GRID-based off-line infrastructure for the PADME experiment at the DAFNE BTF

Alessandro De Salvo[1], Georgi Stefanov Georgiev[2], Emanuele Leonardi[1], Francesco Safai Tehrani[1], Elisabetta Vilucchi[3]

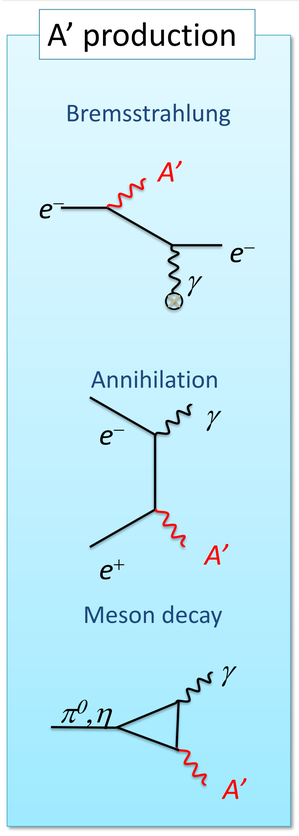
[1]INFN Roma, [2]University of Sofia, [3]INFN LNF

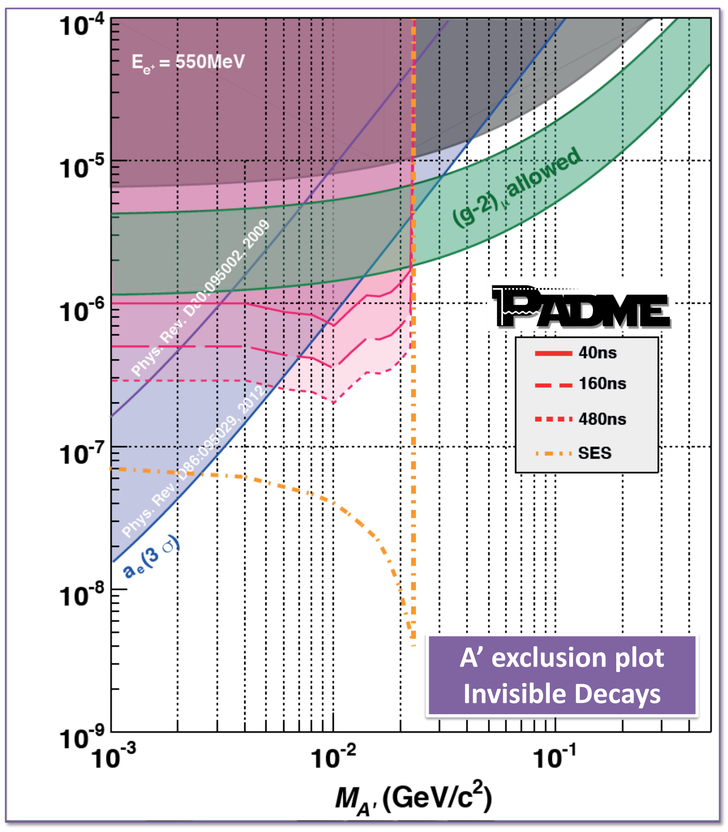
Supported by MAECI Progetto Grande Rilevanza Italia-USA PGR226

The Dark Photon

A possible solution to the Dark Matter problem is the hypothesis that it interacts with the SM gauge fields only via "portals" which link our world to the "dark" world. The simplest model postulates the existence of a U(1) symmetry with its corresponding A' vector boson: SM particles are neutral under this symmetry, so that this field, thanks to the possible mixing with the photon, would couple to the SM only with an effective charge εe. For this reason it is often called «dark photon».

A' production





The PADME Experiment

PADME (Positron Annihilation into Dark Matter Experiment), approved by INFN in 2015, will look for invisible A' decays using the Beam Test Facility (BTF) line of the Linear Accelerator at the INFN Frascati National Laboratory (LNF).

A 550 MeV e⁺ beam will impinge on a thin (100μm) diamond target: the experiment will detect the production of dark photons by measuring the missing mass of the $e^+e^- \rightarrow A'\gamma$ final state.

