ASTERICS - H2020 - 653477

Second main thematic training event

ASTERICS GA DELIVERABLE: D3.7

Document identifier:	ASTERICS-D3.7.docx
Date:	**25 August 2017**
Work Package:	**WP3 OBELICS**
Lead Partner:	**LAPP**
Document Status:	**Report**
Dissemination level:	Public
Document Link:	https://www.asterics2020.eu/documents Second main thematic training event

Abstract

This report gives an overview of the second main thematic event organized by WP3-OBELICS entitled “First ASTERICS-OBELICS International School on Advanced software programming for astrophysics and astroparticle physics”. This event was the deliverable D3.7 under the grant agreement. This report discusses the overall outcome and feedback collected from students, tutors and organizers that will be useful to organize the next thematic training event.

I. COPYRIGHT NOTICE

II. DELIVERY SLIP

	Name	Partner/WP	Date
From	Jayesh Wagh	LAPP	25-08-2017
Author(s)	Wagh, Poireau, Lamanna	LAPP	25-08-2017
Reviewed by	Giuseppe Cimò	JIVE/ASTRON	08-08-2017
Approved by	Cimò		20-09-2017

III. DOCUMENT LOG

Issue	Date	Comment	Author/Partner
1	07-08-2017	First draft	Jayesh Wagh, LAPP
2	24-08-2017	Second draft	Vincent Poireau, LAPP
3	25-08-2017	Third draft	Giovanni Lamanna, LAPP

IV. APPLICATION AREA

This document is a formal deliverable for the GA of the project, applicable to all members of the ASTERICS project, beneficiaries and third parties, as well as its collaborating projects.

V. PROJECT SUMMARY

ASTERICS (Astronomy ESFRI & Research Infrastructure Cluster) aims to address the cross-cutting synergies and common challenges shared by the various Astronomy ESFRI facilities (SKA, CTA, KM3Net & E-ELT). It brings together for the first time, the astronomy, astrophysics and particle astrophysics communities, in addition to other related research infrastructures. The major objectives of ASTERICS are to support and accelerate the implementation of the ESFRI telescopes, to enhance their performance beyond the current state-of-the-art, and to see them interoperate as an integrated, multi-wavelength and multi-messenger facility. An important focal point is the management, processing and scientific exploitation of the huge datasets the ESFRI facilities will generate. ASTERICS will seek solutions to these problems outside of the traditional channels by directly engaging and collaborating with industry and specialised SMEs. The various ESFRI pathfinders and precursors will present the perfect proving ground for new methodologies and prototype systems. In addition, ASTERICS will enable astronomers from across the member states to have broad access to the reduced data products of the ESFRI telescopes via a seamless interface to the Virtual Observatory framework. This will massively increase the scientific impact of the telescopes, and greatly encourage use (and re-use) of the data in new and novel ways, typically not foreseen in the original proposals. By demonstrating cross-facility synchronicity, and by harmonising various policy aspects, ASTERICS will realise a distributed and interoperable approach that ushers in a new multi-messenger era for astronomy. Through an active dissemination programme, including direct engagement with all relevant stakeholders, and via the development of citizen scientist mass participation experiments, ASTERICS has the ambition to be a flagship for the scientific, industrial and societal impact ESFRI projects can deliver.

VI. EXECUTIVE SUMMARY

The first ASTERICS-OBELICS International School on Advanced software programming for astrophysics and astroparticle physics was organized from 6-9 June 2017 in Annecy, France. This school was organized following the brainstorming that took place at the first ASTERICS-OBELICS workshop, where the need to train young and senior researchers in the field of astronomy and astrophysics was discussed. This school was organised and hosted by the CNRS LAPP (Laboratoire d'Annecy-le-Vieux de Physique de Particules) laboratory. This four-day training event brought together over 80 PhD students, postdocs, senior researchers from the domain of astrophysics and astroparticle physics, and gravitational waves community from renowned research institutes in Europe as well as USA. The school provided theoretical and

hands-on training to acquire efficient and fast computer programming techniques, as well as skills for improving scientific data analysis software. Thanks to the expertise shared by the scientific committee as well as the tutors, the school successfully addressed the programming needs of the astrophysics community .

VII. Table of contents

1. Introduction

During the 1st ASTERICS-OBELICS workshop in 2016, the need to train the astrophysics scientists in advanced software programming was extensively discussed. Following this feedback, the second thematic WP3 training event was dedicated to advanced software programming for astrophysics and astroparticle physics.

