

# The GridPP New User Engagement Programme: Selected Case Studies

**T Whyntie<sup>1</sup>**

<sup>1</sup>Particle Physics Research Centre, School of Physics and Astronomy, Queen Mary  
University of London, Mile End Road, London, E1 4NS, United Kingdom

E-mail: [t.whyntie@qmul.ac.uk](mailto:t.whyntie@qmul.ac.uk)

**Abstract.** A number of selected case studies from the GridPP Collaboration's New User Engagement Programme are presented, featuring new user communities from fields including medical physics, computational biology, space and astrophysics. The services used and developed by GridPP to achieve this engagement are also briefly described.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>LUCID</b>	<b>4</b>
<b>3</b>	<b>HTC for Evolutionary Biology</b>	<b>7</b>
<b>4</b>	<b>Galactic Dynamics (UCLan)</b>	<b>10</b>
<b>5</b>	<b>PRaVDA - HTC for hadron therapy</b>	<b>13</b>
<b>6</b>	<b>The Large Synoptic Survey Telescope</b>	<b>16</b>
<b>7</b>	<b>Monopole hunting at the LHC</b>	<b>20</b>



Except where otherwise noted, this work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/). Please refer to individual figures, tables, etc. for further information about licensing and re-use of other content.

## 1. Introduction

The GridPP Collaboration [1, 2] is a community of particle physicists and computer scientists based in the United Kingdom and at CERN. Drawing on expertise from nineteen UK institutions, its vision is to create, manage and oversee the evolution of the computing infrastructure needed to maintain the UK's position as a world leader in particle physics. It does this by using and actively contributing to the development of open source software, applications and middleware needed to power large-scale distributed computing for particle physics and beyond.

GridPP was originally proposed to provide the UK's contribution to computing required for the Large Hadron Collider [3] (LHC) at CERN. However, GridPP's continually developing infrastructure has since evolved to serve a wide range of user communities in many fields of scientific and engineering endeavour. Indeed, a requirement of GridPP's funding agreement is to make up to 10% of its resources available to non-LHC communities, which as of 2012 was roughly 30,000 logical CPUs (equating to 292,000 HEPSPEC06<sup>†</sup>) and 29 PB of storage.

With that in mind, since 2013 the GridPP New User Engagement Programme has encouraged the development and use of a number of tools and services to enhance the potential for engagement with the Grid. These include:

- **The GridPP CernVM:** The GridPP CernVM provides an out-of-the-box Virtual Machine (VM) instance that is ready to be used as a gateway to the Grid. Based on the Scientific Linux 6 operating system, and using technology developed by CERN's [CernVM group](#) [4], the GridPP CernVM is a contextualised VM image that has instant access to the software and tools required to access GridPP's distributed computing resources.
- **The CernVM File System** (CernVM-FS or CVMFS): allows software to be deployed, managed, and used wherever it is needed on the Grid [5].
- **Ganga:** The Ganga software suite<sup>‡</sup> provides a Python-based interface to the Grid. It was developed to provide a simpler interface to enable physicists to submit their jobs across the Grid and to handle the complete life-cycle of each job. Ganga can be configured to submit and manage jobs on a local machine, a local batch system, or distributed GridPP resources via GridPP DIRAC.
- **GridPP DIRAC:** The GridPP DIRAC service is a DIRAC [6] instance hosted at Imperial College London that has been developed to cater for the multiple VOs that represent the non-LHC users of GridPP resources [7, 8]. In combination with the Ganga software suite, users have at their disposal a powerful set of tools for managing complex distributed computing workflows.

A number of selected case studies, demonstrating how new user communities have been engaged using this toolkit, are listed in Table 1 and presented in the following sections of this document.

<sup>†</sup> See <https://w3.hepik.org/benchmarks/doku.php>

<sup>‡</sup> See <http://ganga.web.cern.ch>

Table 1: A list of the new user community case studies included in this document.

User community	Sector	Compute	Storage	CernVM	CVMFS	DIRAC	Ganga
LUCID (p4)	Space	✓	✓	✓	✓	✓	
HTC for Evolutionary Biology (p7)	Biology	✓	✓				
Galactic Dynamics (UCLan) (p10)	Astrophysics	✓	✓	✓	✓	✓	
PRaVDA - HTC for hadron therapy (p13)	Medicine	✓			✓	✓	✓
The Large Synoptic Survey Telescope (p16)	Astrophysics	✓	✓			✓	✓
Monopole hunting at the LHC (p20)	Physics	✓	✓	✓	✓	✓	✓

## 2. LUCID

### 2.1. Overview

The team behind the satellite-based LUCID experiment have run [GEANT4 simulations](#) of their detectors on the Grid in order to better understand the data collected since its 2014 launch aboard [Surrey Satellite Technology Limited's TechDemoSat-1](#).



Figure 1: The [TechDemoSat-1](#) selfie from July 2014. Image credit: Surrey Satellite Technology Ltd.; please contact them regarding licensing/re-use of this image.

### 2.2. The problem

The Langton Ultimate Cosmic ray Intensity Detector (LUCID) experiment uses five Timepix detectors to measure properties of the space radiation environment in Low Earth Orbit (LEO). These detectors, developed by the [Medipix Collaboration at CERN](#), can measure the flux, direction and energy of incident ionising radiation using silicon pixel technology developed for the LHC. LUCID was designed by students from the Simon Langton School for Boys (Canterbury, UK) and built by [Surrey Satellite Technology Limited](#) (SSTL), and is one of the research projects associated with the [CERN@school programme](#).

As with most modern particle physics experiments, a GEANT4 simulation of LUCID was developed in order to understand the performance of the experiment in the wide range of radiation environments to which it would be exposed. These included the outer electron belts of the Polar Regions and the South Atlantic Anomaly (SAA). The number of simulations required to produce reliable statistics was far beyond what the school computing infrastructure could provide and so it was decided to use the Grid.

