Astronomy ESFRI and Research Infrastructure Cluster

Minutes: 2nd OBELICS (WP3) MEETING 2016

14-16 September, UCM Madrid.

List of Participants

Number	Participants
1	Thomas Vuillaume (LAPP, France)
2	Tammo Jan Dijkema (ASTRON, Netherlands)
3	Arpad Szomoru (JIVE, Netherlands)
4	Jaime Rosado (UCM, Spain)
5	Dominique Boutigny (LAPP, France)
6	Peter Hague (Cambridge,UK)
7	Tarek Hassan (IFAE, Spain)
8	Vincent Poireau (LAPP, France)
9	Stefan Geißelsöder (FAU, Germany)
10	Rob van der Meer (ASTRON, Netherlands)
11	José Luis Contreras (UCM, Spain)
12	Giovanni Lamanna (LAPP, France)
13	Nicolas Chotard (LAPP, France)
14	Fabio Pasian (INAF Trieste, Italy)
15	Eric Chassande-Mottin (APC, France)
16	Cristiano Bozza (INFN, Italy)
17	Daniele Cesini (INFN, Italy)
18	Bojan Nikolic (Cambridge,UK)
19	Tino Michael (CEA, France)
20	Jayesh Wagh (LAPP, France)
21	Alessandro Costa (INAF Catania, Italy)

14 September 2016 (12h15-13h00): Welcome address & introduction session

Welcome Address: Prof. José Luis Contreras, UCM Madrid.

Prof. José Luis Contreras welcomed all the OBELICS members and provided with practical information. The welcome address was followed by a round of introduction.

ASTERICS & OBELICS: Dr. Rob van der Meer, ASTRON, ASTERICS Project Manager, ASTRON.

- H2020-AENEAS is approved. This proposal was submitted in response to call for Science data center for SKA-Led by ASTRON and Dr Michael Wise.
- AENEAS will address data generated from SKA.
- Michael Wise and/or Dr Marco de Vos, ASTRON Managing Director, can help to involve H2020-AENEAS project participation at the OBELICS Workshop.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

Objectives of 2nd OBELICS face to face meeting: Dr. Giovanni Lamanna, OBELICS work package leader, LAPP.

- OBELICS has to fulfill commitments through deliverables and supporting ESFRI projects and cooperation within clusters.
- Actions within each task should be upgrading the roadmap, link with ESFRI projects;
- The OBELICS workshop & OBELICS international school was internally announce. The feedback on priorities of ESFRI projects should be the aim of the upcoming OBELICS workshop. Task leaders input for content for workshop and training events.
- Midterm review report: inputs from task leaders are important.
- Need for improving the OBELICS results communication internally and externally. Especially Communication on software repositories can be improved.
- OBELICS activities started up slowly as more time was required to build up team and to be sure on the content to be communicated to external community.
- Virgo-EGO participation in ASTERICS is important. Recruitment table must be updated.
- Discussions on VIRGO & gravitational waves community involvement at the workshop and OBELICS workpackage. Additional partners from Gravitational waves community should be involved in OBELICS without funding. Virgo community networking with data centres for sharing gravitational antennas, and OBELICS can be used to commence discussions on sharing gravitational waves antennas. Address at least one of the topic at the workshop that will be of interest for VIRGO.

14 September 2016 (14h30-18h00): Data GEneration and information eXtraction

Overview of Task 3.2 (D-GEX) Data GEneration and information eXtraction: Prof. Jose-Luis Contreras, UCM Madrid.

- The role of DGEX in OBELICs is the first step in data analysis, it is the smallest WP (140 person months) with activities on Data Model and Formats, Low-Power computing, streaming data.
- No inputs from Dr Paschal Coyle from CPPM – ANTARES/KM3. CPPM has recruited a PhD student Liam Quinn who is working on energy reconstruction. This work relates with D-ANA. Some groups still do not have contact persons. Almost no group has a report of activities in the wiki. Clear separation among tasks, same people and similar goals is required. Internal brainstorming is required to publish and disseminate deliverables.
- KM3NeT reconstruction algorithms do not work standalone but they work on framework and the framework cannot be open sourced and disseminated. In this case such algorithm can be published mentioning the algorithms are inspired by a method that is applied to KM3NeT framework.
- Possibilities of cooperation's: Event based, big data based (HDF5 data storage) & Observatory use cases. Next steps for collaboration can be low cost parallel computing and data works. Distinction between priority projects and added value from ASTERICS should be considered.
- HEP Software Initiative: Various techniques and algorithms can have commonalities with OBELICS.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- Programming language: At LAPP a survey within CTA was conducted to check the cpu performance with C, C++ & Python. The software is a work in progress. A document can be produced in this context since there is a 24 months deliver on Benchmarks. In CTA context python and Java are excluded and C and C++ are preferred languages.

