

XENON computing model

M. Selvi

INFN - Sezione di Bologna

E-mail: marco.selvi@bo.infn.it

Abstract. The XENON project is dedicated to the direct search of dark matter at LNGS. XENON1T was the largest double-phase TPC ever built and operated so far, with 2 t of active xenon, decommissioned in December 2018. It successfully set the best world-wide limit to the interaction cross-section of WIMPs with nucleons. In the context of rare event search detectors, the amount of data (in the form of raw waveform) was significant: order of 1 PB/year, including both Science and Calibration runs. The next phase of the experiment, XENONnT, is under construction at LNGS, with a 3 times larger TPC and correspondingly increased data rate. Its commissioning is foreseen by the end of 2019. We describe the computing model of the XENON project, with details of the data transfer and management, the massive raw data processing, and the production of Monte Carlo simulation. All these topics are addressed using in the most efficient way the computing resources spread mainly in the US and EU, thanks to the OSG and EGI facilities, including those available at CNAF.

1. The XENON project

The matter composition of the universe has been a debate topic among scientists for centuries. In the last couple of decades a series of astronomical and astrophysical measurements have corroborated the hypothesis that ordinary matter e.g. electrons, quarks, neutrinos, etc. represents only 15% of the total matter in the universe. The remaining 85% is thought to be made of a new, yet-undiscovered exotic species of elementary particles called dark matter. These indirect evidences of its existence triggered a world-wide effort to try observe its interaction with ordinary matter in extremely sensitive detectors, but its nature is still a mystery. The XENON experimental program [1, 2, 3, 4] is searching for weakly interacting massive particles (WIMPs), hypothetical particles that, if existing, could account for dark matter and that might interact with ordinary matter through nuclear recoil. XENON1T is the third generation of the experimental program; it completed the data taking at the end of 2018, setting the best world-wide limit to the interaction cross-section of WIMPs with nucleons. The experiment employs a dual-phase (liquid-gas) xenon time projection chamber (TPC) featuring as target for WIMPs two tonnes of ultrapure liquid xenon. The detector is designed in such a way to be sensitive to rare nuclear recoils of xenon nuclei possibly induced by WIMPs scattering within the detector. The TPC is surrounded by a water-based muon veto (MV). Each sub-detector is read out by its own data acquisition system (DAQ). The detector is located underground at the INFN Laboratori Nazionali del Gran Sasso in Italy to shield the experiment from cosmic rays.

XENON1T is an order of magnitude larger than any of its predecessor experiments. This upscaling in detector size produced a proportional increase in the data rate and computing needs of the collaboration. The size of the data set required the collaboration to transition from a centralized computing model, i.e. the entire dataset is stored on a local facility at

Figure 1. Overview of the XENON1T Job and Data Management Scheme.

various institutions, to having to distribute the data across collaboration resources. Similarly, the computing requirements called for incorporating distributed resources, such as the Open Science Grid (OSG) [5] and the European Grid Infrastructure (EGI) [6], for main computing tasks, e.g. initial data processing and Monte Carlo production.

2. XENON1T

For what concern the data flow, the XENON1T experiment uses a DAQ machine hosted in the XENON1T service building underground to acquire data. The DAQ rate in DM mode is 1.3 TB/day, while in calibration mode it can be significantly larger: up to ~ 13 TB/day.

A significant challenge for the collaboration has been that there is no single institution that has the capacity to store the entire data set. This requires the data to either be stored in a cloud environment or be distributed across various collaboration institutions. Storing the data in a cloud environment is prohibitively expensive at this point. The data set size and the network traffic charges would consume the entire computing budget several times over. The only feasible option was to distribute the data across several computing facilities associated with collaboration institutions.

