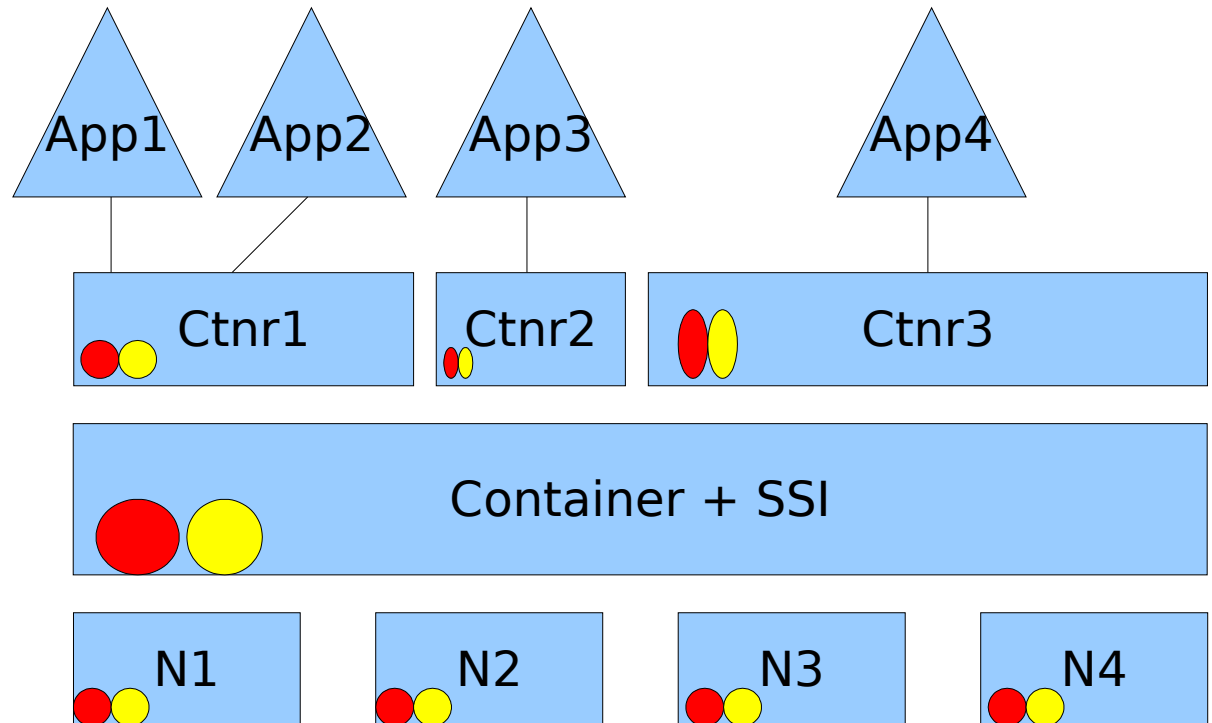
Our approach

Make several combinations between
SSI
and
Container, type1 and **type11** virtualization

Combining Container and SSI

Container on the top of SSI

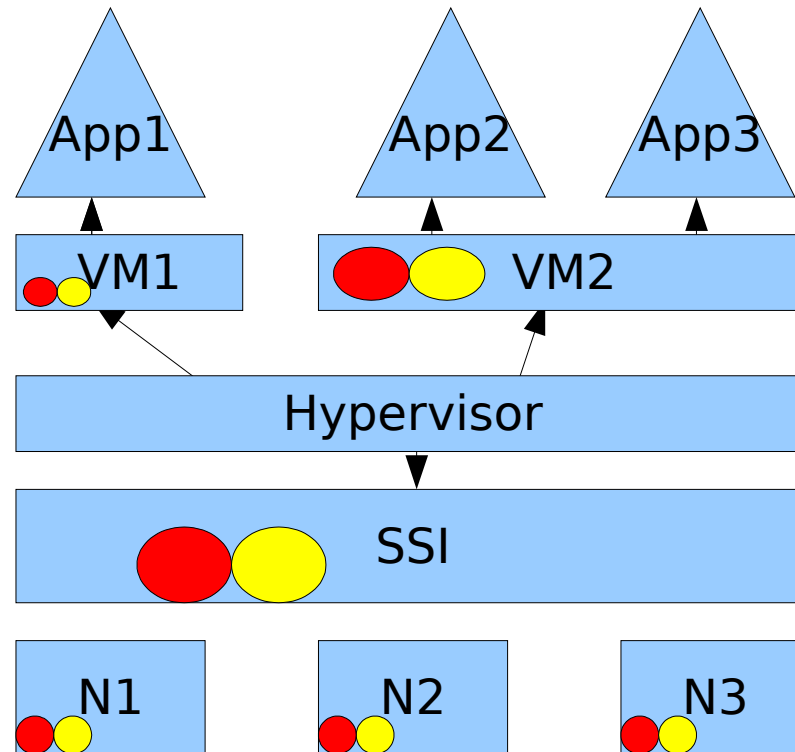
Allow containers to get advantage of the global resources provided by the SSI