CTA @ INAF for ASTERICS: Prof. Jose-Luis Contreras on behalf of Dr L.A. Antonelli.

- INAF is actively involved in high energy Astroparticle Physics and Major Atmospheric Gamma Imaging Cherenkov Telescopes (Magic), ASTRI CTA, GAW etc.
- INAF activities are still centered in the ASTRI pipeline, adapting it to low-power computing CUDA and JETSON. ASTRI is a prototype for CTA. A talk was presented at the SPIE conference on July. For SKA this prototype won't work because of the cost of moving data to memory.
- Data Reduction Pipeline was presented. For image cleaning, CUDA integration is implemented. It allows seamless and flexible integration of parallel code.
- Portability of algorithms is one of the upcoming aim of INAF to apply it to various projects such as MAGIC, LST, MST etc.
- Low power computing is of interest to SKA & INFN; at UCAM workshops and summer schools are organized addressing low power computing.

OBELICS WP 3.2 D-GEX LOFAR status: Dr Tammo Jan Djikema, ASTRON.

- LOFAR Overview: 60 stations across Europe, Groningen is the Central Correlator which is connected to the stations with 240 Gbit/s fibre connections. Streaming to the online storage system at the rate of 80Gbit/s. Online processing cluster has 2.2 PB temp storage.
- Superterp has a setup for measuring cosmic rays and observations from this station have been published in Nature.
- Low Power Architectures: For imaging gridding and degridding are some of the most important tasks.
- Offline Processing: 2 people from ASTERICS are involved. Tammo works on calibration of raw data using StefCal, SageCal and extensions.
- For Data format LOFAR doesn't use HDF5 because when the standard software for Radioastronomy was developed HDF5 wasn't available. Therefore, they use Casacore Data Table Systems (CTDS). However, one of the raw data stream of LOFAR use HDF5. One of the advantage of CDTS is that it provides compressed storage manager.

Activities at IFAE: Dr Tarek Hassan, IFAE.

- IFAE is involved in Task 3.2 & Task 3.3.
- CTA Data models: Data Level 0 (DL0) is generated directly from the telescopes. DL0 is in HDF5 format. DL1 & DL2 are used internally within CTA so may not be useful for other ESFRI projects however DL3 might be of interest to other ESFRI projects. It's an open source code.
- Open DL3 (FITS format): *Event lists (event-wise energy, RA, DEC, time...) of gamma-like events + IRFs describing the instrument performance (Eff. Area, BG rate, direction/energy dispersion) + TECH data describing details of the observations (pointing, obs. conditions, etc.)*

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- For CTA & IACT's IRFs change continuously change and eventually telescopes and conditions of observations affect the performance of the instrument. Standardizing format allows us to perform analysis on Fermi Data.

INFN-CNAF in T3.2: Dr Daniele Cesini, INFN.

- CNAF Bologna is the National Computing Centre (TIER-1) of INFN.
- CNAF is supporting 4 LHC & 34 non-LHC experiments.
- CNAF is supporting the creation of the computing models of experiments involved in ASTERICS such as KM3, in particular for what concern the data management and the services connected to it.
- CNAF will recruit 1 person by the end of November 2016 under ASTERICS project, till then Daniele Cesini will manage the task responsibilities. While recruitment emphasis was given to low power computing and any proposal for collaboration with other task would be welcomed.
- CNAF is leading the COSA project (www.cosa-project.it) which is addressing low power computing. It is an INFN funded project, lasts 3-years starting from Jan 2015.

OBELICS Task 3.2 – LAPP contribution : Dr Thomas Vuillaume, LAPP.

- Data Lossless Compression (work of P. Aubert et al.): Advanced Polynomial Compression+ LZMA yields compression ratio of 4.84 within 24.646 sec i.e. 19x faster. With further vectorization compression speed can further be enhanced (work in progress at LAPP). Time of decompression is also equally faster.
- Data format: LAPP Sorted HPC Autotuned Data Optimized Kernel format. It is used for CTA Data.
- Besides being user friendly, it provides all the required info eg. Telescopes, camera etc. It allows CPU Data prefetching & vectorization of algorithms.