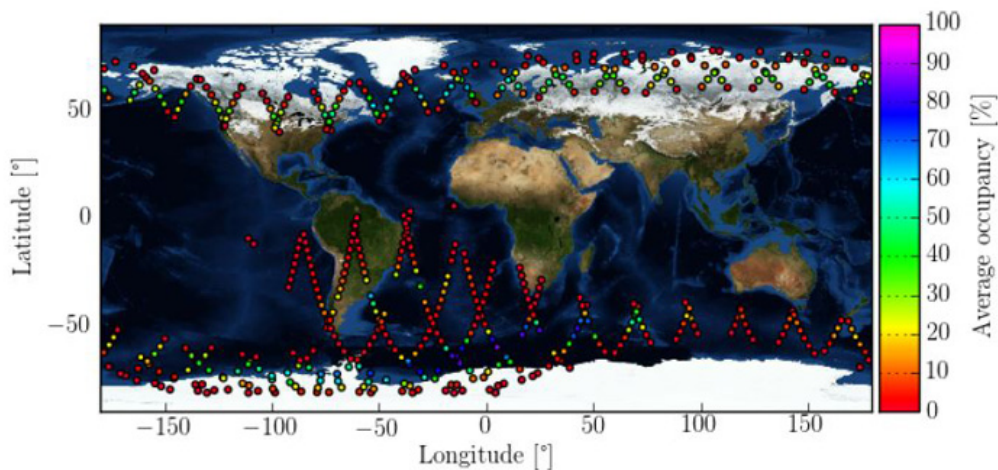


Figure 2: Detector occupancy plot from the LUCID full simulation. Image reproduced from [9]; please contact the publisher for further information regarding the licensing/re-use of this image.

### 2.3. The solution

The GEANT4 simulation of LUCID, built with the Medipix Collaboration's Allpix software suite, was compiled on a Grid-compatible SLC6 machine and deployed using the CERN@school CernVM-FS software repository hosted on RAL's Stratum-0. After determining the number of simulations required, based on the relevant points on LUCID's orbit and the expected particle fluxes from the AP8 (AE8) proton (electron) models, GEANT4 parameter input files were generated and used to submit thousands of jobs to the Grid using the Imperial DIRAC instance. Around 500 orbital points were simulated per run, modelling the interactions of around 5 million source particles per orbital point. Data was stored and managed using the DIRAC File Catalog, with prompt retrieval and analysis made possible using DIRAC's metadata functionality. The results were presented at the International Workshop on Radiation Imaging Detectors 2014 [9].

### 2.4. Further details

#### 2.4.1. Services used

- **CVMFS:** For software deployment and management.
- **CernVM:** For creating a User Interface to the Grid and for software testing.
- **Storage:** The DIRAC File Catalog (DFC) was used for data storage and management.
- **The GridPP DIRAC service:** For Grid job submission and management.
- **Virtual Organisations:** LUCID is part of the CERN@school research programme, and used the cernatschool.org VO accordingly.

#### 2.4.2. *Supporting GridPP sites*

- [UKI-LT2-QMUL](#)
- [UKI-NORTHGRID-LIV-HEP](#)
- [UKI-SCOTGRID-GLASGOW.uk](#)

#### 2.4.3. *Virtual Organisations*

- [cernatschool.org](#) (Ops portal entry)

#### 2.4.4. *Publications*

- Whyntie, T. et al. (2015), “*Full simulation of the LUCID experiment in the Low Earth Orbit radiation environment*”, JINST **10** C03043 DOI: [10.1088/1748-0221/10/03/C03043](#)

#### 2.4.5. *Acknowledgements*

In addition to the support provided by STFC for GridPP, this work was supported by the UK Science and Technology Facilities Council (STFC) grant ST/N00101X/1 and a Special Award from the Royal Commission of the Exhibition of 1851 as part of work with the CERN@school research programme.

### 3. HTC for Evolutionary Biology

#### 3.1. Overview

Whales and dolphins make up the cetaceans, a special group of mammals adapted to life underwater. Their transition from dry land to an aquatic environment is one of the greatest examples of evolutionary adaptation. [Georgia Tsagkogeorga](#) and her colleagues at Queen Mary University of London were interested in how cetaceans evolved in comparison with their closest relatives. In particular, they wanted to know if the molecular adaptations related to aquatic lifestyle appeared before or after whales and dolphins split from the hippos. Using GridPP resources, the team were able to test different evolutionary hypotheses on a vast scale to gain an insight.



Figure 3: A humpback whale. This image is in the Public Domain.

#### 3.2. The problem

To test whether selection pressures acted on a given gene, a series of probabilistic analyses using codon models of evolution need to be performed. This involves fitting two models onto the data: one representing the null neutral evolution hypothesis and a second model assuming selective pressures acting on the gene. By comparing the likelihoods of the null vs. alternative models, the selective regime of the gene in the lineage of interest can be inferred. However, this needed to be done for 11,000 genes across five lineages of aquatic and semi-aquatic mammals.



