Wuppertal
Contributions and Experiences

T. Harenberg, M. Sandhoff, M. Vogel, C. Zeitnitz

What we do

- Group: Experimental Particle Physics (Wagner/Zeitnitz)
- Member of the ATLAS-Collaboration at the LHC
  - ATLAS data analysis
  - ATLAS detector operations
  - Detector upgrade for HL-LHC (Pixel detector)
  - Heavily involved in computing
- Operating an ATLAS-Tier-2 center for the last 10 years
  - Extensive experience in middleware (since D-GRID times), storage technologies and services
- Involved in computing operations and core software development for ATLAS
- Members involved in Computing
  - Torsten Harenberg: Tier-2/ATLAS Germany operations, chair of the GridKa technical advisory board
  - Marisa Sandhoff: Tier-2/ATLAS Germany operations
  - Marcelo Vogel: development and maintenance of core software for ATLAS
Postdoc in the Computing-Verbund
  - Christian Zeitnitz: PI in this project
  - Two PhD-students at CERN (Gentner program): data analysis framework, data management (RUCIO), storage simulation. Neither is directly involved in this project

Contributions to the Computing-Verbund

- Topic area A – starting now
  - Workpackage A1 and A3
    - Containerization of user jobs
    - Containerization of services (e.g. VOMS, database access, monitoring)
- Topic area B – partially in parallel to A1/A3
  - Workpackage B1
    - Testing of implemented container technologies in different systems (Tier-2 cluster, HPC, Cloud)
  - Workpackage B3
    - Testing the virtualization of user jobs
  - Workpackage B4
    - Integrated testing of all packages and technologies in different systems in terms of their reliability, scalability and general performance
    - The ease of installation and operation of these technologies will be of special interest

Experience

- GRID environment (Harenberg, Sandhoff)
  - Running an ATLAS Tier-2 center, inclusive a full set of services
  - Also running a local Tier-3 center shared by more than 10 groups from faculties of the university
  - Running (backup) grid-essential services for the whole of ATLAS
- Data management/storage (Harenberg, Sandhoff)
  - Extensive experience with a large variety of storage types and file systems
  - Developing and maintaining ATLAS-wide storage monitoring solutions
- Container technologies (next slides)

Container technologies

The Wuppertal group has been actively involved in the effort to deploy ATLAS software in containers since 2016. Our activities can be grouped into three major areas:

- Publishing and maintaining general documentation on uses of containers in ATLAS software. The main topics are:
  - Installation of Docker in laptops with tutorial describing basic operations (pulling, committing, tagging and pushing images)
  - How to use prebuilt images for running typical ATLAS production workflows
- Production and testing of standalone images for detector simulation jobs
- Integration of Docker into the ATLAS build system (nightlies)

Full ATLAS production releases in containers

- We currently produce and maintain Docker images with full releases of ATLAS production software
- Goal: to run production workflows without network connectivity, such as detector simulation
  - We install a full release from RPMs on a base image containing basic tools and libraries
  - We install and configure a standalone database package for detector conditions data
- The final image (~16GB) can be run natively in Docker. If pulled from Singularity it will be transformed automatically into a singularity image and run natively in Singularity
- The image was tested successfully at Theta (HPC at Argonne National Laboratory)

Containers in ATLAS Continuous Integration system

- Goal: to use hosts with newer OS versions to build ATLAS software RPMs
- The base image is loaded by the CI system and the build process continues in the container. Directories are mounted to the container to store output
- The Docker container runs in “detached mode” and exits when its “ENTRYPOINT” script ends. The script does three things:
  - Authenticates the local user (kerberos keytab imported from the host)
  - Downloads the scripts to build the rpms
  - Test the freshly built release in the container with a suite of production workflows. Network connectivity is necessary for access to the experiment’s databases. The test results are copied from the container to storage for monitoring
- After the container exits, rpms are copied to storage and later installed in an ATLAS distributed file system (CVMFS)

What is next?

- Get in contact with other groups involved in the work packages
  - First contacts established
- Survey of existing technologies and solutions for containerized services
  - Existing approaches will be the starting point for further developments
- Test out what is existing and adapt and improve for our applications