OBELICS D-GEX UCAM Summary: Dr Bojan Nikolic, UCAM.

- Alma Data Streaming: It has 2 methods of data processing. 1. Pure Streaming 2. Microbatched processing.
- D-GEX task (architecture and streaming data format) is related to Data processing unit of SKA Science Data Processor. It focuses on Image processing, feature extraction, cross matching etc.
- SKA Central Signal Processor streams data at the rate of 1TB/sec and it mostly contains noise. The Objective of data processing unit is to extract information from data and then discard the data.
- Challenge: There are 5 major SKA requirements and each requirement should be solved and addressed by specialized framework or introduce one single framework that will address all the requirements.
- Apache-Spark (JAVA based framework) is being investigated to handle streaming, iterative streaming, quality assessment and databases. Spark works up to 10000 nodes of data lots.
- CTA & SKA computing models show some resemblance and can be further exploited.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

Activities of UCM-GAE within the Task 3.2 D-GEX: Dr Jaime Rosado, UCM.

- UCM is working in CTA: mainly Data model issues. It was responsible of the first deliverable (report on formats and computing frameworks). Now UCM is looking for synergies among projects. Low level formats (HDF5).
- High level formats being developed in gamma-ray astronomy (Fermi,CTA, etc) could be extended to other event based projects. Some use cases were presented on this point.
- HDF5 has advantages for management of large amount of data and is being establishing as a standard with many available tools and libraries.

Towards an open high-level data format for gamma-ray astronomy: Dr Tarek Hassan, IFAE.

- CTA will operate as an open observatory, with most of the open source software and the data will be available in DL3 format.
- DL3 development is result of collaborative efforts between HESS, CTA, MAGIC, VERITAS, FERMI.
- CTA pipeline is being developed and available on https://github.com/cta-observatory
- Science tools : ctools & gammapie γπ
- Event based projects such as KM3NeT can involve in DL3 development.
- Interoperability is the aim to be achieved through DL3.
- Gravitational waves: VO Event format, Meta data, layer across messengers, more complex Meta data reconstruction of source direction for particular event.

15 September 2016 (14h30-18h00): Data systems INTegration session

CTA resource requirements: Dr Thomas Vuillaume, LAPP.

Inputs from ASTERICS partners are most welcome on Complex and efficient Archive Management System of CTA, more specifically

- To map and catalog data.
- To optimize the data reduction and simulation pipelines.
- To keep track of and synchronize changes applied to data during analysis chain – Data Provenance.

EUCLID: Dr Thomas Vuillaume, LAPP.

- EUCLID is an ESA mission will produce 850 Gbit/sec raw compressed science data.
- EUCLID catalogue would contain 10^{10} objects.
- Its data model is similar to CTA data model.
- EUCLID storage needs are divided into pre-launch (all simulation & modeling ~ 18 PB) & post launch (~ 100PB). Number of CPU cores required post launch = 20000.
-

LSST Data and Computing: Dr Nicolas Chotard & Dr Dominique Boutigny, LAPP.

- LSST will capture 10 years movie of the sky in 4D (x,y,z,t). The project is approved and will operate in 2022 – 2031.
- LSST is a significantly wide and fast telescope as wholesome sky mapping can be done in 3 days.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- LSST is a facility under construction in La Cerena, Chili. One archive site in NCSA-US, another archive site is in CC-IN2P3 France. 15 TB of raw data / night will be generated from ~10 million transient events / night.
- LSST Software Stack: Data processing software developed in Python & C++, it's an open source software.
- Three data product levels: nightly data product, annual data product, user created data product
- SQL Database Qserve: Query service, developed at SLAC + IPAC Design optimized for astronomical queries.
- QServe test bench has been deployed at CC-IN2P3 with collaboration with dell
- Prototype data access center at NCSA is testing QServe and will be duplicated at CC-IN2P3
- Scientific validation of the database QServe
- how to articulate the LSST use case within OBELICS work package is one of the point that can be discussed during the workshop.
- What is the difference between licenses of LSST and similar projects? Some parts of the software are being developed from scratch and some parts are being improved from previous versions.
- LSST-SKA collaboration can be possible for QServe development.

OBELICS WP 3.3 D-INT LOFAR status: Dr Tammo Jan Djikema, ASTRON.