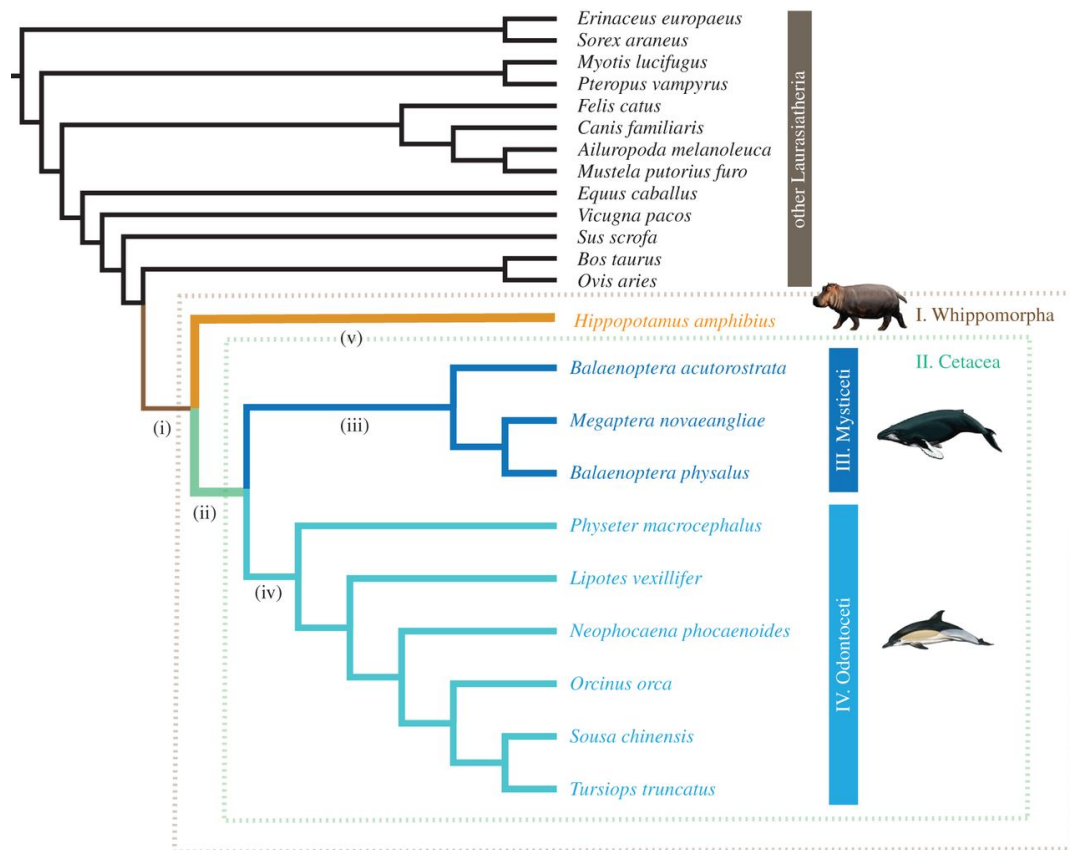


Figure 4: Evolutionary relationships among laurasiatherian mammals as used in molecular evolution analyses performed with GridPP resources. The four clades tested for divergent selection are shown in colour and numbered in uppercase: (I) Whippomorpha (Hippopotamidae + Cetacea); (II) Cetacea; (III) Mysticeti and (IV) Odontoceti. Branches tested for positive selection are numbered in lowercase: (i) Whippomorpha (Hippopotamidae + Cetacea); (ii) Cetacea; (iii) Mysticeti; (iv) Odontoceti and (v) hippo. Reproduced from (Tsagkogeorga, 2015); please refer to the paper regarding licensing/re-use of this image.

### 3.3. The solution

Fortunately the required analyses were highly parallelisable, making them ideal for running on the Grid. With support from the Queen Mary University of London Tier-2 cluster team, Tsagkogeorga and her colleagues were able to run the required 110,000 computations quickly and efficiently. The conclusions, [published in Royal Society Open Science](#), show that the most significant molecular adaptations to aquatic life, for example the evolution of underwater sensory perception or their ability to dive, appeared in cetaceans after the split with the hippos, 55 million years ago.



### 3.4. *What they said*

*“These analyses would have taken months to complete without access to parallel computing. Our study benefited greatly from the 3,000+ CPU cores of QMUL GridPP cluster, by speeding up the analyses to only a few weeks or even days.”*

— Dr [Georgia Tsagkogeorga](#), School of Biological and Chemical Science, QMUL

### 3.5. *Further details*

#### 3.5.1. *Services used*

- **Compute:** The studies were run on the High Throughput Computing (HTC) GridPP cluster at QMUL.
- **Storage:** Results were stored on the QMUL cluster's Lustre-based storage system.

#### 3.5.2. *Supporting GridPP sites*

- [UKI-LT2-QMUL](#)

#### 3.5.3. *Publications*

- Tsagkogeorga, G. et al. (2015), “A phylogenomic analysis of the role and timing of molecular adaptation in the aquatic transition of cetartiodactyl mammals”, R. Soc. open sci. **2** 150156 DOI: [10.1098/rsos.150156](https://doi.org/10.1098/rsos.150156)

#### 3.5.4. *Acknowledgements*

In addition to the support provided by STFC for GridPP, this work was funded by the European Research Council (ERC 1076 Starting grant no. 310482).

## 4. Galactic Dynamics (UCLan)

### 4.1. Overview

The [Galactic Dynamics group](#) at the [University of Central Lancashire \(UCLan\)](#) have been using GridPP resources to simulate orbits of galactic matter in order to better understand how galaxies form.



Figure 5: Barred spiral galaxy NGC 1300. Credit: NASA, ESA, and The Hubble Heritage Team STScI/AURA) 2004; please contact them regarding licensing/re-use of this image.

### 4.2. The problem

Traditionally, galactic simulations require super computers (HPC) to run. This is because the effects of forces such as gravity need to be applied to the entire system being modelled; all objects being modelled will exert a force on all of the other objects present in the simulation. The GalDyn group's research looks at the orbits of particles within galaxies, such as stars and dark matter, and how these are affected by changing a wide variety of orbital parameters. This makes the problem highly parallelisable and ideal for running on the Grid.

### 4.3. The solution

Realising the Grid offered the tools required to run tens of thousands of simulation jobs, and manage the large data sets these jobs would produce, the GalDyn group approached GridPP via the Lancaster Tier-2 site. Following the instructions in the GridPP UserGuide, the GalDyn team were setup with grid certificates and access to the NorthGrid Incubator Virtual Organisation to give them access to GridPP computing resources via the Imperial DIRAC instance. Their custom software, written in C++ and stored in a BitBucket repository, was deployed using a CernVM-FS repository hosted on the GridPP RAL Stratum-0, making it instantly available to any UK Grid site. By using a GridPP CernVM, the GalDyn team were able to locally compile and test their software for Grid use.

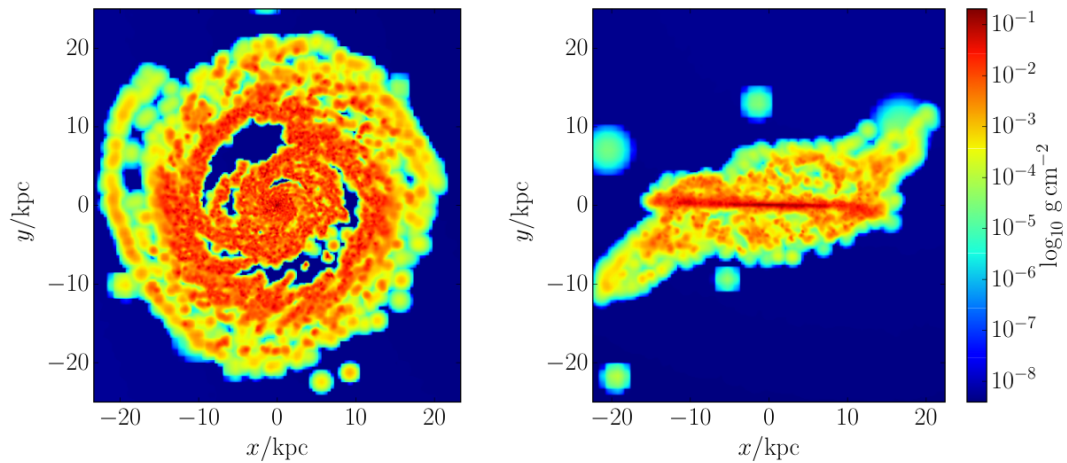


Figure 6: A Smooth Particle Hydrodynamics (SPH) simulation of the gas in a galactic disk as simulated by the GalDyn group. Credit: A. Clarke 2015; please contact them regarding licensing/re-use of this image.

