Status and Activities Topic Area B

Christian Zeitnitz

Bergische Universität Wuppertal

From the Proposal

- Topic Area B: Application and Test of virtualized Software Components in the Context of heterogeneous Computing Resources

| 1) Tests der Technologiekomponenten	2) Job- und Ressourcenmanagement
Implementierung und Tests auf verschiedenen Plattformen von • Speicher- und Cachinglösungen und • virtualisierter Dienste (Datenbanken, Monitoring, Accounting).	Jobverteilung und Überwachung in der Umgebung heterogener Computingressourcen unter Einbeziehung von Containervirtualisierung.
3) Virtualisierung von Nutzerjobs	4) Kombinierte Tests
• Erfassung der Anforderungen, • Bestimmung und Erzeugung der Laufzeitumgebung, • Erstellung des Containers und von Metadaten und • Checkpointing von Containervirtualisierung.	Testen von Gesamtsystemen (Speicher, Dienste, Ressourcemanagement) auf verschiedenen Plattformen in Bezug auf • Installations- und Wartungsaufwand, • Performance, • Skalierbarkeit und • Robustheit.

Status

- Positions in F, FR, M, W have been filled, or will be filled soon
  - Late hiring due to lack of suitable people
  - Overlap with personnel of TA A
  - In general a late start!
- Topic Area B relies on the development TA A in order to apply and test the software packages
  - Main activity of TA B will be in the second half of the project

Activities in the involved Institutes

- Aachen (Thomas Kress)
  - No direct funding for TA B
  - Interested to contribute, within the limited resources, to
    - Interaction between Tier centers and opportunistic resources, e.g. remote storage, protocols
    - Testing of remote functionalities together with other sites
- Frankfurt (Volker Lindenstruth)
  - Position will be filled at Nov. 1st
- Göttingen (Arnulf Quadt)
  - No funding for ErUM-IDT
  - Effort to integrate the enormous resources of the new Nordeutschen Hochleistungsrechner (HLRN, https://support.hlrn.de/home/view/Main/WebHome) into WLCG and make these accessible for particle physics
  - This activities was unfortunately not funded within ErUM-IDT

Status and Planned contributions in Work Area B from Freiburg
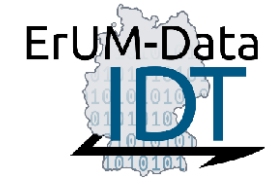
M.Schumacher

Not really started. Plans unchanged. Position only filled 1st September.

WP2: Management of jobs and resources

➢ optimise parameters for job orchestration for different combinations of heterogeneous resources (WLCG, HPC, clouds, …; permanent and short time availability) based on monitoring and unified queues developed in WP1 (1st for ATLAS use case)

- ▪ deployment and transfer of VMSs and Containers to CPU nodes
- ▪ handling of VM/Container and Jobs meta data
 also/in particular for resources which are not available as long as expected

WP4: Combined tests

➢ test of complete workflow based on tools for monitoring, benchmarking and accounting developed in work area A on complex systems

- ▪ evaluation and optimisation of performance
- ▪ investigation of reliability, scalability and maintenance of solutions

CERN

- Markus Elsing, Mario Lassnig, Thomas Beermann, …
- Distributed Scientific Data Management Service
- Open Source
- Broad Community
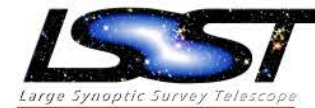
- Ongoing developments
  - o Optimization of disk space usage (replicas, data/tape carousel)
  - o Multi-VO support
  - o „Data embargo" – requested by astro community
  - o Direct integration of research databases (relation data and publications)
- Projects with commercial cloud providers (Google, OCRE)
  - o Topic: cost control
- Ideas for the future
  - o Current model of static storage assignment to projects will be inefficient in case of big data centers
  - o Possible solution: Dynamic storage and data flow. Requires intercommunication between RUCIO instances

Status and Activities in Wuppertal

- Involved people
  - Marcello Vogel
  - (Thomas Beermann (RUCIO) – not funded through ErUM-IDT)
  - Christian Zeitnitz

- Currently addressed projects
  - Containerization of jobs (simulation, reconstruction, analysis)
  - Database extraction for network-less jobs
    - Determine from metadata the required databases and entries
    - Extract data to SQLITE files
    - Pack job and DB files into container
  - Current status
    - Use case ATLAS
    - Successful for simple case of simulation
    - Reconstruction nearly working, but obstacles because of missing meta data and hardcoded paths and DB access
  - Next step
    - Generalization to other experiments
    - Define interfaces