- For LOFAR imaging pipeline, ideally processing time < observing time
- Pipelines is inside the Docker container because it's flexible and allows user to run any obscure software.
- Standard pipelines run on LOFAR software and UNICORE was considered as an alternative.
- A&A is done through grid certificates. No processing on grid yet and Docker doesn't work well on grid.

SKA Data Archives: Dr Peter Hague, UCAM.

- Data produced by SDP archived on site
    - then mirrored
- Scale of the data is of the size of 1PB / 6hours or 6 EB/yr.
    - A Use Case *The HI Nearby Galaxy Survey (THINGS) was presented.*
- This work involves combining datasets from multiple telescopes and it is carried out in the context of VO. The focus of this work is more on the data and less on the instruments.
- Contribution from SKA to ASTERICS and especially Task 3.3 (D-INT) was discussed.
- Not all hiring has completed yet. One more person will be hired by end of October.
- D-GEX is focus of UCAM, UCAM is also involved D-INT & D-ANA, D-INT & D-ANA is part of the SKA and observatory, pipelines

KM3NeT e-Needs: Dr Cristiano Bozza, INFN.

- Neutrino based astronomy.
- 2 detectors: arca (located in Italy) & orca (located in France).
- 3 phases of the project currently at phase.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

Investigation of Docker for data analyses in the context of ASTERICS: Stefan Geißelsöder, FAU.

- Some people from ECAP-FAU are working on FITS format.
- Docker is currently being investigated for KM3NeT in the context of ASTERICS.
- Docker is a suited solution for heterogeneous softwares and data. The problems caused due to upgraded version of softwares can be avoided with docker.
- One common initial Docker image to ensure reproducibility for analysis and results.

CTA@INAF for ASTERICS: Dr Jose-Luis Contreras on behalf of Dr LA Antonelli.

- INAF involved in DGEX and DINT.
- FITS for pipeline = DL4.
- Deliverables & Achievements: Full data reduction chain for CTA pre-production ASTRI telescopes on different HW platforms., libraries & softwares, report on benchmarking, Participation and presentation of the work to specific conferences (e.g. GTC, SPIE, …).
- Future activities: Explore reconstruction methods for imaging for IACTs instrumentation.

Deliverable - Analysis Report on Resource Requirements: Dr Tammo Jan Djikema, ASTRON.

- Thomas Vuillaume (CTA), Peter Hague or Bojan Nikolic (SKA) can participate in drafting the report.
- EUCLID inputs are missing. Maybe Dr Fabio Pasian can do that but he will need authorization from the responsible.
- From each ESFRI project one person should provide a table of relevant bottlenecks, solutions under test and or under development. Selected topics to be identified (survey) that can be part of upcoming actions and addressed during workshop and it should be related to potentially data transfer, data sharing, data archive and database, user access to data, workload management services and any other computer model interpretation. Take the essentials from various inputs, summaries the issues and prepare conclusion.
- Tammo Jan Dijkema to share a google doc which will be accessible to different members of ESFRI projects.

15 September 2016 (14h30-18h00): Data ANAlysis / Interpretation

Summary & Planning: Dr Bojan Nikolic, UCAM.

- Missing contacts from IAP & CPPM
- Task 3.4.1: Peter Hague is working on Stat planner libraries and Bojan is involved in MW-Inference libraries.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

OBELICS Task 3.4: LAPP contribution – Dr Thomas Vuillaume.

- HPC libraries for python: Can perform calculations such as reduction, barycenter and first and second momenta calculation. This general library can be applied to other projects.
- Relation with Processor architecture & processor time: Intel processor and less than 10 years old, vectorization will be faster with next generations of Intel processor,
- Halide Framework that's used in mobile phone image processing and GPUs, comparison with Halide could be useful for benchmarking report, Bojan to provide weblink to Thomas.

CEA Status Report (CTA Group): Dr Tino Michael, CEA.

- 12 people, 1 ASTERICS postdoc from April 2016 (Tino).
- Wikipages= existing libraries to be improved, packaging or enhancement of the libraries
- To do list
  - Optimise parameters for wavelet cleaning – already implemented in cosmos smart library
  - adapt wavelet algorithm for hexagonal image grid
  - evaluate background rejection and shower reconstruction
  - in light of wavelet cleaning
- Sky Images
  - continue to work on improving sensitivity to transient sources -
  - disentangle complex, large FoV

Obelics WP 3.4 D-ANA LOFAR status: Dr Tammo Jan Dijkema, ASTRON.

