

Virtualization

- Not to be confused with clouds
- Talking about virtual machines here
- Can be provisioned in a static or dynamic way

Virtualization for services:

- Done already at many sites, not controversial
- Transparent for the VOs

Virtualization of batch resources:

Several possibilities, tried at different sites

- Let the batch system start a VM and start payload inside
- Use the batch scheduler to launch VMs, and connect them to the batch farm
- Or launch VMs externally, and run the batch software inside

Batch virtualization: why ?

Whats the problem you want to solve ?

- Adaption to software and OS environment
- Encapsulation of user payload
 - Containers could eventually solve that as well (as of SL6)
 - Better adaption to what the payload expects
- Reduction of operational costs
 - Automation of routine tasks, eg intrusive updates
 - Easy draining of physical nodes for interventions (?)
- Better support of new hardware (use recent OS on hypervisors without having to wait for applications)
- Way to provide whole nodes for VOs

Drawbacks

- Virtualization penalties
 - CPU : seen down to 3% but needs tuning
 - Local I/O: typically seen 20-30% but do we really care ? (if a site buys disks which are 30% less performant, would the VOs realize ?)
 - Networking: little overhead in tests at CERN
 - CPU affinity helps
- Paravirtualization gives best performance results
- Stuff to be carefully considered:
 - Security, accounting, placing
 - Use of cgroups and ksm
- For virtualization, the batch system scalability is a source of concern

To be discussed

- Pure site business ?