

Experience with Rucio in the wider HEP community

Martin Barisits¹, Thomas Beermann², David Cameron³, James Alexander Clark^{4,5}, Riccardo Di Maria¹, Gabriele Gaetano Fronzé⁶, Ian Johnson⁷, Mario Lassnig¹, Cédric Serfon⁸, and Eric W Vaandering⁹

¹European Organization for Nuclear Research (CERN), Geneva, Switzerland

²Bergische Universität Wuppertal, Wuppertal, Germany

³University of Oslo, P.b. 1048 Blindern, 0316 Oslo, Norway

⁴California Institute of Technology, CA, USA

⁵Georgia Institute of Technology, GE, USA

⁶Istituto Nazionale di Fisica Nucleare (INFN), Torino, Italy

⁷STFC Rutherford Appleton Laboratory (RAL), Didcot, England

⁸Brookhaven National Laboratory, Upton, NY, USA

⁹Fermi National Accelerator Laboratory, Batavia, IL, USA

Abstract. Managing the data of scientific projects is an increasingly complicated challenge, which was historically met by developing experiment-specific solutions. However, the ever-growing data rates and requirements of even small experiments make this approach very difficult, if not prohibitive. In recent years, the scientific data management system Rucio has evolved into a successful open-source project that is now being used by many scientific communities and organisations. Rucio is incorporating the contributions and expertise of many scientific projects and is offering common features useful to a diverse research community. This article describes the recent experiences in operating Rucio, as well as contributions to the project, by ATLAS, Belle II, CMS, ESCAPE, IGWN, LDMX, Folding@Home, and the UK's Science and Technology Facilities Council (STFC).

1 Introduction

Due to the ever-growing data rates of modern scientific experiments, the need for efficient, scaleable and user-friendly data management has never been bigger. While the processes and workflows creating and processing the data are becoming more complex, the scientific organisations involved are under increasing financial pressure to choose storage technologies that best cope with the large volumes of data. Thus the requirement of a data management system that seamlessly and flexibly integrates a wide landscape of heterogeneous resources, while also satisfying the growing needs of large-scale transfers, metadata handling, and data life-cycle management, is more important than ever.

The open-source scientific data management system Rucio [1] has been developed to meet those objectives. While initially developed to meet the needs of the ATLAS [2] experiment at the Large Hadron Collider (LHC) [3], at CERN, the project has evolved into a fully-fledged community project, with developments and contributions coming from a wide range of scientific projects and organizations. Today, Rucio is a comprehensive solution for



data organization, management, and access which is built on top of common software and tools which makes it easy and convenient to interact with. The guiding principle of Rucio is to give scientific organisations the autonomy to orchestrate their dataflows in a way best fitting their use-cases, while trying to automate as many workflows as possible. Since these objectives are common to many scientific endeavors, Rucio has been adopted, and is now being successfully operated, by an increasing amount of scientific organisations.

In this article we describe the recent efforts made by some of these scientific organisations in deploying, maintaining, and evolving their Rucio instances, as well as their development contributions to the project. To foster this international collaborative effort, the Rucio team hosts annual Rucio Community Workshops, Rucio Coding Camps, as well as weekly meetings where development and DevOps issues are discussed. The remainder of this article highlights the advancements made by ATLAS, Belle II, CMS, ESCAPE, Folding@Home, IGWN, LDMX and the Multi-VO Rucio project. We close the article with a summary and a conclusion.

2 ATLAS



Figure 1. Volume of worldwide ATLAS data managed by Rucio

The deployment model for Rucio in ATLAS is currently in a migration from Puppet-managed VMs to containers running on Kubernetes. Figure 1 shows the volume of data managed, over time, of this deployment. There are several advantages of the Kubernetes implementation: the containers provide a sand-boxed environment which does not collide with system dependencies, the servers and daemons can quickly be scaled to match operational needs, and the initial setup of new Rucio instances becomes much easier and quicker.

There are container images provided on Docker Hub for the servers, daemons, web interface, and clients. These images can also be used directly without Kubernetes. Additionally, there are Helm charts, developed together with CMS, that provide an easy-to-install package for the different Rucio components on Kubernetes. The configurations are hosted on Gitlab at CERN and FluxCD is used to automatically apply them to the clusters.

The ATLAS deployment currently spans three Kubernetes clusters hosted on the CERN Openstack infrastructure: one integration cluster and two production clusters. The integration cluster runs all servers and daemons with a low number of threads and only gets a small fraction of the total workload. There are two production clusters for redundancy and they are equal in terms of their configuration and run the exact same share of the workload. Flux makes it easy to manage the configurations across multiple clusters. For the monitoring of the clusters Prometheus and Grafana are used, which come pre-installed. For logging

and alarming the central CERN IT Elasticsearch and Kibana service is used. All logs are collected using Filebeat, aggregated and enriched in Logstash, and then stored centrally in Elasticsearch.