- 2 ASTRON persons working on all the three tasks
- Image domain gridding – patent filed and awaiting decision
- CASA: Gridding is part of whole imaging process, to test it on real data set existing image software is required.
- Adapted Casa = Casasynthesis (Borrow open source code from CASA and split it) + LOFAR library AWImager2
- In the meantime, WSclean has been developed, WSclean (C++).

Obelics WP 3.4: Dr Arpad Szomoru, JIVE.

- 4 people involved in OBELICS from JIVE
- Under OBELICS: Create a system that allows modern notebook-style approach to post-processing of VLBI data,
- Speed up pipeline development under HILADO with Bojan to Eliminate most of the redundant re-calculation
- Notebook framework Jupyter : being adapted to the version of Python used by CASA.
- Ongoing work: Plumbing Casa into the Jupyter framework, additional layer on top of CASA
- Jupyter & Notebook based work should be interesting to ESFRI projects. LSST is interested in exploring more about Jupyter. Jupyter for gateway in CTA can be one possibility.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- Common Jupyter or Python action within D-ANA can be discussed
- Jupyter is for remote access protocol. Jupyter works very well with the containers and docker.

INFN for OBELICS -D-ANA: Dr Cristiano Bozza, INFN.

- Activities under task 3.4
- **CORELib :** Background to many experiments, Also a tuning benchmark, Potentially useful to other communities, Currently using CORSIKA as generator
  - Production available on many «private» institutional servers
  - how to make many TB (20-30 TB) available for Open Access? EUDAT B2xxx services seem viable, discussion needed, efficient open access where any citizen can download terabytes of data = community participation is required to access such data, ASTERICS is distributing codes but data should also be shared,
  - It is important to know what the user requirements are.
  - CTA : Building cta computing grid for monte carlo simulations,
  - Archive for cosmic rays is an action to be executed, survey of user requests should be done within ASTERICS.
  - Central Repository for simulated data to be kept under CORSIKA open access archive. For CTA may not be useful, showers are re-simulated, storage that will be required will be huge. In the benchmarking report computational requirements vs storage requirement can be included.
  - Budget availabilities should also be considered
  - Inputs required: Make data «more easily» available on the Web
- **ROAst:** ROot extension for Astronomy: Classes to access astronomical catalogues; Generators of primary particles (neutrinos will be implemented, others will be supported only as placeholders)
- **pLISA:** parallel Library for Identification and Study of Astroparticles Generic interface for machine learning and MVA-based study/id of events Optimized GPU code to run "under-the-hood" to implement algorithms Goal: allow generic detectors and easy user extension to library

FAU/ECAP for OBELICS -D-ANA: Dr Cristiano Bozza on behalf of Kay Graf.

- FAU/ECAP is highly involved in Computing and Software in KM3NeT, e.g. simulations coordinator, computing and software coordinator, physics and software coordinator are at the institution

Task 3.4.1 Summary & Planning: Dr Bojan Nikolic, UCAM.

- Jayesh to share with task leaders – dropbox the common report.
- Impact, scope dissemination in first 18 months: Share the grant agreement with task leaders.
- Final release of software libraries is one of the deliverable.
- Contributions missing from IAP-EUCLID & CPPM = 6 months for Task3.4= Paschale Coyle hired a PhD student Liam Quinn), inputs for pLISA.
- Repository fulfilled the deliverable, KM3Net participation should be included.
- In the report, a summary should be presented with future contribution from KM3NeT.
- Limited overlap in Task 3.4, Potential synergies & reuse should be investigated.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- PLIBS is advanced but there is scope for further developments.
- Discussion over the schematics and defining the links.
- ROAST is an enhancing root for radio astronomy. It can use PLISA but not the other way round.
- ROOT & CASA: Most of the radio astronomy is root and casa based. We can't merge both.
- Additions to CASA can be introduced that can work as a separate module and pLIBS as well.
- For INFN, not enough EU funding for manpower to achieve the goals – will add resources from staff and try to attract interest from other KM3NeT members.
- Start with building some of the connections instead of targeting entire schematics.
- Repository to be communicated through wiki page, a logical diagram can be made for the wiki page
- ESO-ESA to be involved in the project.

INAF for D-ANA Task 3.4.2: Dr Fabio Pasian and Dr Alessandro Costa, INAF.