Combining Virtualization and SSI

Typell-Virtualization upon SSI 1/2

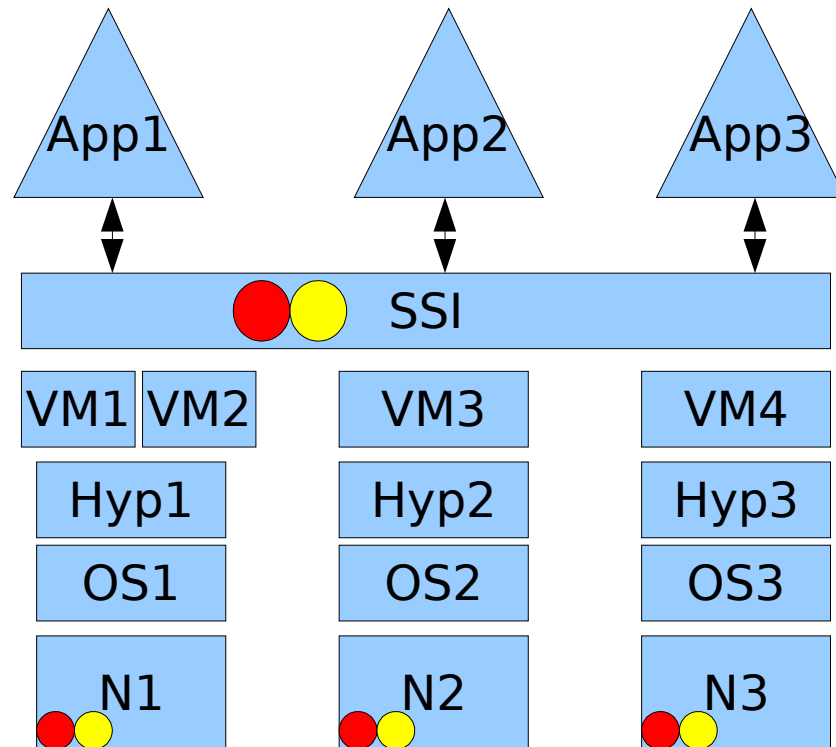
Allow VMs to get advantage of the global resources provided by the SSI



Combining Virtualization and SSI

SSI upon Type1-Virtualization 2/2

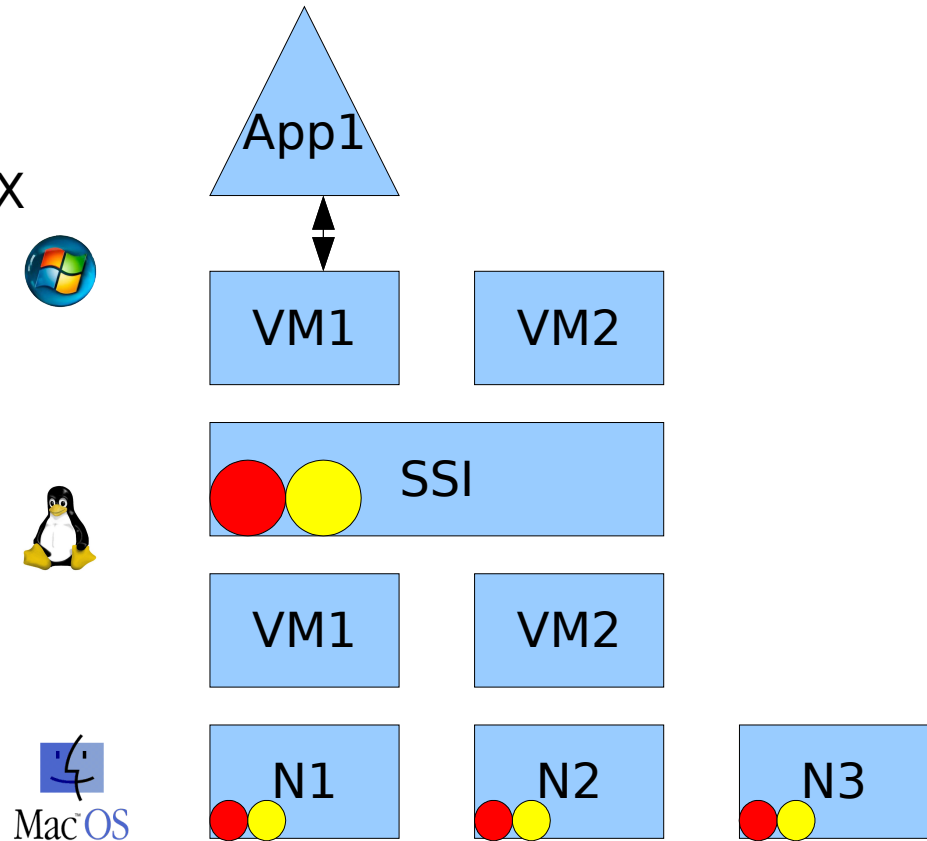
Allow the SSI to be run on the top of several hardware architecture



