Essay: HPC clusters and virtualization.

*Vergazov Artem, Advanced Computational Science, MSc 1st year.*

1. What are the differences and similarities between Zhores and the HPC Cluster @ Swinburne?

To begin with, the most obvious difference betweer these two clusters is their computational configurations. At first glance, Zhores cluster seems technically more sophisticated than the Swinburne cluster (<https://supercomputing.swin.edu.au/ozstar/>). Another difference that follows from the video is their workload manager systems. According to the video, they run PBS (Portable Batch System) as a queue system on Swinburne, meanwhile SLURM is used on Zhores for that purpose, because all the examples from the video were utilizing PBS. However, if you refer to the link above, it says that the latest and the only Swinburne’s supercomputer OzSTAR runs on SLURM Workload Manager. It also says that this supercomputer was installed in 2018 and the date of the video release is also 2018. So, my theory is the video relates to the previous Swinburne’s cluster Green II which has been decommissioned (<https://supercomputing.swin.edu.au/docs/1-getting_started/Swinburne-HPC.html>).

What comes to similarities between these clusters, from the user’s prospective there is almost no difference which Workload Manager is used for the queue system. The interaction scenario is every time the same: the user writes the shell script that performs the needed jobs, specifies the necessary configuration for this job (number of nodes, cores, amount of estimated time and RAM) and submits this script to the queue system which allocates resources and puts the output to the log file. In both clusters, this process is the same and does not depend on the Workload Manager (despite the different names of commands, like qsub and sbatch).

1. Are scripts generated by BYU job generator applicable to Zhores?

Yes, they are, because this generator just takes the configuration you enter, i. e. number of CPUs, GPUs, time, RAM, and other machine-independent common stuff and translates this information into “the SLURM language” which will run on Zhores too since Zhores uses SLURM as the Workload Manager. The only thing left to do is to write the body of the script – the job that is needed to be done in the script.

What are the array tasks, is it possible to setup array tasks on Zhores and what applications are suitable for that?

Array tasks are sets of “embarrassingly parallel” tasks with approximately the same number of cores and GPUs required. It is possible to setup them on Zhores too. Example of application for that: running Monte-Carlo simulations (for pi computation for example) with different RNG seeds in order to compare or average the results. For every such job the task is the same and requires the same amount of resources and only the RNG differs for each job.

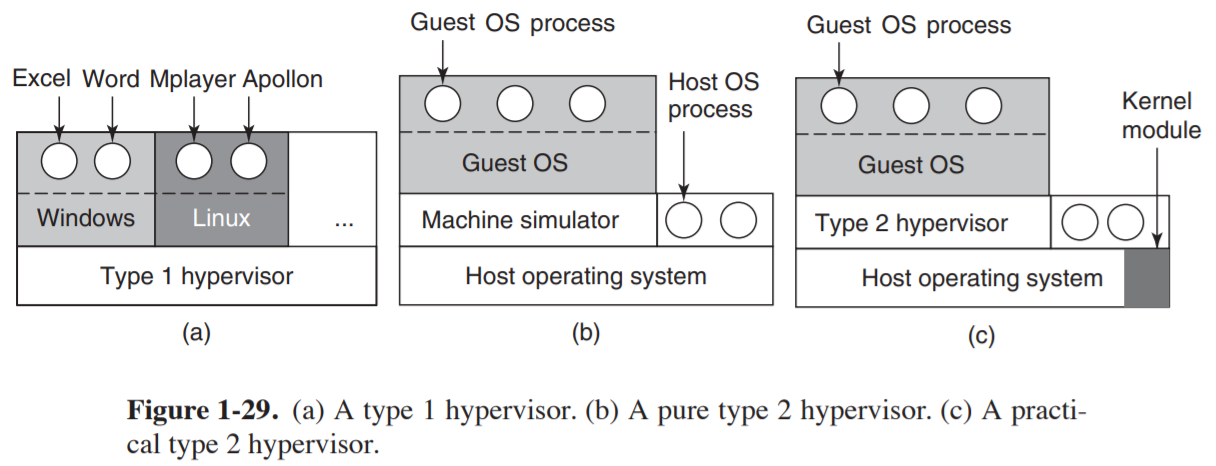
1. What containers can be run on Zhores and why?

You cannot run Docker containers on Zhores explicitly because of security issues: under specific circumstances one can get root rights and mess up the system. Instead, you have to convert your Docker container to a Singularity container and run it.

1. Is Docker a type 1 or type 2 virtual machine? Does it fulfill requirements for a VM?

Virtualization (also known as a hypervisor) allows a single computer to host multiple virtual machines, each potentially running a completely different operating system. Type 1 hypervisors run on the bare metal, and type 2 hypervisors may make use of all the services and abstractions offered by an underlying operating system.

Illustration from the book:



In practice, the real distinction between a type 1 virtual machine and a type 2 virtual machine is that a type 2 makes uses of a host operating system and its file system to create processes, store files, and so on. A type 1 hypervisor has no underlying support and must perform all these functions itself.

However, Docker is not a full VM, it is a container management tool. A VM can operate as a separate computer because each VM has its guest operating system above the host operating system, meanwhile Docker containers share the host operating system which makes them more lightweight but tied to the specific choice of the host OS which is usually UNIX-like. It is also important here that there is still variety of possible choices for the container OS because there are numerous Linux distributives that can be built on top of each other and share the same core.

That being said, Docker is not a full VM but behaves similarly to a type 2 VM, being more lightweight at the same time.