Virgo Cluster

M. Al-Turany on behalf of CIT

Disclaimer

What you see here is the result of the work of many people listed (alphabetically) here and others from the CIT

D. Bertini, M. Dessalvi, S. Fleischer, R. Grosso, C. Huhn, R. Karabowicz, D. Klein, J. Knedlik, T. Kollegger, D. Kresan, V. Penso, C. Preuß, K. Schwarz, C. Tacke, F. Uhlig

Current situation: GSI -CIT

- Very limited manpower
- Rapidly changing hardware and software environments
- Increasing and very different requirement from users
  - Pretty old software!
  - Newest packages and compilers
  - GPUs
  - ….

Constraints

- Cluster Maintenance: We need to be able to "quickly" re-install nodes on the cluster
- Even though separated from storage it needs to integrate with the storage (lustre)
- We must guarantee "Long-term stability" (e.g. distribution lifetime)

Our Users

- Mostly scientific users who need to develop and execute custom software
- Knowledge level varies very strongly:
  - Change parameter in script and submit
  - …
  - Develop/compile/submit
  - …
  - Sophisticated workflows (MPI, FairMQ, …)
- External users (e.g: ALICE grid)

Virgo Cluster: Design decisions

- **Virtualized** approach (**Singularity containers**).
- **Separate the Cluster from the rest of the infrastructure**.(i.e: Interactive machines, group servers, desktops, etc)
- **CentOS** as the host OS for the batch farm.
  - Prerequisite for OpenHPC
  - Better hardware (driver) support
  - Better compatibility with HEP community
- **SLURM** as resource management system.

Virtualized approach based on **Singularity containers**

- Decouple minimal host system installation from **V**irtualized **A**pplication **E**nvironment (**VAE**)
- Host system: Core system packages without application software
- VAE: Runtime environment + Application software (CVMFS)
- Few VAEs (<5) to support all users maintained by IT
- Custom user containers are not excluded but not supported in any form!

Design decisions: Software distribution

- **Spack** as a package manager for application software (https://spack.readthedocs.io)
- **CVMFS** as software distribution service inside and outside the GSI. (https://cernvm.cern.ch/portal/filesystem)

Software distribution and CVMFS

- User = Repository Owner := An experiment such as CBM, ALICE, . . .
- Each user gets a shared secret needed for publishing *(Each user is free to use Spack or whatever he likes to manage his software on CVMFS)*
- Each user can set up as many remote publishers as he needs (for example one per platform)
- It is the responsibility of the user to make sure that there are no unwanted collisions if there are several people/mechanisms publishing to their cvmfs repository

Software distribution and CVMFS

- CVMFS Procedures are documented at:

  https://git.gsi.de/dc/cvmfs-server/

- Each experiment have to decide for themselves whether to make each of their repositories public or internal.

Containerized environments:

Virtual Application Environments (VAE) is selected by login to specific submit nodes:

Submit node	VAE	Description
virgo-debian8.hpc.gsi.de	Debian 8	Compatible to Kronos
virgo-centos7.hpc.gsi.de	CentOS 7	

Kronos Compatibility: VAE mimicking the Kronos cluster

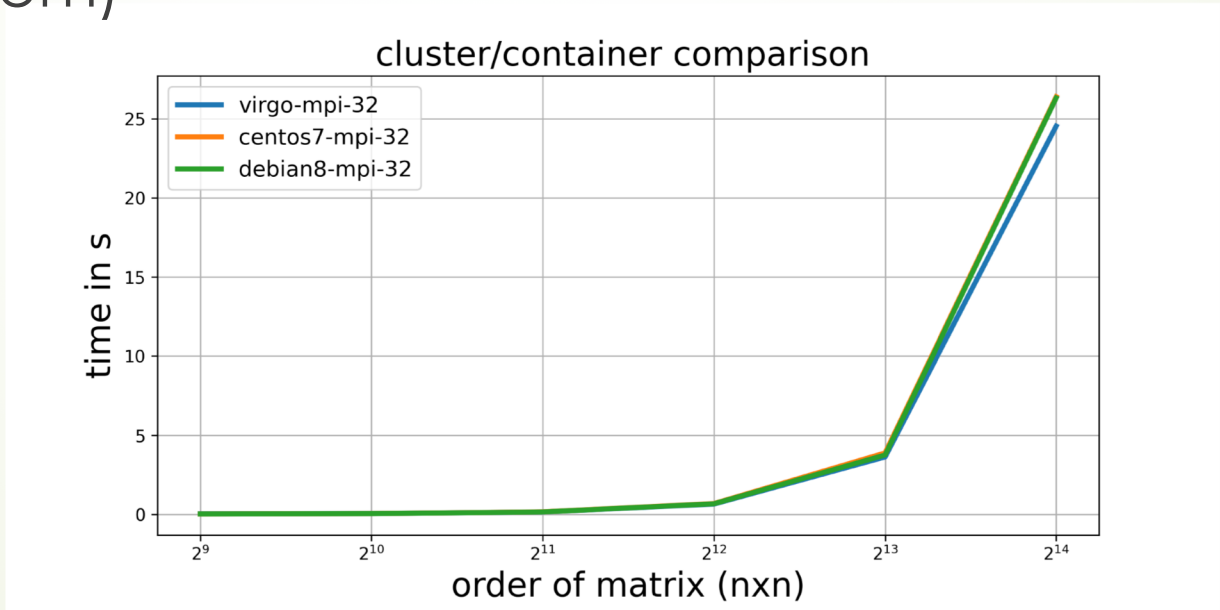
- Transition to the new cluster
- All software previously available via CVMFS on Kronos is mounted in VAE
- *Applications which run on Kronos should work* without modification.

Kronos Compatibility: OpenMPI test

- VAE successfully used for typical GSI TNSA Simulation
- ~ 3 days running time
- Physical results OK

Kronos Compatibility: MPI/ScaLapack tests

- Using AMD AOCL optimized BLAS(BLIS) via ScaLapack's pzgesv() (solve linear matrix with complex)
- Nearly identical behavior native/container (CPU-bound problem)



Dmytro Kresan
Florian Uhlig

Kronos Compatibility: ROOT/FairRoot Test

- Root/FairRoot-based applications run without any issues

Kronos Compatibility: VAE mimicking the Kronos cluster

This will be supported for limited and short time:

- **Debian 8 support ends this month**
  - **Kronos will be shut down as soon as all of you move to the Kronos VAE**

- **Kronos VAE should not stay much longer! (maximum until the end of this year)**

Software Installation Request Template

- Users should fill out the SIR template to request new software:

https://git.gsi.de/SDEGroup/SIR

User issues

- Virgo cluster and CVMFS
  - cluster-service@gsi.de
- Lustre
  - lustre-service@gsi.de

Future plans (very short term)

- Make compilers and basic scientific software available for CentOS7 on the new CVMFS (Next week)
- Shut down Kronos as soon as possible and integrate the hardware into Virgo

Documentation

- **CVMFS at GSI:** https://git.gsi.de/dc/cvmfs-server/
- **Virgo cluster**: https://hpc.gsi.de/virgo/

- **Spack:** https://spack.readthedocs.io
- **CVMFS:** https://cernvm.cern.ch/portal/filesystem