Combining Virtualization and SSI

Three layers of flexibilities

Possibility to run a windows application on a linux SSI cluster running upon MACOS X systems

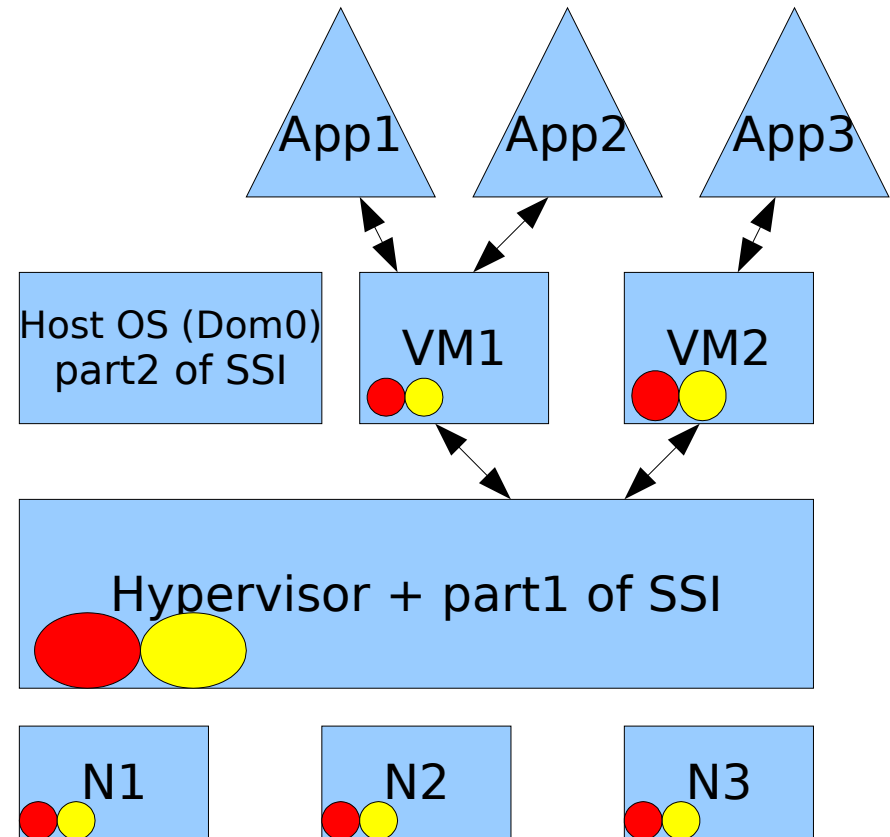


Combining Virtualization and SSI

Type1-Virtualization upon SSI 1/2

Part1 of SSI:
hide the distributed aspect
of the cluster
Does ScaleMP should be
a SSI Part1?