#### 4.4. What they said

*“We can run our code on the Grid and are looking to move into production with tens of thousands of jobs required for our final results. It’s nice to know the facility is there and quick to get setup, instead of having to worry how long it will take to get a load run!”*

— A. Clarke, GalDyn

#### 4.5. Further details

##### 4.5.1. Services used

- **CVMFS:** For software deployment and management.
- **CernVM:** For creating a User Interface to the Grid and software testing.
- **Storage:** The DIRAC File Catalog (DFC) was used for data storage and management.
- **The GridPP DIRAC service:** For Grid job submission and management.
- **Virtual Organisations:** GalDyn made use of the NorthGrid incubator VO during the workflow development phase.

##### 4.5.2. Supporting GridPP sites

- **UKI-NORTHGRID-LANCS-HEP**

- [UKI-NORTHGRID-LIV-HEP](#)

#### 4.5.3. *Virtual Organisations*

- [vo.northgrid.ac.uk](#) (Ops portal entry)

## 5. PRAVDA - HTC for hadron therapy

### 5.1. Overview

Researchers at the [University of Birmingham](#), part of the [Wellcome Trust-funded Proton Radiotherapy Verification and Dosimetry Applications \(PRAVDA\) Consortium](#), are helping to design, build, and commission a device that measures the position and energy of protons used to attack tumours that other forms of cancer treatment can struggle to treat. Such measurements can help ensure that the tumour is being treated optimally, reducing the risk for patients. To optimise their designs, computer simulations of the experimental setup can be run to investigate the effect of changing various parameters without the need for extensive laboratory testing.

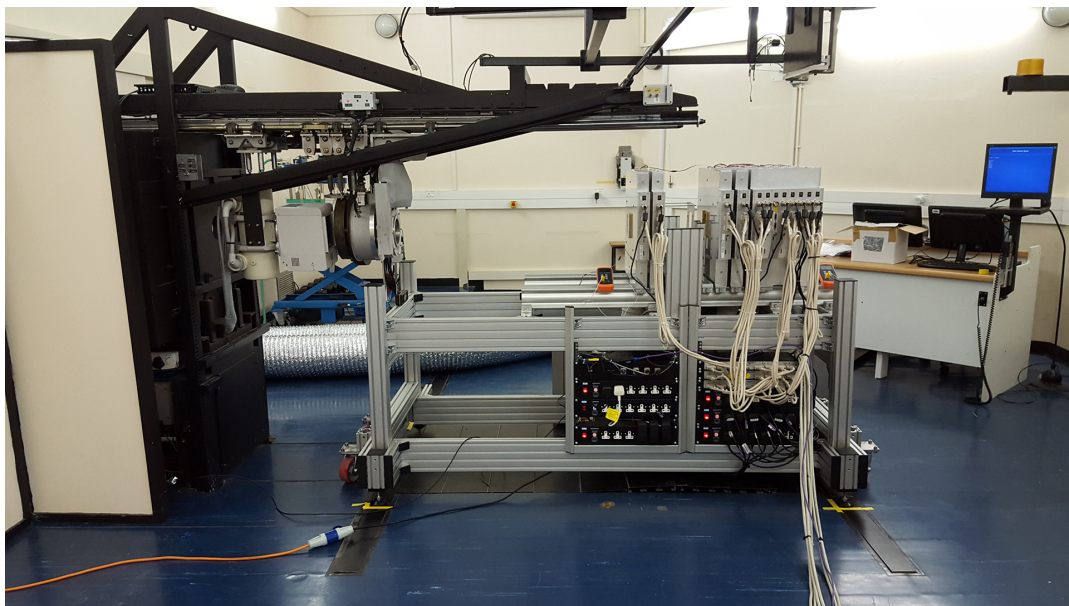


Figure 7: The PRAVDA experimental setup at the iThemba LABS beamline, South Africa. Image credit: T. Price/PRAVDA Consortium 2016; please contact them regarding licensing and re-use of this image.

### 5.2. The problem

The PRAVDA team's experimental setup is modelled using the GEANT4-based Super Simulation (SuSi) software suite. Thousands of parameter settings, covering changes to the detectors used (silicon strips and pixels) and their settings, the types and sources of particles, and the group's novel event reconstruction algorithms, needed to be run with simulations featuring millions of particle interactions. Such a workload proved too much for the university's local computing cluster, and other UK-based compute resources were not optimised for such a massively-parallelisable task on the scale required by PRAVDA.