The experiment is starting its preliminary data taking in July 2018.

Data Production

In the first year of run PADME will collect a total of 10^{13} e⁺ on target corresponding to $O(10^9)$ spills. DAQ consists in a total of 900 ADC channels, for a total DAQ rate of ~80MB/s, and will produce ~300TB of raw data which will include all ADC samples (after applying zero suppression). Data will be written to a 40TB local RAID disk buffer, then moved to the WLCG Tier2 in LNF, and copied to the tape library at INFN-CNAF (Bologna). An emergency copy of the data will also be kept on the KLOE2 tape library at LNF. The local disk buffer allows data taking even in absence of a network connection, up to a maximum of ~2 weeks.

MonteCarlo Production

The full layout of the experiment has been modeled with the GEANT4 simulation package. Simulating the interaction of a full e⁺ spill with the experiment takes ~5sec including the digitization step (measured on a not-so-recent Xeon E5-4610 v2 2.3GHz CPU core). We foresee a total production of at least twice the DAQ statistics (i.e. 2 x 10^{13} e⁺ on target) using a total CPU power of 3 kHEPSpec over 1 year and producing 80TB of MC data.

Event Reconstruction

Both DAQ and MC events will be fully reconstructed. The reconstruction procedure is still at a preliminary stage, so it is not possible to have a precise estimate of the CPU needs and the total amount of storage produced. Given the relatively simple structure of the experiment and the fact that reconstructed events will not include the full set of ADC samples, we expect the CPU needs to be a fraction of that for MC production, and storage to be roughly 1/10th of the raw data.

GRID infrastructure

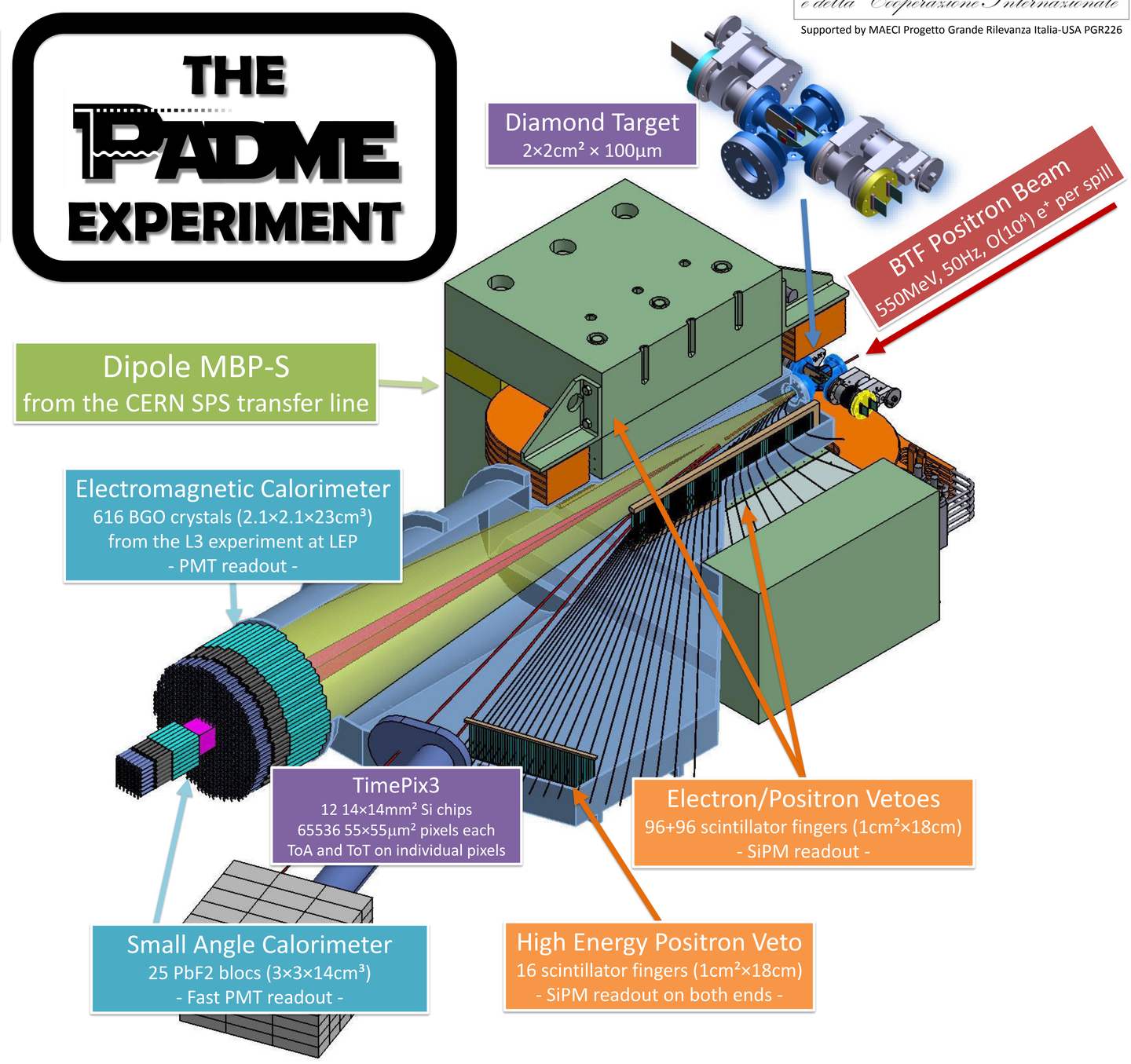
To support all the processing and storage needs of the experiment we developed a GRID-based infrastructure. This allowed access to distributed resources with a homogeneous interface. Resourced used included:

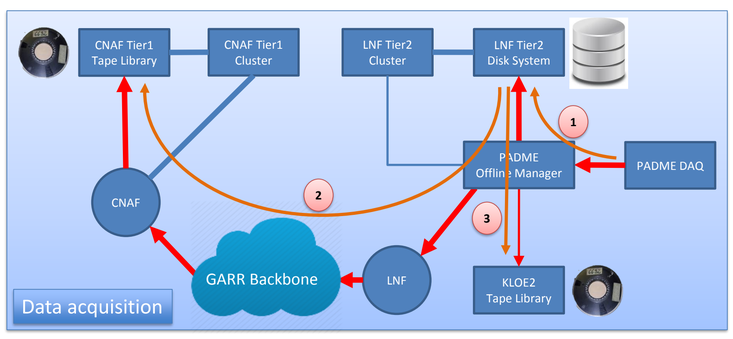
- CPU farms in INFN-LNF, INFN-CNAF, and Sofia University
- Disk storage systems in INFN-LNF and INFN-CNAF
- Tape libraries in INFN-CNAF and INFN-LNF

All data (Raw, MC, Reco) will be saved to tape. All Reco data will be kept on disk for data analysis, while MC and Raw data will reside on disk only for the time strictly needed to reconstruct them.

PADME software is distributed over the GRID using CVMFS (hosted at INFN-CNAF).

THE PADME EXPERIMENT

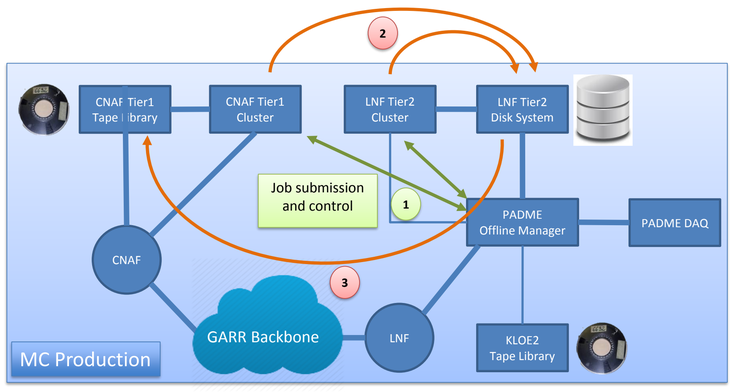




Data coming from the PADME DAQ system are initially copied (pull) to the INFN-LNF Disk System (1) and then replicated to the INFN-CNAF tape library (2) and to the KLOE2 tape library at INFN-LNF (3).