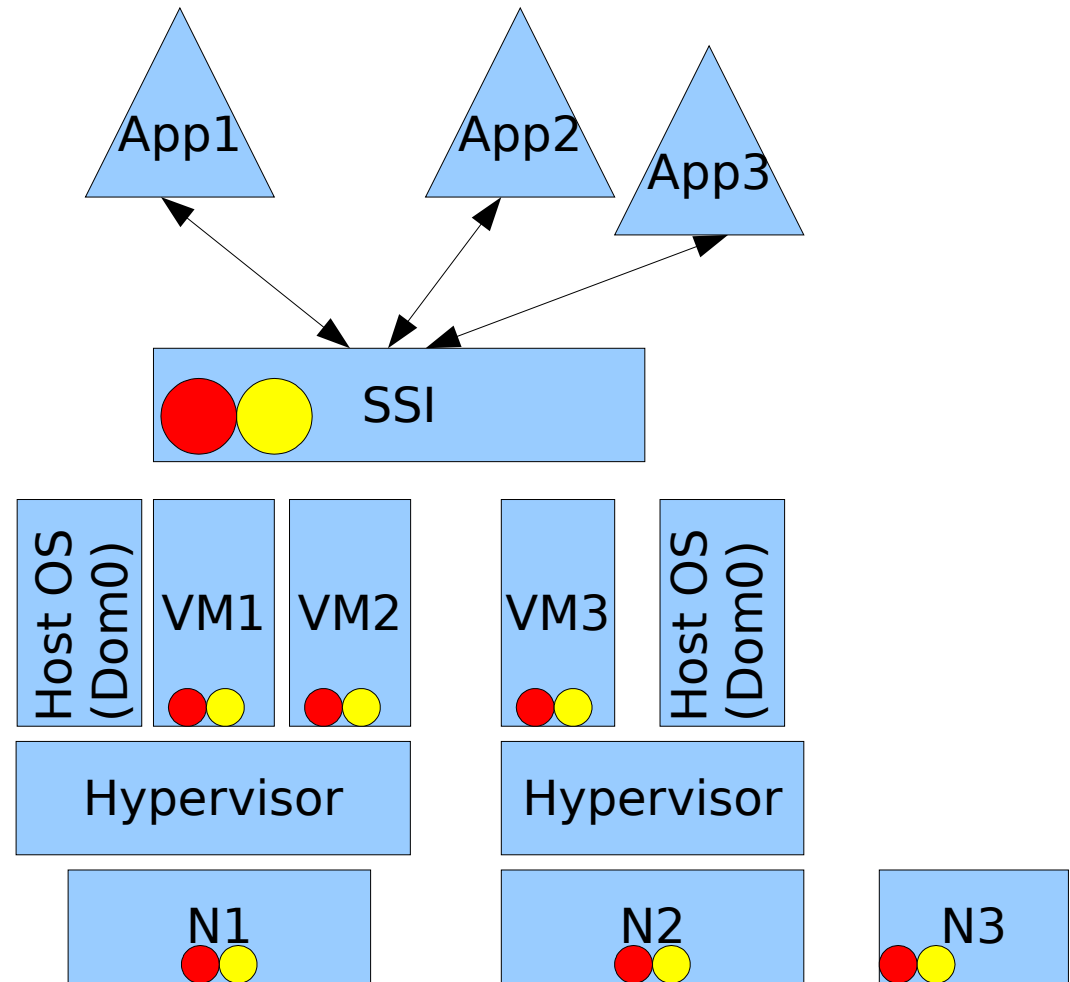
Part2 of SSI:
give tools like checkpoint,
migration...



Combining Virtualization and SSI

SSI upon Type1-Virtualization 2/2

Same as already seen



Lessons Learnt 1/4

- Containers on top of single system image clusters
 - The SSI system combines all the advantages enabled by containers on a real SMP machine in a cluster environment

Lessons Learnt 2/4

- Virtual Machines on top of SSI clusters
 - Virtualization solves application portability issues
 - example :
 - Possibility to run an application developed for process technology A and OS B on top of computers running SSI OS based on OS C and developed for architecture processor D.

Lessons Learnt 3/4

→ SSI on top of VMs

- VM migration and suspend functionalities provide:
 - Flexible, simple and "On Demand" resource allocation mechanisms for the applications
 - Transparent adaptation mechanisms in case of hardware changes (for instance, node additions / removals).

Lessons Learnt 4/4

- Virtualization and SSI complement each other
- SSI abstracts resources
- Virtualization adds another level of hardware flexibility

Conclusion

- Performance evaluation:
- Work in progress

- Combining VM/SSI setup:
 - Very attractive
 - Installation and use are quite simple (evaluated with Typell)

