Euclid UK Computing

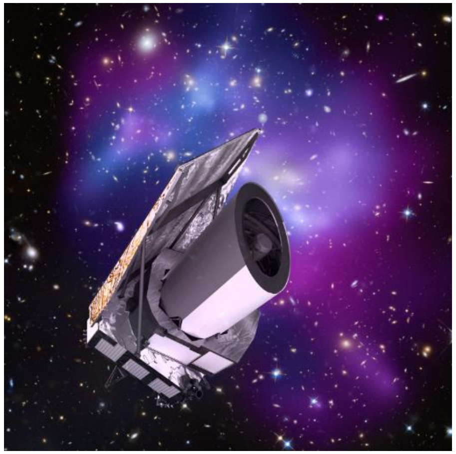
Mark Holliman

SDC-UK Deputy Lead

Wide Field Astronomy Unit

Institute for Astronomy

University of Edinburgh



Euclid Summary

- ESA Medium-Class Mission
  - In the Cosmic Visions Programme
  - M2 slot (M1 Solar Orbiter, M3 PLATO)
  - Due for launch 2021
- Largest astronomical consortium in history: 15 countries, ~2000 scientists, ~200 institutes. Mission data processing and hosting will be spread across 9 Science Data Centres in Europe and US (each with different levels of commitment).
- Scientific Objectives
  - To understand the origins of the Universe's accelerated expansion
  - Using at least 2 independent complementary probes (5 probes total)
  - ***Geometry of the universe***:
    - Weak Lensing (WL) Galaxy Clustering (GC)
  - ***Cosmic history of structure formation***:
    - WL, Redshift Space Distortion (RSD), Clusters of Galaxies (CL)

Controlling systematic residuals to an unprecedented level of accuracy, impossible from the ground

Euclid UK

- Four main areas of work:
  - **Instrument building**: VIS detector
  - **SDC-UK**: UK data processing operations for Euclid Mission. Requires running the full Euclid pipeline for approximately 10% of the total mission data and storing/serving the resulting archive data during the mission.
  - **OU-SHE/OU-LE3**: Develop the weak lensing software pipeline (algorithms and functional pipeline components), and some of the Level 3 data processing functions. Requires significant computing resources for shear estimation codes, galaxy simulations, etc.
  - **Science Working Group**: Requires significant computing resources for setting mission requirements, science analysis of mission archive data, and operations support.

SDC-UK Computing

- The UK is committed to process and host ~10% of the mission data from 2021 – 2027.
  - Early "Science Challenges" and "IT Challenges" will be run to test the pipeline and infrastructure across all 9 SDCs (2018: SC4/5/6, 2019: IT8 and SC7, 2020: IT9 and SC8, 2021: End-to-end test). These will use significant computing in 2020-2021 as end-to-end tests are completed (~1000 core years per year, and ~2PB of data).
  - Our current estimates put the UK data processing commitment at ~2,000 core years per year from 2021-2027, and will ultimately store ~10PB of data by the end of the mission (some of this can be moved to tape/offline storage at different stages, but exact percentages are only speculative at them moment).
  - The final data will be transferred to ESA for hosting the science archive for science users, and most UK hosted data will be deleted at the end of the mission (though some will likely remain as inputs for other Astronomy projects like LSST and SKA, amounts unknown for now).

SDC-UK Compute Requirements

Euclid data processing requires the following:

- **Worker Environment**: customized CentOS7 environment (called EDEN in Euclid parlance) with access to CVMFS repository, local scratch space on worker nodes, and shared storage for storing/moving data products
- **IAL (Infrastructure Abstraction Layer) Service**: in-house developed control service for submitting jobs to DRM queues, managing data movement between archive storage and worker nodes, and ingesting output products into the archive. Requires SSH access to a submission host, and preferably access to a shared storage area visible to worker nodes. Also requires access to the DSS service (described below).
  - This service is currently evaluating a “pilot-jobs” type model for job submission, which may become standard by the end of 2018.
  - One instance of this service runs at each SDC in the Euclid consortium
- **DSS (Data Storage Server) Service**: in-house developed data management service for storing/retrieving/registering files and metadata products into the archive, from the clusters. Works over https/ftp and requires external network connections to all SDCs.
  - A minimum of one instance of this service runs at each SDC in the Euclid consortium
- SSH access for 2-5 operators and deployment/configuration control for the IAL and DSS services.