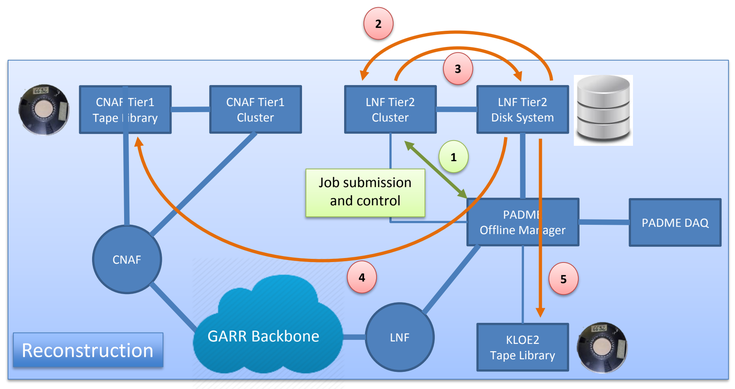
GFAL2 tools are used for data transfer jobs.

N.B. the PADME DAQ system and the KLOE2 tape library are not "grid-ified"→ use **GFAL2 SSH** interface.



MC production jobs are submitted from the PADME Offline Manager, configured as a GRID User Interface, to an available GRID site (1). All produced data are copied to the INFN-LNF Disk System for reconstruction (2) and then replicated to the INFN-CNAF tape library (3).

N.B. Given the small number of production sites, we do not use a Resource Broker.



Reconstruction jobs are submitted from the PADME Offline Manager to the INFN-LNF Tier2 cluster (1). Raw/MC data are read from the INFN-LNF Disk System (2), Reco data are written back to it (3) and then replicated to INFN-CNAF and KLOE2 tape libraries (4,5). When the reconstruction is over, Raw/MC data are removed from the Disk System.

N.B. All Reconstruction jobs are currently run at INFN-LNF as they rely on a calibration DB not visible from outside.

Conclusions

- In terms of available man power, PADME is a small experiment, but the amount of data produced and the computing needs are not negligible and will benefit from a GRID-based infrastructure.
- The initial learning curve of the GRID is quite steep: information is somewhat scattered around and one has to rely on help from GRID site managers and knowledgeable people from other (mainly LHC) experiments.
  - A "GRID 101" (aka "**GRID for Dummies**") manual would greatly improve the situation.
- A few basic GRID-related services (file manager, job submission manager, data transfer manager) were missing or were too complex to be used in a low manpower environment. A simple version of these services was therefore implemented within the collaboration. As all these services are very likely needed by any experiment, with very small variations, creating easily customizable standard versions of these packages should not be difficult and would greatly reduce the time for the initial GRID-ification of an experiment.

REFERENCES

- Raggi M, Kozhuharov V. *Results and perspectives in dark photon physics.* Riv. Nuovo Cim. 38 (2015) n.10, 449-505.
- Raggi M, Kozhuharov V, Valente P. *The PADME experiment at LNF.* EPJ Web of Conferences 96, 01025 (2015).
- Leonardi E, Raggi M, Valente P. *Development and test of a DRS4-based DAQ system for the PADME experiment at the DAΦNE BTF.* J. Phys.: Conf. Ser. 898 032024 (2017)
- Leonardi E, Kozhuharov V, Raggi M, Valente P. *GEANT4-based full simulation of the PADME experiment at the DAΦNE BTF.* J. Phys.: Conf. Ser. 898 042025 (2017).
- PADME Experiment Home Page: http://www.lnf.infn.it/acceleratori/padme/

CHEP 2018 — 23RD INTERNATIONAL CONFERENCE ON COMPUTING IN HIGH ENERGY AND NUCLEAR PHYSICS — 9-13 July 2018, National Palace of Culture, Sofia, Bulgaria