# Federated infrastructures and cloud computing

## Organisation and preparation of BMBF proposal

**Alexander Schmidt, RWTH Aachen, 21.09.2020**

# 2017-2021: IDT-UM

- Task areas A, B, C, D

- Fed. infrastuctures, cloud computing are in task areas A+B

- Project ends 30.09.2021

- We need to plan for seamless continuation and hope for support from BMBF

## Innovative Digitale Technologien für die Erforschung von Universum und Materie

**Gemeinsamer Antrag von Gruppen aus den Bereichen Elementarteilchenphysik, Hadronen- und Kernphysik und Astroteilchenphysik**

- Rheinisch-Westfälische Technische Hochschule Aachen, Prof. Dr. Martin Erdmann
- Rheinische Friedrich-Wilhelms-Universität Bonn, PD Dr. Philip Bechtle
- Friedrich-Alexander-Universität Erlangen-Nürnberg, Prof. Dr. Gisela Anton
- Goethe Universität Frankfurt am Main, Prof. Dr. Volker Lindenstruth
- Albert-Ludwigs-Universität Freiburg, Prof. Dr. Markus Schumacher
- Georg-August-Universität Göttingen, Prof. Dr. Arnulf Quadt
- Universität Hamburg, Jun.-Prof. Dr. Gregor Kasieczka
- Karlsruher Institut für Technologie, Prof. Dr. Günter Quast
- Johannes Gutenberg-Universität Mainz, Prof. Dr. Volker Büscher
- Ludwig-Maximilians-Universität München, Prof. Dr. Thomas Kuhr
- Bergische Universität Wuppertal, Prof. Dr. Christian Zeitnitz

Assoziierte Partner sind

- CERN, Dr. Markus Elsing
- DESY, Dr. Volker Gülzow
- GridKa, Dr. Andreas Heiss
- GSI Helmholtzzentrum für Schwerionenforschung, Darmstadt, Dr. Kilian Schwarz
- Forschungszentrum Jülich, Dr. Elisabetta Prencipe
- Westfälische Wilhelms-Universität Münster, PD Dr. Christian Klein-Bösing

# 2020: ErUM-Data?

- ErUM-Data "action plan" expected later this year

- However, we were "suggested" to submit the new proposals anyway on **1. November** along with the other ATLAS/CMS proposals in ErUM-Pro

- The proposals could possibly be moved into ErUM-Data at a later time

- This was similar in 2017

# Two or more proposals

- There will be two separate proposals ("Verbünde") for ErUM-Data related work:

Federated infrastructures and cloud computing

Big data analytics

- There will be yet another proposal for experiment-specific computing infrastructure investments (Tier-2 hardware), and related personnel (by ATLAS/ CMS together)

# Communication

- New mailing list was necessary:
  federated-infrastructures-proposal-2020@lists.rwth-aachen.de

- Self-subscribe at
  https://lists.rwth-aachen.de/postorius/lists/federated-infrastructures-proposal-2020.lists.rwth-aachen.de/

- All communication will happen there, so you need to subscribe

- We will need a couple more meetings to coordinate the details

# Who does what?

- Coordination of the content of the new proposal has started

- **Tables with "topic areas" and interested partners on google doc:**
  https://docs.google.com/document/d/1SFdqMC5brzzg1X0cxgbqgejeTOCseymIoTqT7QI-2Rs/edit?usp=sharing

- A (still rough) draft of the general part of the proposal in overleaf:
  https://www.overleaf.com/read/ynwmhrryzyhs

- We still need some more "responsible editors" for some of the topic sections, I will contact some of you for this. **Please volunteer**, it's not an extreme amount of work.

We don't need to reinvent everything. But some adjustments are necessary for the new proposal.

# Suggested schedule

- Political boundary conditions are half-clear, so start writing now :-)

- Until end of September: draft of the general part of the application ready

- Until middle of October: fine-tuning of the details

- Middle October: Submit proposal to University administrations

- End of October: Submit proposal to BMBF

# Investitionsmittel

- the LHC Tier-computing resources for Run 3 are applied for in a separate proposal

- Do we need/want Hardware also in the ErUM-Data cloud computing proposal? Some thoughts on this are:

  - We develop models to efficiently operate and use computing infrastructure (hardware) so we need to **test** this on hardware

  - For LHC, we could use existing Tier-computing hardware, however these are officially "pledged" to the experiments

  - We could also use smaller scale HPC center resources (such as University computing clusters) for testing purposes

  - Astroparticle physics does not have any LHC-like Tier computing

some more details
(partially based on slides of C. Zeitnitz)

# Structure

- We are developing **the future** contributions from Germany to computing in different fields in the research of the structure of matter

- Community-overarching
  - High energy (specifically High-Luminosity LHC, from 2028)
  - Astroparticle
  - Hadron and nuclear physics (FAIR)

# Structure

- because of the gigantic resource needs, the future model will be centered around around large computing centers
  - specifically for mass storage (**data lakes**)

- CPU can also be available in a distributed way across several locations
  - unified,transparent access to CPU and storage necessary

- need to take into account
  - heterogeneous resources
  - temporarily available resources
  - opportunistic/parasitic usage of resources
  - various science clouds and also commercial providers

# Structure

- **LHC Run-3**
  - computing as now
  - Tier-2 at Universities deliver service and resources
  - financing is not yet clear, hope for ErUM-Data

- **LHC Run-4**
  - new models with data lakes at big centers
  - Universities can be "federated" and play an important role
  - opportunistic contributions should be feasible also temporarily from "any" resource

# Long term schedule

- Computing frameworks for LHC Run 4 (2028) being developed in parallel to Run 3 (2022 - 2024)

- Larger test installations will be necessary from 2024
  - combined test under realistic conditions with federated centers

- Resources for Run 4 need to be fully ready in 2028
  - Installation of resources needs to start 2025/26

- Finance resources for this need to be clarified ~now