- Migrate an SSI cluster from Rennes to Sophia

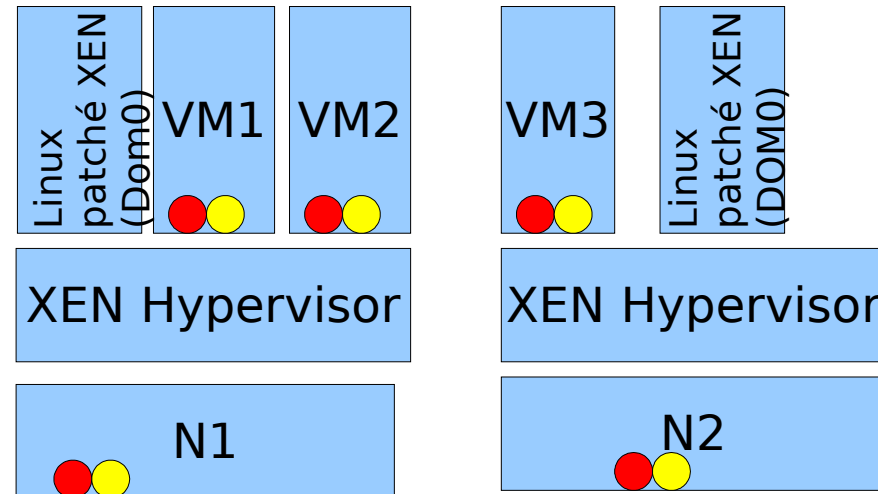
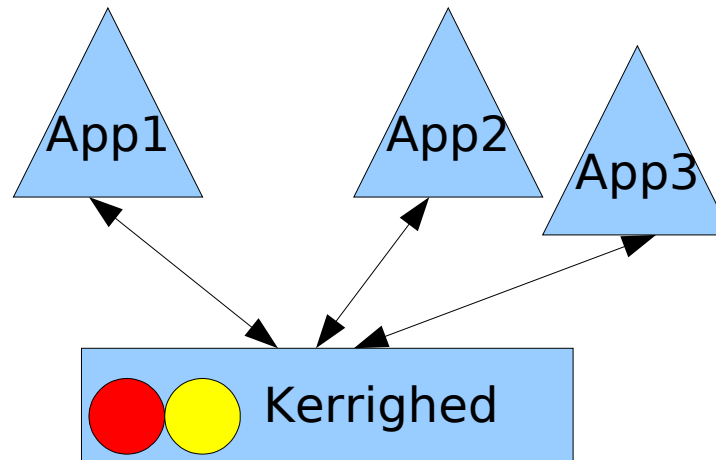
References

- [Grit2007] L. Grit, D. Irwin, V. Marupadi, P. Shivam, A. Yumerefendi, J. Chase, and J. Albrecht. *Harnessing virtual machine resource control for job management*. March 2007.
- [Goldberg1973] R. P. Goldberg. *Architecture of virtual machines*. AFIPS National Computer Conference, July 1973.
- [OpenVZ2007] OpenVZ. OpenVZ welcome page, 2007. Available as http://wiki.openvz.org/Main_page.
- [Chroot2007] GNU. Chroot, 2007. Available as <http://www.gnu.org/software/coreutils/manual/coreutils.html#chroot-invocation>
- [Barham2003] P. Barham, B. Dragovic, K. Fraser, S. Hand, T. Harris, A. Ho, R. Neugebauer, I. Pratt, and Warfield. *Xen and the art of virtualization*. Bolton Landing, New-York, USA, October 2003. SOSP'03.
- [VMware2007] VMware. *VMware welcome page*, 2007. Available as <http://www.vmware.com>
- [Bellard2005] F. Bellard. *QEMU, a fast and portable dynamic translator*. Technical report, USENIX Association, 2005.
- [Huelsenbeck] J. P. Huelsenbeck and F. Ronquist. *MrBayes: a program for the Bayesian inference of phylogeny*. <http://golab.unl.edu/teaching/SBseminar/manual.pdf>

Experimentations to do SSI upon type1-virtualization

→ Kerrighed upon XEN

- Does it work ?
- LiveMigration ?
- Performance ?



Advantages:

- + Isolation
- + server consolidation
- + Suspend/restart
- + Application portability

Disadvantage:

- VM portability

Experimentations to do SSI upon type1-virtualization

→ Performance evaluations

Make comparison between clusters with:

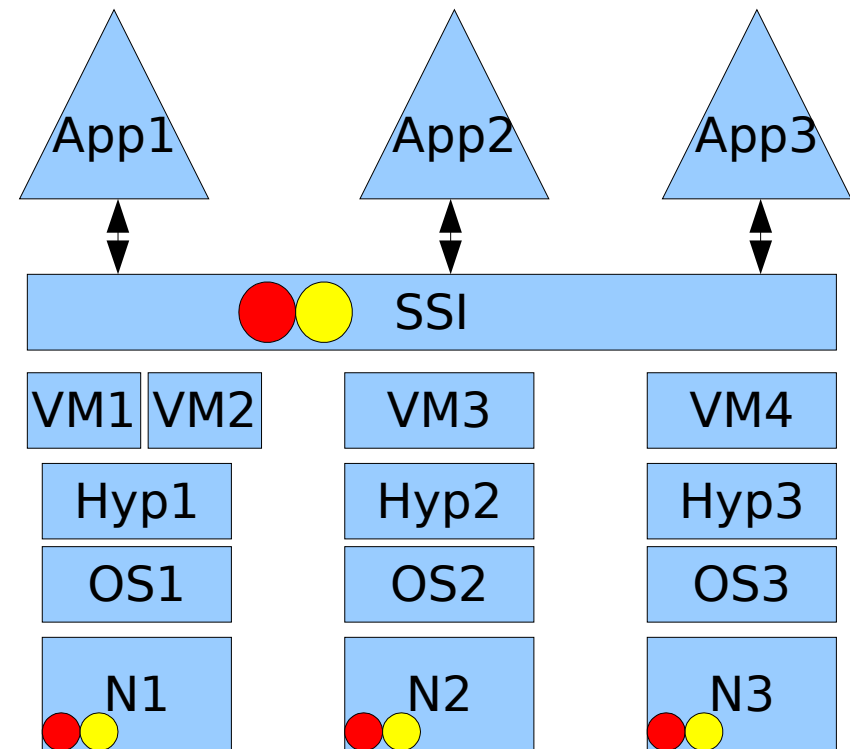
- Windows OS / VM / Kerrighed / Apps
- Linux / VM / Kerrighed / Apps
- Kerrighed / Apps