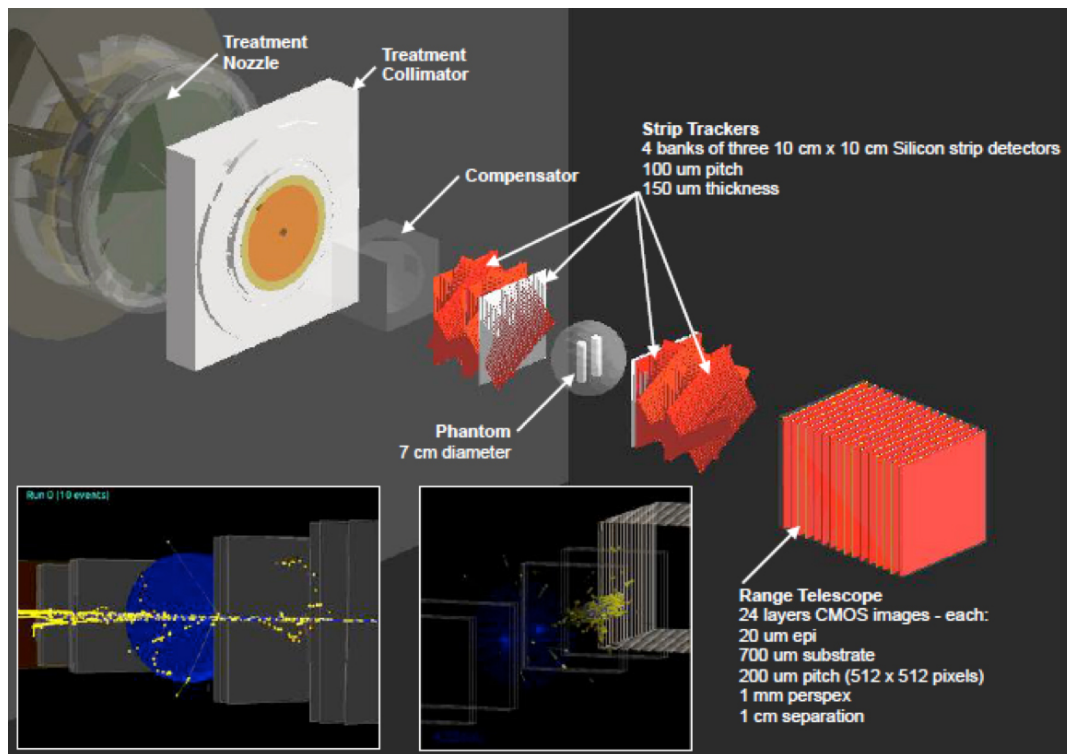


Figure 8: A visualisation of the PRAVDA Super Simulation model. Inset: some example proton beam interactions with GEANT4. Image credit: T. Price/PRAVDA 2015; please contact them regarding licensing and re-use of this image.

### 5.3. The solution

GridPP, who regularly handle billions of GEANT4-based simulations for the Large Hadron Collider (LHC) experiments and other particle physics-based projects, were suggested to the PRAVDA team as potential collaborators. After a successful initial meeting to establish the group's requirements, the PRAVDA SuSi workflow was ported to run on GridPP resources via the GridPP DIRAC service. Simulation jobs were configured and managed using the Ganga interface, and the team's software was distributed to the grid using the CernVM File System (CVMFS). As a result, the team has been able to move to simulations using five times more particles-per-simulation while reducing total run times from weeks to hours.

### 5.4. What they said

*"GridPP has been invaluable to us as a project. There were experts on hand to get me up and running, which was great as time on the project is short. It's easy to quickly run test jobs, and running on 64 cores at a time has vastly reduced wait times. CVMFS has greatly simplified our software deployment too. A huge thank you to those who have supported us!"*

— Dr Tony Price, Uni. Birmingham/PRAVDA

### 5.5. Further details

#### 5.5.1. Services used

- **CVMFS:** CVMFS was used for software deployment and management.
- **Ganga:** Ganga was used for grid job submission and management.
- **The GridPP DIRAC service:** Jobs were managed using the GridPP DIRAC service.
- **Virtual Organisations:** Jobs were run using the GridPP incubator VO.

#### 5.5.2. Supporting GridPP sites

- [UKI-SOUTHGRID-BHAM-HEP.uk](https://ukgrid.bham.ac.uk)

#### 5.5.3. Virtual Organisations

- [gridpp](#) (Ops portal entry)

#### 5.5.4. Publications

- Poludniowski, G. et al. (2015), “Proton radiography and tomography with application to proton therapy”, Br. J. Radiol. **88** 20150134 DOI: [10.1259/bjr.20150134](https://doi.org/10.1259/bjr.20150134)
- Price, T. et al. (2015), “Expected proton signal sizes in the PRaVDA Range Telescope for proton Computed Tomography”, JINST **10** P05013 DOI: [10.1088/1748-0221/10/05/P05013](https://doi.org/10.1088/1748-0221/10/05/P05013)

#### 5.5.5. Acknowledgements

In addition to the support provided by STFC for GridPP, This work was supported by the Wellcome Trust Translation Award Scheme, grant number 098285.



## 6. The Large Synoptic Survey Telescope

### 6.1. Overview

The [Large Synoptic Survey Telescope](#) (LSST) is an international physics facility being built at the Cerro Pachón ridge in north-central Chile. It will use a three-billion-pixel camera to conduct a ten-year survey of the night sky in the visible electromagnetic spectrum. The resulting catalogue of images – an estimated 200 petabytes of data – will enable scientists to answer some of the most [fundamental questions we have about our Universe](#). One such use of this data is being explored by researchers at the [University of Manchester](#), who are studying the effect of dark matter and dark energy on the observed shape of galaxies. By examining how galaxies are distorted by an effect known as [cosmic shear](#), insights can be gained into this missing ninety-five percent of everything we believe exists – the so-called [Dark Universe](#).

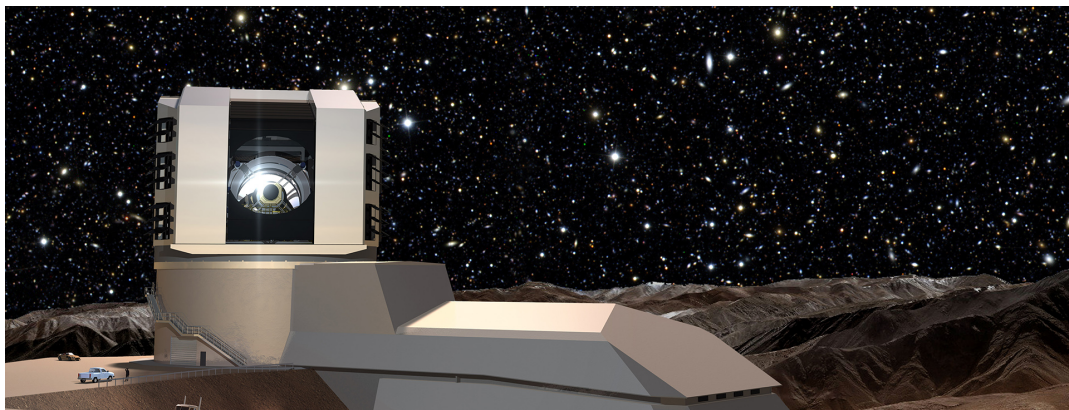


Figure 9: An artist's impression of the [Large Synoptic Survey Telescope](#) (LSST). LSST is under construction at the Cerro Pachón ridge in north-central Chile, and is [due to start scientific operations in 2023](#). Image credit: [Mason Productions Inc./LSST Corporation](#); please contact them regarding licensing/re-use of this image.