The raw data are copied into *Rucio*, a data handling system. There are several Rucio endpoints or Rucio storage elements (RSE) around the world, including LNGS, NIKHEF, Lyon and Chicago. The raw data are replicated in at least two positions and there are two mirrored tape backups, at CNAF and in Stockholm, with 5.6 PB in total. When the data have to be processed, they are first copied onto Chicago storage then they are processed using the OSG. The processed data are then copied back to Chicago and become available for the analysis. In addition, for each user there is a home space of 100 GB available on a disk of 10 TB. A dedicated server will take care of the data transfer to/from remote facilities. A high memory 32 cores machine is used to host several virtual machines, each one running a dedicated service: code (data processing and Monte Carlo) and documents repository on SVN/GIT, the run database, the on-line monitoring web interface, the XENON wiki and GRID UI. In fig. 1 we show a sketch of the XENON computing model and data management scheme.

The resources at CNAF (CPU and Disk) are used so far mainly for the Monte Carlo simulation of the detector (GEANT4 model of the detector and waveform generator), and for the real-data storage and processing. Currently we used about XX TB of the XX TB available for 2018. There were some improvements performed recently by the Computing Working group of the experiment. The CNAF Disk at the beginning was not integrated into the Rucio framework because it was not large enough to justify the amount of work needed for the integration (it was 60 TB up to 2016). For this reason we required for 2018 an additional amount of 90 TB, to reach a total 200 TB which is considered significant by the collaboration to consider a full integration of the Disk space.

The second improvement has been to perform the data processing on both the US and EU GRID (previously it was done in the US only). Some software tools have been successfully developed and tested during 2017, and they are used for a fully distributed massive data processing. To fulfil this goal, we required 300 HS06 additional CPUs, for a total of 1000 HS06, equivalent to the resources available on the US OSG.

The request of Tapes (1000 TB) in 2018 was done to fulfil the requirement by INFN to have a copy of all the XENON1T data in Italy, as discussed inside the INFN Astroparticle Committee. A dedicate automatic data transfer to tapes has been developed by CNAF.

The computing model described in this report allowed for a fast and effective analysis of the first XENON1T data in 2017, and the final ones in 2018, which lead to the best limit in the search of WIMPs so far [7, 4].

3. XENONnT

The planning and initial implementation of the data and job management for the next generation experiment, XENONnT, has already begun. The experiment is currently under construction at LNGS, and it's scheduled to start taking data by the end of 2019. The current plan is to increase the TPC volume by a factor 3 to have 6 t of active liquid xenon. The new experimental setup will also have an additional veto layer called Neutron Veto. The larger detector will require modifications to the current data and job management. The processing chain and its products will undergo significant changes. The larger data volume and improved knowledge about data access patterns has informed changes to the data organization. Rather than store the full raw dataset for later re-processing, the data coming from the detector will be filtered to only include interesting events. The full raw dataset will only be stored on tape at one or two sites, where one of these sites is for long-term archival. The filtered raw dataset will be stored at OSG/EGI sites for later reprocessing. The overall data volume of the reduced dataset will be similar to the current data volume of XENON1T.

4. References

- [1] Aprile E. et al (XENON Collaboration), *Dark Matter Results from 225 Live Days of XENON100 Data*, 2012, Phys. Rev. Lett. **109**, 181301
- [2] Aprile E. et al (XENON Collaboration), *Physics reach of the XENON1T dark matter experiment*, 2016, JCAP **04**, 027
- [3] Aprile E. et al (XENON Collaboration), *The XENON1T Dark Matter Experiment*, Eur. Phys. J. C77 **12**, 881 (2017)
- [4] Aprile E. et al (XENON Collaboration), *Dark Matter Search Results from a One Ton-Year Exposure of XENON1T*, 2018, Phys. Rev. Lett. **121**, 111302
- [5] Ruth Pordes et al., *The open science grid*, Journal of Physics: Conference Series 78, 1 (2007), 012057.
- [6] D. Kranzlmuller et al., *The European Grid Initiative (EGI)*, Remote Instrumentation and Virtual Laboratories. Springer US, Boston, MA, 6166 (2010).
- [7] Aprile E. et al (XENON Collaboration), *First Dark Matter Search Results from the XENON1T Experiment*, 2017, Phys. Rev. Lett. **119**, 181301