VM:

- Vmware (virtualization only)
- QEMU (emulation)

Apps:

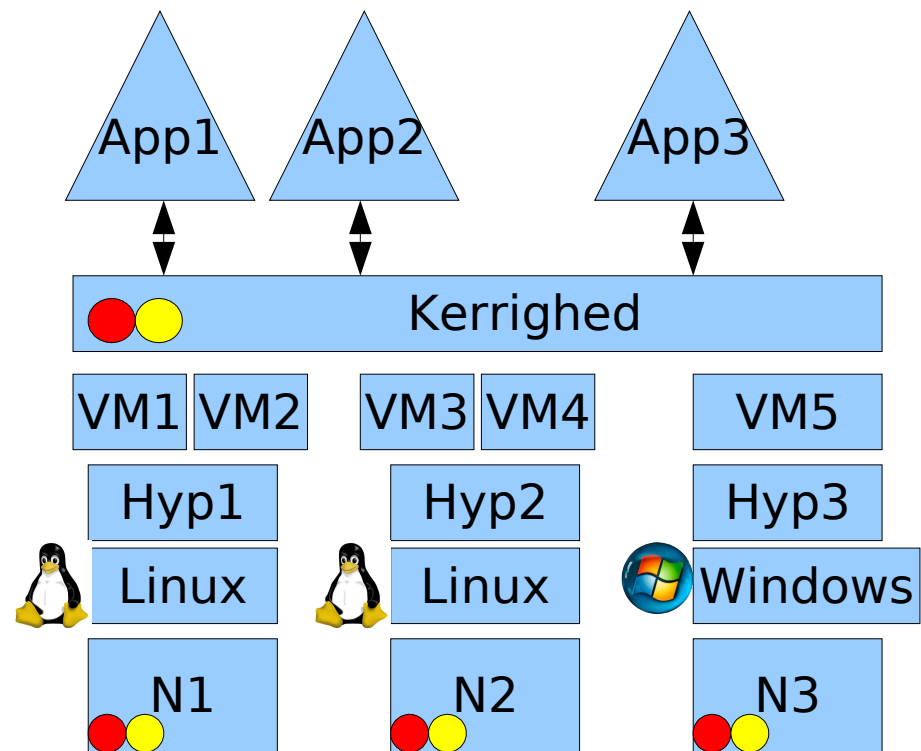
- MrBayes in MPI (CPU intensive)
- "make" kernel
- Apache



Experimentations to do

SSI upon typeII with multi-host OS

- Kerrighed upon Windows and Linux clusters
- Performance evaluations



Advantages:

- + Application portability
- + Isolation
- + server consolidation
- + Suspend/restart
- + VM portability

Experimentations to do typell upon SSI

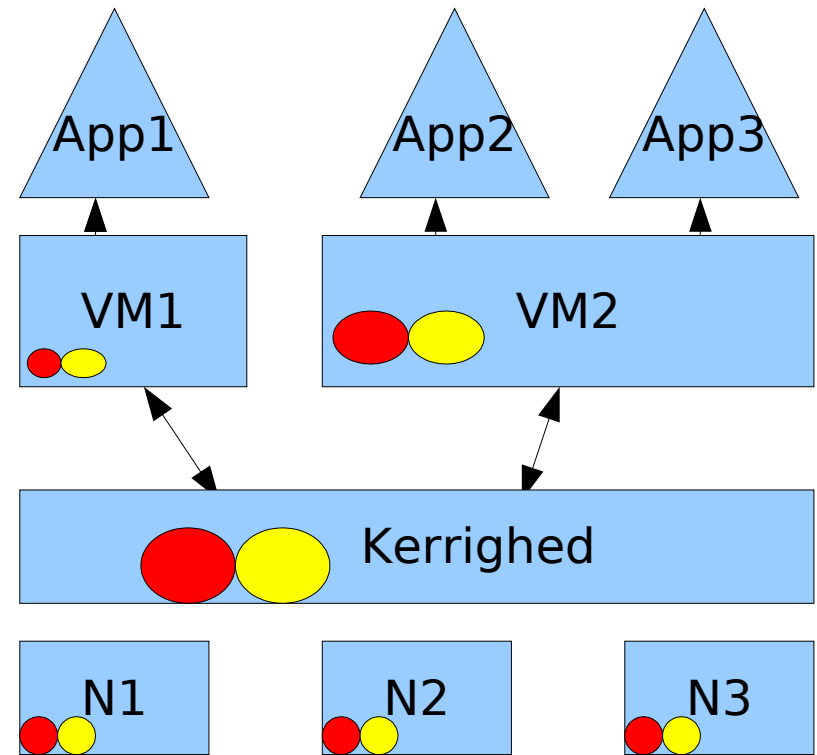
- VMs upon de Kerrighed
- LiveMigration OK ?
- Kerrighed_migration OK ?
- What is the most interesting ?