An international programme committee comprising of members from various research institutions and ESFRI projects was formed to identify the list of subjects for theory and hands-on sessions to be included in the school programme. Considering the feedback from the international programme committee, Python was preferred as the programming language for the hands-on sessions.

The school announcement was widely disseminated to H2020-ASTERICS members, ESFRI projects as well as European astronomical societies in the month of February 2017 and March 2017.Task 3.1 MAUD members along with the LAPP staff managed the organizational arrangements for the school participants. As for the registrations, we attained maximum number of registration requests a week before the registration deadline.

The school was organized and hosted by the Laboratoire d'Annecy de Physique de Particules in Annecy, a renowned touristic location in France. This four-day training event brought together over 80 PhD students, postdocs, senior researchers from the domain of astrophysics and astroparticle physics, and gravitational waves community from renowned research institutes in Europe as well as USA.

The format of the school (section 3) included theory sessions as well as hands-on sessions so that participants with diverse background could have sufficient theoretical knowledge to participate in the hands-on sessions.

Post-event feedback from the participants (section 4. Feedback) has convinced us about the sheer success of the event in terms of organization as well as scientific programme. There is still some room for further improvement in the scientific programme and we aim to further improve on the suggestions from participants.

2. Organizational Approach

While organizing an IT-training event, organizational approach plays a key role. We established two committees to ensure timely execution of various stages of organization. An international programme committee was set up comprising of members from H2020-ASTERICS as well as ESFRI projects and PRACE. A local organization committee comprising of staff members from LAPP and WP3-OBELICS members at LAPP.

2.1 Venue & Duration

The event was organized from 6-9 June 2017 in Annecy-le-Vieux, a picturesque, and usually sunny, location in the month of June.. The second week of June was selected because in the first week another international event "GWPAW"(https://indico.in2p3.fr/event/13954/) was organized by LAPP in Annecy. In the week of 12-17 June LSST collaboration meeting was organized in Lyon. The dates were selected so that the participants from those two events could also attend the ASTERICS-OBELICS school.

2.2 Dissemination

For dissemination purpose, we relied on email communications as well as social media dissemination. The first email communication was sent out in the month of February 2017. We requested ASTERICS members to disseminate the school announcement within their respective institutions, ESFRI projects as well as contacts in the national astronomical societies.

2.3 Registration

We opened the school registration in the Month of February 2017. Interested PhD students and researchers were able to confirm their participation through online registration and payment of €150 of registration fee per participant on CNRS-Azure Colloque portal. The comprehensive registration fee covered all the following charges during the school

- Course fees (all modules): 6-9 June 2017.
- Registration kits.
- Shared accommodation (5 nights) at Jean 23 guesthouse.
- Lunch: 6-9 June 2017.
- Coffee breaks: 6-9 June 2017.
- Social Events

For 10 meritorious applicants we have provided scholarship in the form of full registration fee waiver. These applicants were selected by the OBELICS members at LAPP.

3. Scientific Programme

The school programme was based on three modules, addressing efficient code writing, parallel and GPU programming for Python and Python libraries for astronomy and astrophysics. Python was the language of reference during the school. All the three modules included a combination of theory and hands-on sessions. For the hands-on sessions, the participants were divided in two groups, beginner group and intermediate/advanced group. All the presentations and course material is made publicly available on the school website and it is free to download.

Module 1: efficient code writing

This module included a combination of theory sessions (5hours) and hands-on sessions (1hour 30 minutes) addressing following subtopics:

- Good code practice, traps to avoid
- Project management/organisation, continuous integration

- Profiling & debugging

Module 2: parallel & GPU programming for Python

This module included a combination of theory sessions (4 hours) and hands-on sessions (5 hours 25 minutes) addressing following subtopics

- Learn how to accelerate your code to be executed on several CPUs or on GPU cards
- Hands-on sessions with real exercise in physics problematics

Module 3: Python libraries for astronomy & astrophysics

This module included a combination of theory sessions (3hours 15 minutes) and hands-on sessions (5 hours 45 minutes) addressing following subtopics

- Learn about libraries such as numpy, pandas, astropy, pyVO.

3.1 List of tutors

In addition to identifying school modules, the international programme committee also contributed to suggestions of the tutors for the school from their networks. PRACE (Partnership for Advanced Computing in Europe) members have shared their expertise for the organization of the school. Full list of tutors with their brief bio is provided in Appendix 1.