### 6.2. The problem

The LSST will produce images of galaxies in a wide variety of frequency bands of the visible electromagnetic spectrum, with each image giving different information about the galaxy's nature and history. In times gone by, the measurements needed to determine properties like [cosmic shear](#) might have been done by hand, or at least with human-supervised computer processing. With the hundreds of thousands of galaxies expected to be observed by LSST, such approaches are unfeasible. Specialised image processing software ([Zuntz 2013](#)) has therefore been developed for use with galaxy images from telescopes like LSST and its predecessors. This can be used to produce maps like those shown in Figure 2. The challenge then becomes one of processing and managing the data for hundreds of thousands of galaxies and extracting scientific results required by LSST researchers and the wider astrophysics community.

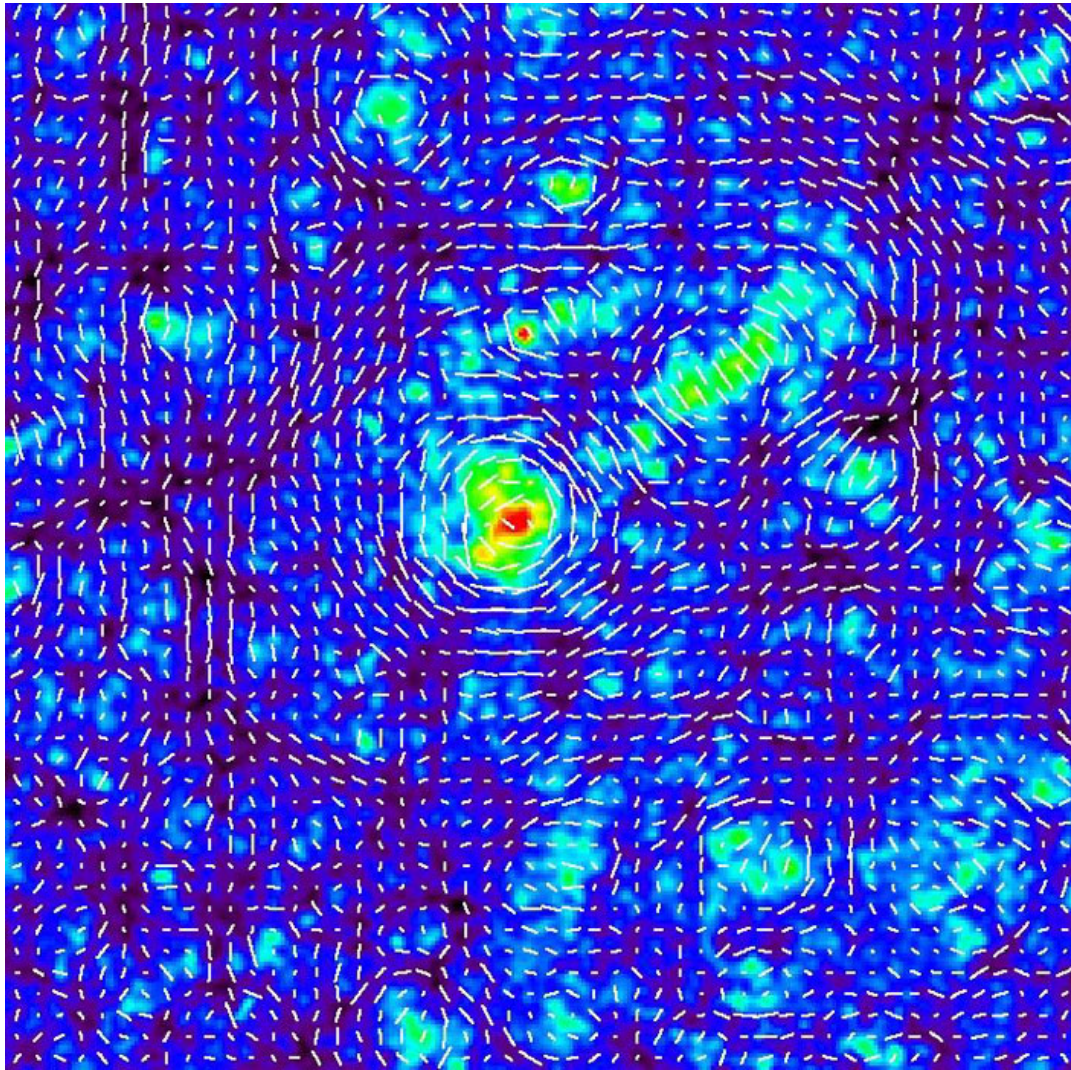


Figure 10: A “whisker plot” showing the magnitude and direction of the [cosmic shear](#) (white tick-marks) caused by the mass distribution (coloured background) from a N-body astrophysical simulation. The degree of [cosmic shear](#) inferred from the analysis of LSST survey data can be used to estimate the effect of dark matter and dark energy in our Universe. Image credit: [T. Hamana/arXiv](#); please contact them regarding licensing/re-use of this image.

### 6.3. The solution

As each galaxy is essentially independent of other galaxies in the catalogue, the image processing workflow itself is highly parallelisable. This makes it an ideal problem to tackle with the kind of High-Throughput Computing (HTP) [resources and infrastructure offered by GridPP](#). In many ways, the data from CERN’s [Large Hadron Collider](#) particle collision events is like that produced by a digital camera (indeed, pixel-based detectors are used near the interaction points) – and GridPP regularly processes billions of such events as part of the [Worldwide LHC Computing Grid \(WLCG\)](#). You can see how the ATLAS experiment does this in the latter part of [this video](#).