Every daemon in Rucio generates periodic heartbeats. These heartbeats help to automatically distribute the workload across multiple instances of the same daemon. This aids the process of gradually moving daemons from the Puppet VMs to Kubernetes. Currently, more than half of the workload is already running on Kubernetes. The migration will continue and is expected to be finished in the first half of 2021.

3 Belle II

The Belle II experiment [4] is a B-physics experiment at the SuperKEKB [5] accelerator complex at KEK in Tsukuba (Japan). Belle II has been in data taking mode since 2019.

The computing model of Belle II is based on Dirac [6] and more specifically on an extension of Dirac called BelleDirac that has been responsible for the Workflow Management and Data Management. BelleDirac was well suited for Belle II Data Management's need, but to tackle the massive increase of data during the next years a extensive new developments were foreseen. Therefore, it was decided to transition to Rucio since Rucio provides many of the features needed by Belle II and has demonstrated the scalability at scale beyond Belle II's need.

In order to interface the BelleDirac Workload management part with Rucio, a new Rucio File Catalog plugin in Dirac [7] was developed, as well as synchronization agents between the Dirac Configuration Service and Rucio. In addition, some specific new features needed by Belle II were introduced in Rucio, such as chained subscriptions and the new monitoring infrastructure. These are described in more details in [8].

After more than six months of certification, the transition to Rucio finally occurred in January 2021 during the accelerator winter shutdown. The migration took 4 days of downtime of the Belle II computing activities. During that downtime more than 100M file replicas (see Figure 2) representing 15 PB were imported into the Rucio database corresponding to about 1M rules. No major issues were found after the end of the downtime, and the infrastructure, composed of 2 virtual machines, for the Rucio servers and 2 other virtual machines used for the Rucio daemons was able to handle the load.

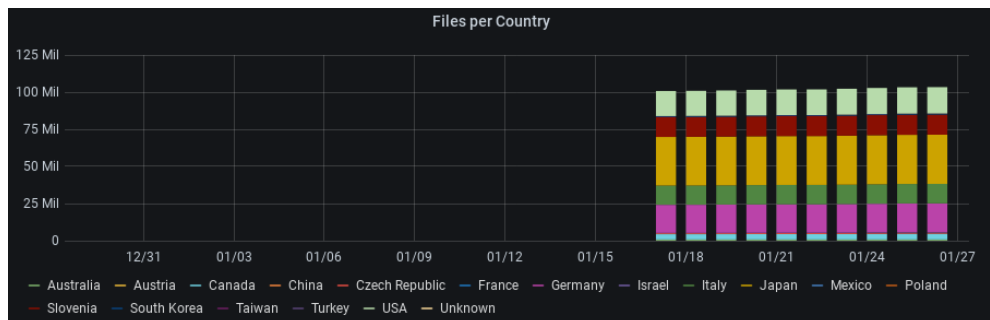


Figure 2. Evolution of the number of files registered into Belle II Rucio instance just after the migration

4 CMS

During 2020, CMS fully transitioned its data management system from legacy systems (PhEDEx [9] and Dyanmo [10]) to Rucio. The experiment began in April by transitioning the NANOAOB data tier to Rucio. NANOAOB was chosen as a first test case because it is small at about 1 kB/event, never needs to be read back in by the production system, and can be quickly regenerated if lost or corrupted. Data locations were synced from PhEDEx into Rucio. Rucio subscriptions were applied to the Rucio datasets to provide a reasonable distribution of data with at least one copy per CMS region. CMS Tier1 and Tier2 sites were divided into four regions based on network connectivity. The CMS production system, WMAgent, was modified to inject NANOAOB directly into Rucio. Finally, all NANOAOB metadata was removed from PhEDEx.

This experience was vital in developing interfaces to Rucio for the rest of the CMS offline software stack as well as for generating confidence for the succeeding steps. No problems were encountered and, for the most part, this change was not noticed by end users. It was also invaluable in giving the CMS data operations team experience with the new tools.

In October of 2020, after further development of software tools which needed to interact with Rucio, CMS began the process of moving all its data under Rucio management. Following a similar process to that developed during the NANOAOB transition, metadata was synced from PhEDEx to Rucio, and WMAgents were individually migrated to begin injecting new data into Rucio. This transition was seamless and was accomplished without downtime. The transition period is illustrated in Figure 3 which shows the smooth transition from PhEDEx to Rucio.