3.2 Course content

In order for tutors to develop the course material, it was important for them to understand the level of understanding of registered participants. We had sent out a survey to school participants to understand their knowledge level for all the modules and submodules. A summary of these inputs was shared with tutors. For hands-on sessions, we had separated the participants in two groups beginners and advanced groups with exercises adapted to their level.

3.3 Infrastructure

All the participants had their own laptop. A list of recommendations on tools to be installed on the laptop was provided to the participants well in advance. LAPP had ensured a strong wifi network along with eduroam network. SURFsara had made cartesius machine access available to conduct the hands-on sessions on GPU and Parallel programming. Gitter chatrooms were set up for participants to interact and discuss during the school.

4. Feedback

After the event, an online questionnaire was sent to the school participants to have their feedback on the scientific programme as well as the organization of the school. We received inputs from 36 participants, a summary of these inputs is presented below.

4.1 Event Organization

Majority of the participants were satisfied with the event organization, accommodation arrangements as well as the social events. For the next edition of the school, about 50% of the participants suggested to organize a 5 days school instead of 4 days school.

4.2 Expectations of the participants

Despite of having participants from diverse background as well as research projects, the school managed to fulfil the expectations of most of the participants. Nonetheless, we have received some useful feedback that would prove to be useful for the organization of the next training event

- More hands-on would be useful.
- Github tutorial is required.
- The theory session duration can be reduced and more time should be dedicated to hands-on sessions.
- The difficulty levels of the courses can further be optimized.

4.3 Course Content

Vast majority of the participants informed that the training event has upgraded their programming knowledge and skills. We have summarised the feedback on each module below:

Module 1 - Efficient code writing: Most of the participants were satisfied with course content. Some of the critical feedback included more time allotment on hands-on session.

Module 2 - Parallel & GPU programming for Python: We received mixed inputs for this particular module with some of the participants finding course content satisfying while some of them finding the courses too advanced. Similar to module 1, more time allotment on hands-on session was one of the suggestion on this module.

Module 3 - Python libraries for astronomy & astrophysics: From the feedbacks, we found that this module was extremely well designed and adopted to the background of the participants. This was a very useful module of the school for most of the participants.

For the next edition of the school participants have provided us following suggestions for the course modules: Object oriented programming , Machine learning , Data mining, Integration between Julia - Python - R - C, Parallelisation on clusters, Matplotlib, Hands-on on Github/code testing, Pynbody package, How to well structure and build your own python library , Python and C++ interface, VO tools.

5. Self-Evaluation

Overall, we believe the training event addressed the vital needs of a wider community of astrophysics and astroparticle physics. We have received more than expected number of request for registrations. Most of the participants informed us that they have never been exposed to a formal training in software programming and everything they knew about programming languages by self-study. Concerning gender balance, about 30% of the participants were women participants. Due to its international nature, the event also managed to attract the regional press agencies. More details on press coverage of the event is available on Press Releases (Appendix 2). In addition, the participants also widely shared their experience during the school on Social Media such as Twitter and Facebook.

Based on the feedbacks from the participants during and after the school, we obtained SWOT analysis.

Strengths: The scientific content of the school for the theory as well as hands-on sessions on Python programming was very well selected and designed to adapt the needs of the participants. The tutors selection for the courses, the picturesque location of Annecy as well as social events providing opportunity to the participants to network and share their experiences also played a key role in the success of the event.

Weaknesses: As per the feedback from the some of the participants, it was observed that the module 2 was not well adapted for the knowledge level of the participants from diverse background.

Opportunities: From this first experience with school on software programming for Astronomy and Astroparticle physics we are convinced that there is a huge need for similar training events not only for young researcher but also the senior researchers. We intend to address these needs of the wider astronomy community in our next training events.

Threats: The only potential threat we identified is not being able to optimize the difficulty level for the courses. We will be addressing this threat using more optimized input form for

the participant that will be shared with the tutors, providing them precise idea on the background of the participants.

Figure 1 Group photo of the participants of the 1st ASTERICS-OBELICS International School, 6-9 June 2017, LAPP, Annecy.

Appendix 1: List of tutors

Module 1: efficient code writing

- **Name**: Tim Jenness, Data Management Systems Engineer for Large Synoptic Survey Telescope, Tucson, USA .
  Bio: Tim is a member of the LSST Data Management System Architecture team. He investigates new technologies; define development policies and processes, and worries about software interfaces. Additionally, he is responsible for software change control and system requirements along with verification planning.

