

# USATLAS Virtualization Discussions

---

During the USATLAS Meeting on Virtualization and Configuration Management at BNL on June 15<sup>th</sup>, 2011 we formed a subgroup to discuss “Virtualization” use and plans within our infrastructure. Virtualization can refer to many different technologies and techniques used to abstract underlying hardware or software from the higher level application view. We choose to focus on two specific areas:

- **Service virtualization**
- **Worker-node virtualization**

## Service/Server Virtualization

We first discussed “Why” in the context of service/server virtualization. Many groups within USATLAS have either explored virtualization services or have actually virtualized some service hosts already. The following list provides the primary motivations for doing this:

- High-availability/resiliency
- Ease of backup
- Speed of creation/access
- Enabling much easier testing/upgrading
- Cloning capabilities
- Infrastructure consolidation
- Possibility of “turn-key” service appliances for small sites

In the context of “High-Availability” we discussed two forms: a) Service resiliency where virtualized instances of the service node are monitored by the infrastructure. If the service node crashes the infrastructure quickly and automatically starts a new instance (NOTE this doesn’t preserve any state stored in memory from the crashed service node), b) Application resiliency by running 2 virtualized instances in lock-stop on 2 different physical VM servers. Loss of one physical instance doesn’t impact the application.

Some important discussion surrounded many other advantages of service virtualization: ability to quickly proto-type service changes and upgrades (using cloning and quick provisioning), ease of backup (online) for VM instances, consolidation of physical infrastructure (less power, space, cooling and maintenance). It was also pointed out that a very important capability provided by service virtualization is to allow expert built and maintained VMs to be created and broadly deployed (Tier-3 context) without needing the required local expertise to deploy and configure such services at each local site.

## Worker-Node Virtualization

The discussion around worker-node virtualization pointed out the very different considerations which are involved compared to server/service virtualization. The primary goal for worker-node virtualization is to enable access to diverse resources. Virtualized worker-node VMs may run in the cloud or upon numerous different virtualization platforms. These worker-node VMs will allow VOs to tune, customize and test a specific environment to support their applications which can then be broadly run. The “Enterprise” features that are important for Server/Service virtualization are typically not relevant or required for this kind of virtualization. Instead the following are the important issues:

- Low cost (free is best)
- Need an infrastructure that can support large-scale instantiation of various VMs

- Need secure access to trusted VM image repositories
- VMs must be designed to allow easy site-specific configuration and tuning

The discussion of worker-node virtualization produced a number of questions that we felt needed addressing: How many jobs per VM are possible/desirable? How does a pilot system work with virtualized worker-nodes? Could produce a single VM generic enough to support many users/VOs? How do VMs interact with local schedulers (Condor/LSF/PBS...)? What kind of “test suite” is needed to validate ATLAS workflows in such VMs? There was significant discussion about how the PANDA pilot system might adopt to such a new infrastructure (using a 2 pilot system; one to instantiate the needed type of VM and one inside the VM to get the job payload).

Another discussion occurred about why, and to what extent, USATLAS would be interested in doing this. The group felt the primary goal was to allow us access to Clouds and other resources (Super-Computing centers) that would be otherwise unavailable to us. However there was also a strong belief that we should/would not convert most of our “dedicated” resources to support worker-node VMs because any such change (due to virtualization inefficiencies and VM instantiation and removal overhead) would result in a loss of resources.

## Summary

The discussions covered a broad range of topics and we feel we didn’t have time to explore many of the important questions in sufficient detail. In addition to the advantages discussed above we did come up with some possible down-sides for both types of virtualization:

- **Downsides for Server/Service Virtualization**
  - Server virtualization really needs many “Enterprise” features and these are expensive
  - Learning curve for enabling many of the powerful features are steep
- **Downsides for Worker-node Virtualization**
  - Assumed to be less efficient than physically installed, persistent configurations
  - Lots of possible overhead in managing, instantiating and removing VMs
  - Significant changes to existing infrastructures and tools are required to transition to this model.

It should be noted that the groups opinion was that the potential benefits of deploying appropriate virtualization techniques in our infrastructure significantly outweigh the possible downsides.

## Next Steps

We list two sets of next steps depending upon the type of virtualization:

### Server/Service Virtualization

- Create a reference repository for current best practices
  - Document what has been done
  - Provide diagrams
  - Provide enough detail to allow others to replicate setups
- Create a mailing list to allow users/administrators to share ideas and support one another
  - Begin threads on good ATLAS services to virtualize and any unresolved issues that should be addressed
- Provide pointers to recommended “How-to” guides for virtualization setup

- Benchmark impact of virtualization:
  - For the case of specific ATLAS services
  - In general, documenting overhead(s) incurred

## Worker-node Virtualization

- Identify which virtualization technologies are well match matched to worker-node virtualization
- Document existing worker-node virtualization details from sites/users who have already done this
- Provide a single location to gather pointers to relevant documentation, examples and “how-tos”
- Task a group to research, discuss and document answers to the following questions:
  - How many jobs per VM are possible, desirable?
  - How does a pilot system work with virtualized worker-nodes?
  - Could produce a single VM generic enough to support many users/VOs?
  - How do VMs interact with local schedulers (Condor/LSF/PBS...)?
  - What kind of “test suite” is needed to validate ATLAS workflows in such VMs?