Advantages:

- + Isolation
- + server consolidation
- + Suspend/restart
- + Application portability

Disadvantages:

- VM portability



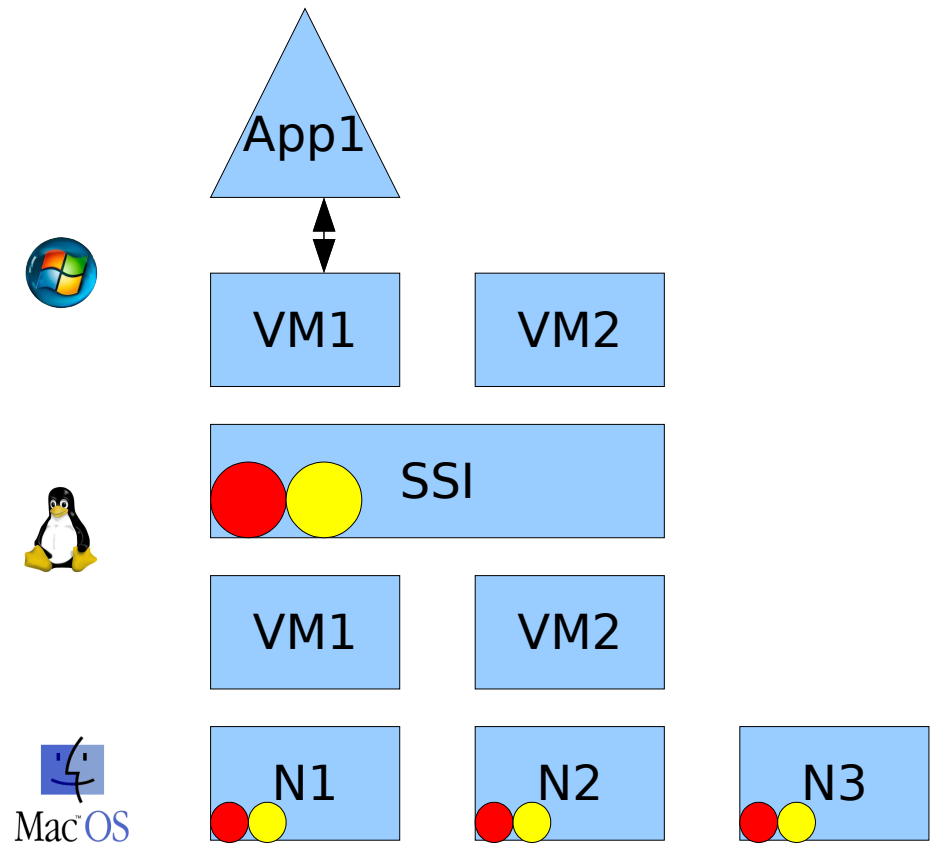
Experimentations to do

Three layers of flexibilities

It seems that it's not possible to run VMware upon VMware

Why this limitation ?

It seems it is possible to run QEMU upon VMware



Internship subject

- Automatic deployment VM/Kerrighed
- Extension of the work of Nicolas Aupetit
 - system able to deploy VMs upon grid5000
 - system able to deploy Kerrighed upon nodes (work of Nicolas Aupetit)
- For instance:
 - `start_kerrighed -vm Vmware -nbvm 6 -nbnode 3`