A pilot exercise, led by [Joe Zuntz](#) at the [University of Manchester](#) and supported by Senior System Administrator Alessandra Forti, saw the porting of the image analysis workflow to GridPP's distributed computing infrastructure. Data from the [Dark Energy Survey](#) (DES) was used for the pilot. After transferring this data from the US to GridPP Storage Elements, and enabling the [LSST Virtual Organisation](#) on a number of GridPP Tier-2 sites, the IM3SHAPE analysis software package (Zuntz, 2013) was tested on local, grid-friendly client machines to ensure smooth running on the grid. Analysis jobs were then submitted and managed using the Ganga software suite, which is able to coordinate the thousands of individual analyses associated with each batch of galaxies. Initial runs were submitted using Ganga to local grid sites, but the pilot progressed to submission to multiple sites via the [GridPP DIRAC](#) (Distributed Infrastructure with Remote Agent Control) service. The flexibility of Ganga allows both types of submission, which made the transition from local to distributed running significantly easier.

By the end of pilot, Zuntz was able to run the image processing workflow on multiple GridPP sites, regularly submitting thousands of analysis jobs on DES images. He has since requested to continue using GridPP to support his scientific work.

#### 6.4. What they said

*"The pilot is considered to be a success. By the end of the exercise, analyses could be carried out significantly faster than was the case previously (for example, when using highly-contended US-based HPC resources). Thanks are due to the members of the GridPP community for their assistance and support throughout."*

— Dr George Beckett, Project Manager, Uni. Edinburgh/LSST

#### 6.5. Further details

##### 6.5.1. Services used

- **Ganga:** The Ganga software suite was used for grid job submission and management to both local sites (development) and the [GridPP DIRAC](#) service (production).
- **Storage:** Multiple GridPP sites were configured for storing DES image data so that it could be used in [GridPP DIRAC](#) jobs. The [GridPP DIRAC](#) File Catalog (DFC) was used to catalogue and manage the datasets once transferred from the US-based network storage.
- **The [GridPP DIRAC](#) service:** Once local job submission was running successfully, Ganga was configured to use the [GridPP DIRAC](#) service for multi-site workload and data management.
- **Virtual Organisations:** Multiple GridPP sites were configured to enable members of the pre-existing LSST VO to access computing and storage resources.



### 6.5.2. Supporting GridPP sites

- [UKI-LT2-BRUNEL](#)
- [UKI-LT2-IC-HEP](#)
- [UKI-LT2-QMUL](#)
- [UKI-NORTHGRID-LANCS-HEP](#)
- [UKI-NORTHGRID-LIV-HEP](#)
- [UKI-NORTHGRID-MAN-HEP](#)
- [UKI-SCOTGRID-ECDF](#)
- [UKI-SOUTHGRID-OX-HEP](#)
- [UKI-SOUTHGRID-RALPP](#)

### 6.5.3. Virtual Organisations

- [Isst](#) (Ops portal entry)

### 6.5.4. Publications

- Zuntz, J. et al. (2013), “*IM3SHAPE: a maximum likelihood galaxy shear measurement code for cosmic gravitational lensing*”, MNRAS **434** 1604–1618 DOI: [10.1093/mnras/stt1125](https://doi.org/10.1093/mnras/stt1125)
- Kilbinger, M. et al. (2015), “*Cosmology with cosmic shear observations: a review*”, Rep. Prog. Phys. **78** 086901 DOI: [10.1088/0034-4885/78/8/086901](https://doi.org/10.1088/0034-4885/78/8/086901)

### 6.5.5. Acknowledgements

In addition to the support provided by STFC for GridPP, this specific work was supported by the European Research Council in the form of a Starting Grant with number 240672.

## 7. Monopole hunting at the LHC

### 7.1. Overview

The MoEDAL experiment ([Monopole and Exotics Detector at the LHC](#)) is the seventh major experiment at CERN's Large Hadron Collider. It is designed to probe the LHC's particle collisions for signs of Paul Dirac's [hypothesised magnetic monopole](#), as well as other highly ionising signs of new physics that the other LHC experiments cannot easily look for. In order to quickly obtain results from the LHC's 13 TeV run, the MoEDAL Collaboration needed to setup and run millions of [GEANT4-based simulations](#). Thanks to the infrastructure provided by GridPP, this was possible within a matter of days.



Figure 11: The MoEDAL experiment in the VeLo cavern at IP8. Credit: The MoEDAL Collaboration 2016; please contact them regarding licensing and re-use of this image.

### 7.2. The problem

MoEDAL is a relatively new - and relatively small - experiment at the LHC. Housed at Interaction Point 8, it shares the experimental cavern with the [LHCb experiment](#). Despite its size, it has a [comprehensive physics programme](#) and needs to perform full particle transport simulations of the detector subsystems and their surroundings in order to understand how monopoles (or other new physics that might show up through highly-ionising signatures) will interact with MoEDAL's experimental apparatus. Different parameters - such as the monopole mass or magnetic charge - each require hundreds of thousands of events. Local batch systems used by collaboration members were struggling to cope with the scale of computing required for the latest 13 TeV dataset.

### 7.3. The solution

Fortunately, GridPP were able to help. While technically an LHC experiment, the size of the MoEDAL Collaboration meant that the infrastructure provided by the New User Engagement Programme was ideal for quickly getting MoEDAL's software group up and running on the Grid. Using the vo.moedal.org Virtual Organisation and a CVMFS software repository at CERN, the MoEDAL and LHCb software could be run on the Grid using GridPP DIRAC. Ganga provided a user-friendly interface for development and testing of the simulation workflow on a local batch system before switching to the GridPP DIRAC back-end. The scalability of the system, with Ganga able to manage hundreds of jobs at a time, meant that a wide range of parameters could be explored in parallel. In all 36 million events were simulated in a few days using GridPP resources, providing the collaboration with the results they needed for a demanding publication schedule.

### 7.4. What they said

*"Having worked with the Grid before, I knew that GridPP's resources could deliver what we needed in the timescale we had to work with. The fantastic team behind GridPP DIRAC were able to configure our VO at supporting sites with ease and get us up and running with large-scale simulations fast. Ganga's ability to switch between local and Grid running meant that the development and testing phases were made simpler too; by changing one configuration parameter we could switch from running a few local jobs to hundreds on the Grid with ease."*

— Dr Tom Whyntie, Software Coordinator, MoEDAL/QMUL

### 7.5. Further details

#### 7.5.1. Services used

- **CVMFS:** The MoEDAL CVMFS repository was used for deployment and management of the MoEDAL software and access to the LHCb, Ganga and DIRAC software.
- **CernVM:** These provided a Grid-friendly system MoEDAL members to use when not using the CERN lxplus batch system, as CernVMs can be used for running Ganga and DIRAC via CVMFS.
- **Ganga:** Ganga was used for local and batch submission during development of the workflow, and then for job submission and management.
- **Storage:** The DIRAC File Catalog (DFC) was used for storage and management of the LHE simulation files (used as input to the simulation jobs) and output ROOT files containing the simulation results.