- **Name**: Zheng Meyer-Zhao, HPC consultant at SURFsara
  Bio: Zheng is an HPC consultant at SURFsara. She studied Computer Science at VUB in Brussels and became interested in parallel programming and machine learning since then. The main focus of her work is to provide technical support and trainings to users of the Dutch national supercomputer. She also works for the PRACE (Partnership for Advanced Computing in Europe) Training project. Before joining SURFsara, Zheng worked as a software engineer in the VLBI group at Academia Sinica Institute for Astronomy and Astrophysics (ASIAA) in Taiwan.

- **Name**: Karl Kosack , Astrophysics Division, CEA Saclay, France
  Bio: Karl is a researcher astrophysics, specializing in the study of galactic sources of very-high-energy (VHE) gamma rays, particularly with the HESS telescope array and the upcoming Cherenkov Telescope Array (CTA) projects. In both of these projects (as well as for the Whipple 10m and VERITAS telescopes) he has been strongly involved in the software development effort, and is currently the coordinator of the data reduction pipelines project in CTA.

- **Name**: Frossie Economou, Technical Manager for the Science Quality and Reliability Engineering group at Large Synoptic Survey Telescope, Tucson, USA .
  Bio: Ever since discovering that telescopes are a lot more fun that academia, Frossie Economou has spent her career working on astronomical software, from data reduction pipelines, to observation management systems to public archives. She is now at the Large Synoptic Survey Telescope (LSST) where she set up Data Management's Science Quality and Reliability Engineering team (SQuaRE). Among other activities, SQuaRE supports a varied developer infrastructure for the far-flung DM team. She in an advocate for open source collaboration, sustainable software development and agile and devops practices in astronomy.

Module 2: parallel & GPU programming for Python

- **Name**: Valeriu Codreanu, SURFsara
  Bio: Valeriu has a Phd in Electrical Engineering with a thesis on developing a novel multi-threaded computer architecture. Afterwards he did a postdoc on GPU computing at the University of Groningen, where he helped develop a CPU-to-GPU compiler called GPSME, under an EU project with the same name. He then did another postdoc on Embedded Computing at Technical University of Eindhoven focusing on real-time embedded systems. Since the end of 2014 he has joined SURFsara as an HPC consultant, and is part of various projects such as PRACE. He was also PI of the GPU Research Center at SURFsara, and also of the current Intel Parallel Computing Center. His interests lie in efficient computing, scaling, and in the application area of deep learning.

- **Name**: Damian Podareanu , SURFsara
  Bio: Damian has a BSc in mathematics and computer science from the University of Bucharest, followed by a MSc in Advanced Computer Architectures at the Polytechnic University of Bucharest, and MSc in Artificial Intelligence at the University of Groningen. Through this period he worked as a software developer or industrial researcher for various companies developing code in a multitude of programing languages ranging from the digital design level - VHDL -, to high level interpreters such as Python or R. Since joining SURFsara in 2016 as a HPC consultant, he has been involved in PRACE, the IPCC 'Scaling up deep learning on the 2nd generation Xeon Phi (Knights Landing) platform' project, and several other application optimization and performance tuning projects.

Module 3: Python libraries for astronomy & astrophysics

- **Name**: Tamas Gal , KM3NeT ; FAU-ECAP, Germany
  Bio: Tamas is PhD Student in Physics at the Erlangen Centre for Astroparticle Physics (ECAP) and the main developer of the Python/Cython based KM3Pipe analysis framework used in the KM3NeT project mainly for online and quasi-online data analysis. The aim of his thesis is online monitoring and live reconstruction and analysis of neutrino events seen by the KM3NeT neutrino telescope. His favourite languages are Python and JuliaLang for scientific computing and Haskell as a hobby.

- **Name**: Jean Jacquemier, LAPP
  Bio: Jean is a software developer. He has been working for french fundamental particle and astroparticle physics research from last 14 years. He is mainly developing C/C++ and Python data reconstruction software

- **Name**: Hendrik Heinl, http://www.g-vo.org
  Bio: Hendrik studied German Language and Literature and Political studies at the Universiti Heidelberg with a focus of digital transcription of medieval manuscripts (when there wasn't something like Digital Humanities). Following up hes was self employd as web developer. In 2013 he joined ARI and as of today he is responsible for Education and Outreach in the GAVO project.