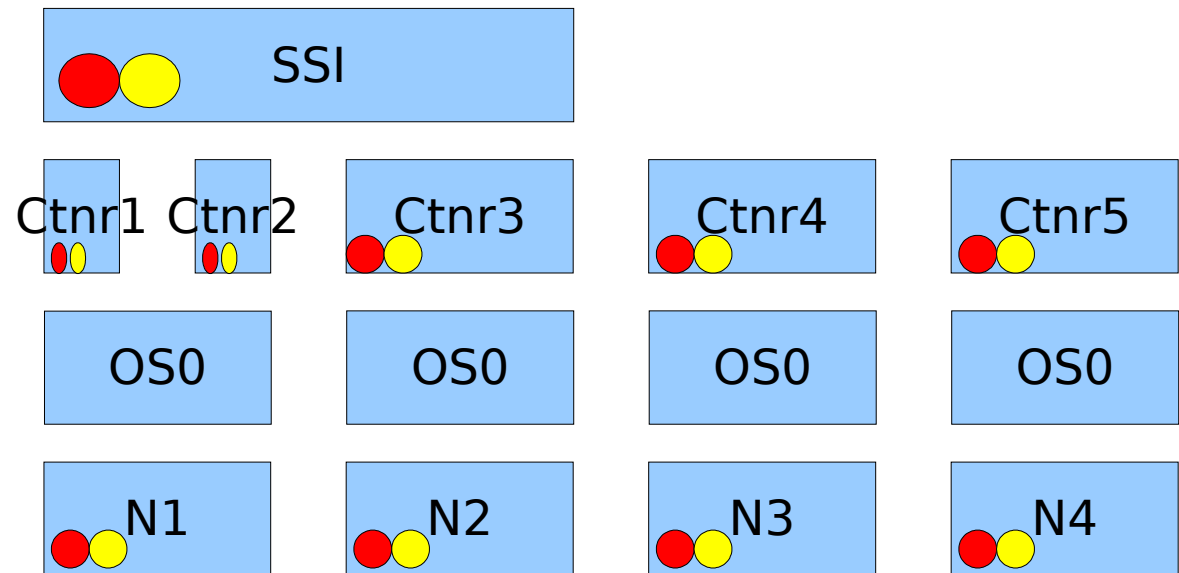
Experimentations

First results

Combining container and SSI

SSI on the top of Container 2/2

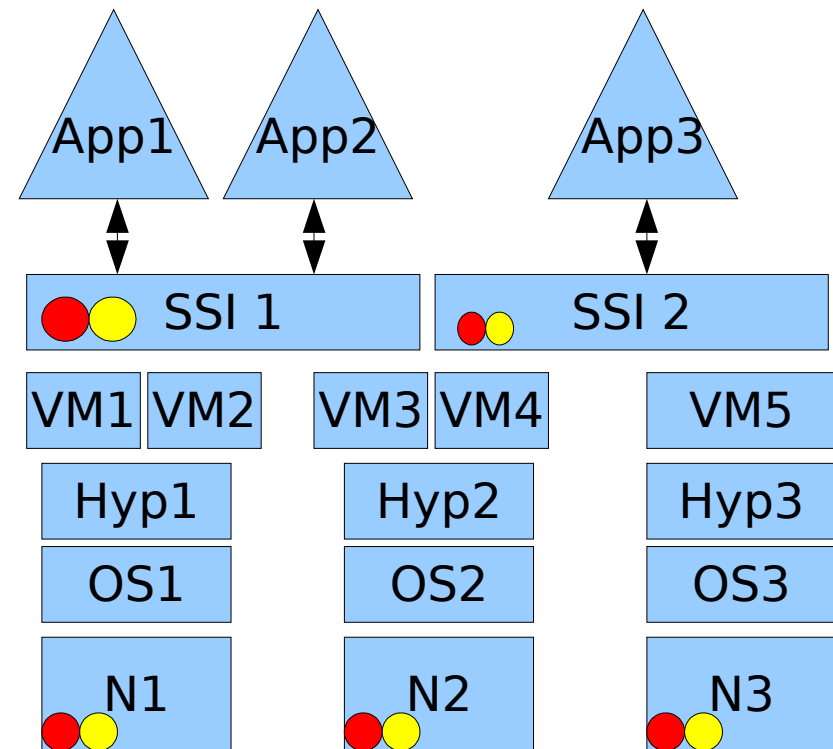
Architecture irrelevant:
The same kernel is shared
between all containers
(they have not their own
kernel)



Combining Virtualization and SSI

Isolation of Two Distinct SSIs

Possibility to run several SSIs on the same cluster.



Future Works 1/2

- Extend the Goldberg's classification to present in a uniform way:
 - The hardware,
 - The emulated hardware,
 - The OS and SSI,
 - The different kinds of virtualization techniques,
 - The containers.

Future Works 2/2

- Investigate SSI scheduler and hypervisor scheduler:
 - does the SSI scheduler directly manage VMs ?
- Extend to the Grid: interests of such hybrid approaches (VM/SSI) in XtremOS
- What about ScaleMP ?

VMs vs SSI

Is virtualization killing system image research?

What's about combining both approaches ?