- **The GridPP DIRAC service:** GridPP DIRAC was used for Grid job submission and management, via the Ganga interface.
- **Virtual Organisations:** The MoEDAL Virtual Organisation to represent MoEDAL's use of computing resources on the Grid.

#### 7.5.2. Supporting GridPP sites

- [UKI-LT2-QMUL](#)
- [UKI-NORTHGRID-LIV-HEP](#)

#### 7.5.3. Virtual Organisations

- [vo.moedal.org](https://vo.moedal.org) (Ops portal entry)

#### 7.5.4. Publications

- The MoEDAL Collaboration (2016), “Search for magnetic monopoles with the MoEDAL prototype trapping detector in 8 TeV proton-proton collisions at the LHC”, JHEP **2016** 67 DOI: [10.1007/JHEP08\(2016\)067](https://doi.org/10.1007/JHEP08(2016)067)

#### 7.5.5. Acknowledgements

This work was supported by the UK Science and Technology Facilities Council (STFC) via GridPP5 and grant number ST/N00101X/1.



## References

- [1] The GridPP Collaboration. GridPP: Development of the UK Computing Grid for Particle Physics. *J. Phys. G*, 32:N1–N20, 2006.
- [2] D. Britton et al. GridPP: the UK grid for particle physics. *Phil. Trans. R. Soc. A*, 367:2447–2457, 2009.
- [3] L. Evans and P. Bryant. LHC Machine. *JINST*, 3:S08001, 2008.
- [4] G. Ganis et al. Status and Roadmap of CernVM. *J. Phys.: Conf. Ser.*, 664:022018, 2015.
- [5] J. Blomer et al. The Evolution of Global Scale Filesystems for Scientific Software Distribution. *Computing in Science and Engineering*, 17:61–71, 2015.
- [6] Tsaregorodtsev, A. and others. DIRAC pilot framework and the DIRAC Workload Management System. *J. Phys.: Conf. Ser.*, 219:062049, 2010.
- [7] D. Bauer et al. The GridPP DIRAC project: Implementation of a multi-VO DIRAC service. *J. Phys.: Conf. Ser.*, 664:062009, 2015.
- [8] D. Bauer et al. The GridPP DIRAC project - DIRAC for non-LHC communities. *J. Phys.: Conf. Ser.*, 664:062036, 2015.
- [9] T. Whyntie et al. Full simulation of the LUCID experiment in the Low Earth Orbit radiation environment. *JINST*, 10:C03043, 2015.

## Acknowledgements

The author wishes to thank the following people and organisations for their support with the work described here: the European Grid Initiative (EGI) and the Worldwide LHC Computing Grid (wLCG); the CernVM group (CERN) for developing the CernVM and CernVM-FS services that have revolutionised the way users work with software on the Grid; Catalin Condurache for supporting GridPP's CernVM-FS capabilities at the Rutherford Appleton Laboratory (RAL); Daniela Bauer, Simon Fayer (Imperial College London), and Janusz Martyniak (now at RAL) for their tremendous work developing and supporting GridPP DIRAC; the Ganga development team<sup>†</sup> for their work on and support for the Ganga user interface; Sam Skipsey and Jens Jensen from the GridPP Storage group for VO storage-related support; Christopher J. Walker (QMUL), Alessandra Forti (Uni. Manchester) and Andrew Lahiff (RAL) for their support with setting up various new user communities at various points in the programme; Andrew McNab (Uni. Manchester) with technical support with the GridPP website and online user guide<sup>‡</sup>; Dan Taylor, Alex Owen, Cozmin Timis, and Terry Froy (QMUL) for support with development work on the Queen Mary University of London cluster; Jeremy Coles (Uni. Cambridge) for leading GridPP Operations and coordinating new user community activities; Peter Gronbech (Uni. Oxford) for support with Project Management; Suzanne Scott and Louisa Campbell (Uni. Glasgow) for administrative support for Collaboration activities (particularly the GridPP Collaboration meetings where the programme was discussed and developed, as well as welcoming contributions from newly engaged users); David Britton (Uni. Glasgow) for leading the GridPP Collaboration and supporting the vision of the New User Engagement Programme as it developed, and, finally; Steve Lloyd (QMUL) for chairing the GridPP Collaboration Board, providing numerous testing and support mechanisms for Grid activities, and many, many, many fruitful discussions.

This work was supported by the UK Science and Technology Facilities Council (STFC) via the GridPP Collaboration [1, 2] and grant ST/N00101X/1 as part of work with the CERN@school research programme.

<sup>†</sup> See <https://github.com/ganga-devs/>

<sup>‡</sup> See <https://www.gridpp.ac.uk/userguide/>

---

## List of Figures

Note that the images listed here are not subject to the Creative Commons 4.0 license under which this document (“the work”) is issued. For information about licensing and re-use, please contact the image owners as described in the figure captions.

1	The TechDemoSat-1 selfie from July 2014. . . . .	4
2	Detector occupancy plot from the LUCID full simulation. . . . .	5
3	A humpback whale. . . . .	7
4	Evolutionary relationships among laurasiatherian mammals . . . . .	8
5	Barred spiral galaxy NGC 1300. . . . .	10
6	Smooth Particle Hydrodynamics simulation of a galactic disk. . . . .	11
7	The PRaVDA experimental setup. . . . .	13
8	A visualisation of the PRaVDA Super Simulation model. . . . .	14
9	An artist’s impression of the LSST. . . . .	16
10	An example of a whisker plot used in cosmic shear analyses. . . . .	17
11	The MoEDAL experiment in the VeLo cavern at IP8. . . . .	20

---

## Version History

Table 2: Version history.

Version	Description	DOI	Author
1.0	Initial version.	<a href="https://doi.org/10.5281/zenodo.220995">10.5281/zenodo.220995</a>	TW