OU-SHE/LE3 Computing

- IfA in Edinburgh are the lead developers of the Weak Lensing Pipeline component within the overall Euclid pipeline (Andy Taylor is the PI)
  - Aims to map dark matter in the universe by measuring distortions ("shear") of light originating from other galaxies passing through/by large gravity wells.
  - Galaxy images (postage stamps) are extracted from Euclid images and then analysed. Each galaxy will take a minimum of 1 second to process, and there are ~40,000 galaxies per image, and ~40,000 images in the Euclid survey.
  - Full universe galaxy simulations are used to train the algorithms and minimize errors/bias. There is a rough 10 to 1 ratio of simulations to real measurements in order to achieve the accuracy required for Euclid science goals.
  - Development effort is currently performed on a 1,200 core cluster located at the ROE in Edinburgh (managed by me)
- Multiple UK institutions are involved in the LE3 (Level 3) pipeline components. The computing requirements for these are currently quite speculative, some of which will likely require HPC facilities and specialized infrastructure to run (such as ARCHER). More information will be available on these by the end of 2018.

OU-SHE Compute Requirements 2018-2019

- Custom execution environment on worker nodes (CentOS7, CVMFS, Python 3.6.2)(called EDEN 2.0 in Euclid parlance)
- 200GB local storage on worker nodes for scratch space
- 200TB network storage for input/output data products
- Resource requirements split into two main profiles:
  1. Simulations – Multi-threaded Python application with significant memory requirements. Our reference config for this is 44 cores and 112GB RAM
  2. LensMC/Sensitivity testing – Single threaded applications with minor memory requirements and trivial parallelization. Our reference config for this is 1 core and 2GB RAM
- DRM Queue (preferably SLURM)
- Our goal for summer 2018 is to run a simulation of 10^9 galaxies and perform shear estimation on this data set. With our current algorithms, this would require 30,000 cores running for 2 weeks to complete (a single galaxy currently takes ~3-5s to process)
- We are likely to run a single 10^9 galaxy simulation again in 2019 and 2020, and then it would step up to a 10^{10} galaxy simulation in 2021. On top of the extra CPUs required, by that time our storage requirements would grow to ~1PB as well. There could be multiple 10^8 galaxy simulations in the interim (2018-2020) as our algorithms are refined and tested at smaller scale (some of these will run on our existing cluster, but others may require outside resource)

Science Working Group Computing

- Three areas of need:
  1. Simulations, before and after launch (2018-2027)
     - Euclid will require between 10^4 to 10^6 n-body (or better hydrodynamical) simulations per cosmology for estimating observation probabilities and modeling structural changes due to dark energy variations
  2. Requirements specification, before launch (2018-2020)
     - Before launch *requirements* need to be specified to a very high level of precision and accuracy (instrument cannot be altered/tweaked once in orbit)
     - We need to compute expected error bars in order to design to survey optimally
  3. Likelihood analysis after launch (2021-2027)
     - Sample cosmological parameter space using complex approaches, e.g. Bayesian Hierarchical modeling (1000s, 10000s free parameters - National-level HTC required)

SWG Computing Requirements 2018-2020

- **Simulations** - Priority is 10^3 to 10^4 n-body simulations for covariance matrix testing
  - ~100 pilot simulations have been done on "modest Xeon cluster, using 4 nodes with dual Xeon E5520 2.27 GHz per node and 24Gb shared memory" (Kiessling et al.)
  - Totalling ~400 core years per year 2018-2020
  - 500TB to 1PB of storage required if storing all data from simulations
    - Otherwise on-the-fly recalculation required

- **Requirements testing**
  - Totalling ~800 core years each year 2019-2020
  - 500TB to 1PB of storage required during simulations and for a limited time afterwards for analysis

Ideal TK0 Resource Config

- OpenStack service with access for 2-3 project operators who can deploy both short and long term VMs/containers/services. For Euclid, this would include:
  - DRM host running SLURM with self-managed user accounts and a permanent IP address/DNS alias.
  - 0…n worker nodes deployed on demand using a local VM image/Docker repo, attached to the DRM host for queues and account management, dynamically networked and allocated. In a truly ideal scenario the worker deployment would occur elastically as jobs enter the DRM queues, and workers would shutdown after a defined period of idleness.
  - Shared storage visible on all worker nodes, ideally as a POSIX filesystem (using CephFS, or NFS, or…). Local storage on each worker node mounted at /tmp for scratch space.
  - IAL host with a permanent IP address/DNS alias and firewall that can be modified by project operators. Needs access to the shared storage space, the DRM host, and the DSS host
  - DSS host with a permanent IP address/DNS alias and firewall that can be modified by project operators. Needs to be visible from the IAL host, and also to the other DSS servers at other SDCs.
  - Resource allocation would be controlled and monitored through the different queues, with at least one queue for each of the 3 main Euclid work areas (SDC-UK, OU-SHE, SWG). Resource accounting would be determined from usage of the queues.

Looking to the future…

Within Euclid we have learned a lot from seeing how CERN operates and we have even successfully adopted several computing tools/methodologies developed there (CVMFS, Pilot Jobs). We see UKT0 as a means of not only resource sharing, but also knowledge sharing, and we look forward to the opportunities this will provide.