- INAF is the only ASTERICS partner involved in Task 3.4.2 of D-ANA, i.e. the activities related to analysis and harmonization of Authorisation & Authentication (A&A) and Workflows.
  - For A&A, tasks are to analyse existing requirements and protocols, and to propose and integrate, as required, a global infrastructure using agreed-upon standards
  - For Workflows, tasks are to analyse and integrate Workflow architectures for the orchestration of compute-intensive data analysis on distributed computing infrastructures
  - Both activities require liaison and coordination with ASTERICS WP4 (DADI).
- A selected set of A&A and Workflow tools were included within deliverable D3.4 (initial Software Libraries), freely available from the ASTERICS Wiki. It is a set of pointers to distributions and can be of use within ASTERICS (in OBELICS, but also across WPs: of interest for DADI and probably CLEOPATRA).
- INAF staff working for T3.4.2: 5 permanent staff plus 2 persons hired in Trieste: Marco Molinaro (from July, in charge of coordination with DADI), Andrea Bignamini (from October, in charge of A&A and Workflows and coordination with H2020 project Indigo-DataCloud). An additional person will be hired in Summer 2017.
- Several technologies have been investigated for Authentication (Shibboleth, X.509 certificates, enID Connect, …), Authorisation (Grouper, GMS, custom), Integration with VO standards (VOSpace, ObsCore).
- For CTA:
  - Use Case Collections, identified as: a) constraints on a system solution; b) a representative set of activity sequences corresponding to constraints and Scenarios;
  - User Requirements;
  - An AAI prototype has been implemented, based on eduGain Authentication and Grouper for Authorisation;

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- To the AAI prototype a CTA Workflow Management System prototype has been attached.

- For SKA:
  - Use Case: Enterprise solution for user access to resources and zones (buildings, restricted areas etc..) ;
  - Prototype status: SAML based user authentication plus use of MySQL for group management;
  - Integrated with federations (EduGain);
- The H2020 multidisciplinary AARC2 project has been approved (letter of support from ASTERICS). The CTA A&A case is one of the funded pilots.
- INAF is supporting a cloud federation activity between EGI and CANFAR, based on sharing of resources through the Virtual Observatory VOspace standard and allowing A&A interoperability between the EduGain-based EGI mechanism and GMS on the CANFAR side.
- Workflows Management Systems and "languages" were specifically evaluated, also exploiting results of the European and international big infrastructures and projects (EGI, ErFlow, Workflow4ever, IVOA, CTA & SKA prototypes): the main systems are TAVERNA, KEPLER, PEGASUS (available in software repository).

Task 3.4.2 Summary & Planning: Dr Fabio Pasian, INAF.

- Both A&A prototypes (CTA, SKA, EGI-CANFAR) and Workflow systems have been proven to work.
- Key issues for the future for both A&A and Workflow systems are: suitable maintenance plans, availability of existing tools and solutions, long-term support, backwards compatibility.
- T3.4.2 work needs to take into account the deliverables from other H2020 projects working on A&A (INDIGO-Data Cloud, AARC, EGI-Engage), analysing their integrability with the needs of the projects supported by ASTERICS.
- The technology integration tests have shown that a key issue is the integration between A&A and Workflows, Computing, Storage systems, Archives (long term preservation). Do we have resources (i.e. personnel) for that? Is it important for the project goals?

16 September 2016 (9h30-12h00): MAnagement User engagement & data Dissemination (MAUD)

TASK 3.1: Jayesh Wagh & Dr Vincent Poireau, LAPP.

- 3 General workshops & training events to be organized
- D3.2 (training event) + D3.6 (workshop) will be merged and organized from 12-14 December 2016 in Rome.
- Format of the event could be 2 days of workshop & 1 day training event
- Audience could be mix of OBELICS members & PhD students
- Potential common topics of interest for ESFRI projects could be vectorization, GPU, Data storage, SPARK, QServe which were discussed and presented during the meeting.
- Expression of Interest for Industrial participation is work in progress. It could be in the form of PhD project.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477

- Technical report for 18 months review is due on 28th October 2016.
- Financial report is due on 9 December 2016.
- Intra-wp3 communication to be improved and table of actual number of personnel should be updated.
- OBELICS international school to be organized in June 2017 at LAPP, Annecy-le-Vieux.
- This event will be a combination of theoretical lectures followed by practical sessions.
- Potential topics for this school could be: Hadoop, Python for blender, Spark, Modern C++.
- A keynote lecture from UBISOFT and visit to CERN computing facilities can be organized.

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 653477