- **Name**: Johannes King, Max Planck Institute for Nuclear Physics, Heidelberg, Germany
  Bio: Johannes gained first programming experience in an undergraduate course using C/C++. Since then he has written most of his analysis code in the ROOT framework until the start of his PhD. Then, he switched mostly to Python and started contributing to open-source projects. In gammapy he is working on the spectrum module which is responsible for 1D spectral analysis. In astropy he is involved in the newly created regions subpackage.

- **Name**: Axel Donath, Max Planck Institute for Nuclear Physics, Heidelberg, Germany
  Bio: Axel is a PhD student in Gamma-ray astronomy working on the H.E.S.S.Galactic plane survey . He is the co-developer of gammapy, an open source python package for gamma-ray astronomy . He is an Astropy core contributor, where he mainly worked on `astropy.convolution` and `astropy.modeling`. His github profile: https://github.com/adonath

- **Name**: Nicolas Chotard, ASTERICS postdoc in LSST, ASTERICS , LAPP , LSST
  Bio: Nicolas is a researcher in observational cosmology and Python developer. His skills are focused on development, validation, and maintenance of scientific softwares and data production pipelines, along with data analysis of cosmological probes.

- **Name**: Tammo Jan Dijkema , ASTRON
  Bio: Tammo Jan is a scientific software engineer at ASTRON, working mainly on calibrating and imaging the visibilities from LOFAR telescope. His background is in numerical mathematics. Tammo Jan is the project manager for the LOFAR Calibration & Imaging team, is involved in the casacore package (which forms the base of CASA) and contributes to the development of SKA. His computer interests include git, travis, qgis, Excel and of course python.

- **Name**: Pierre Aubert , LAPP, France
  Bio: Pierre is a PhD student in High Performances Computing (HPC), involved on the H2020 ASTERICS and CTA project. The aim of his thesis is to use HPC to improve the speed of the CTA data reconstruction

Appendix 2: Press Releases

Le dauphine, 12 June 2017

ANNECY-LE-VIEUX - DESTINÉE AUX PHYSICIENS, POUR RÉPONDRE AUX DÉFIS LIÉS AU BIG DATA

Le LAPP accueille la première école internationale Asterics-Obelics

Des ateliers rassemblant des chercheurs internationaux. Photo Le DL/Marc KIEVITCH

Le LAPP accueille la première édition d'une école destinée aux physiciens du monde afin de leur donner les outils nécessaires pour répondre aux défis liés au Big data. Ils sont 80 participants de onze pays, dont une quinzaine d'enseignants (cinq sont issus du LAPP). Les doctorants, post-doctorats et chercheurs exercent dans les domaines de l'astrophysique, de la physique des astroparticules et des ondes gravitationnelles. Cet événement d'envergure internationale est une reconnaissance mondiale du LAPP dans son domaine de recherches et d'enseignement.

Le Dauphine, Sunday 23 July 2017

ANNECY-LE-VIEUX | Le LAPP a organisé la première école internationale pour 11 pays différents

Une plateforme d'enseignement scientifique au laboratoire de physique des particules

Le LAPP (Laboratoire de physique des particules) dont le rayonnement international dans le domaine de la recherche scientifique est bien établi, veut être désormais une plateforme d'enseignement scientifique. À destination des scientifiques du monde mais aussi à destination des établissements scolaires de la région.

C'est ainsi que le LAPP a organisé la première école internationale Asterics-Obelics sur la programmation avancée pour l'astrophysique et la physique des astroparticules pour plus de 80 participants, de 11 pays différents dont les États-Unis.

Asterics de son véritable nom : "Astronomy ESFRI and research infrastructure cluster", est une infrastructure, financée par la Commission européenne dans le cadre d'un programme de financement pour la recherche et l'innovation.

Elle rassemble pour la première fois 22 instituts d'astronomie, d'astrophysique et de physique des astroparticules afin de créer une synergie entre eux.

L'enjeu est la résolution des défis liés au "big data" pour l'astronomie européenne. L'école Asterics est plus particulièrement organisée par le groupe de travail Obelics ("Observatory e-environments LInked by common challengeS") dirigé par le LAPP.

Cette première expérience d'école au plus haut niveau international conforte l'image du LAPP à travers le monde.

Fort de cette notoriété le LAPP souhaiterait aussi innover et accompagner sur l'information concernant le métier de chercheur à travers les lycées de l'agglomération.

M.K.

Le docteur Giovanni Lamanna (à gauche), directeur du LAPP souhaite étendre la connaissance du métier de chercheur aux scolaires de l'agglomération. Photo Le DLM K.