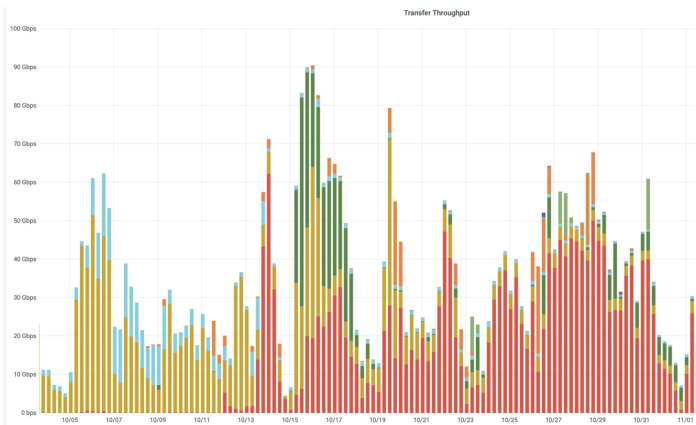


Figure 3. CMS transfer rates over time. In early October the vast majority of data was being transferred by PhEDEx (mustard and cyan) while by late October the vast majority of data was being moved by Rucio (the other colors represent various classes of activity). Only functional tests remained in PhEDEx by November; these activities have also been replaced with Rucio equivalents.

While all data management for CMS is being done by Rucio in 2021, additional work is still needed. The data popularity system CMS was able to bring up with Rucio is not as full-featured as provided by Dynamo, the legacy bespoke system. Also, CMS has not fully implemented consistency checking between the state of storage systems and the state of the Rucio database. These will be the major areas of effort for CMS in 2021.

For CMS, the Rucio installation has always used the Kubernetes and Helm configuration described in Section 2. This method of deployment and configuration has been essential to

CMS to allow it to easily and quickly upgrade as well as to add other third party services to the Rucio cluster.

5 ESCAPE

The European Science Cluster of Astronomy & Particle physics ESFRI research infrastructures (ESCAPE) [11, 12] is a European-funded (Horizon 2020 programme) project aiming to address computing challenges in the context of the European Open Science Cloud [13]. ESCAPE targets particle physics and astronomy facilities and research infrastructures, focusing on the development of solutions to handle data at Exabyte-scale. The needs of the ESCAPE community in terms of data organisation, management and access are fulfilled by a shared-ecosystem architecture of services: the Data Lake.

The infrastructure core is represented by storage services of various software technologies, i.e. dCache [14], DPM [15], EOS [16, 17], StoRM [18], and XRootD [19], orchestrated by Rucio. ESCAPE counts 15 Rucio registered storages, which contribute to a total quota of 821.2 TB and support gridFTP [20], HTTP(S), and xroot access protocols. The redundancy and working-threads of the Rucio daemons, as well as the specific subset chosen, are dictated by the Data Lake implementation. Furthermore, the activity and needs of 9 different experiments and sciences involved in ESCAPE, i.e. ATLAS, CMS, CTA [21], EGO-Virgo [22], FAIR [23], LOFAR [24], LSST [25], MAGIC [26], and SKA [27], are addressed by this specific Rucio configuration. X.509 is used as the standard authentication method, and tokens support is being implemented at the current stage of the project. Moreover, a proof-of-concept for storage Quality of Service (QoS) intelligence has been integrated as a natural consequence of the diversified availability of storage technologies. A continuous testing suite inspects the health of the Data Lake for various areas of the technology stack by checking the infrastructure at Rucio, FTS [28, 29], and GFAL [30] levels separately. The results are displayed through Grafana in ESCAPE monitoring dashboards, hosted by the CERN Agile Infrastructure.

A critical point for the project sustainability and for the successful exportability of the Data Lake model is to demonstrate resource awareness for infrastructure deployment. The Kubernetes cluster hosting the Rucio components uses only 52 virtual CPUs, 104 GiB RAM, and 460 GB local storage, and it is shared with other fundamental ESCAPE services, such as the testing suite.

A full dress rehearsal of the Data Lake has been carried out to assess the infrastructure and the services. The goal of the exercise was to mimic daily activity of ESCAPE sciences and experiments. As an example, ATLAS successfully performed data injection and data movement exercises to reproduce realistic data life-cycles by taking advantage of QoS functionality of the Data Lake.

6 International Gravitational-Wave observatory Network

The International Gravitational-Wave observatory Network (IGWN[31, 32]) handles the computing and storage tasks for the worldwide collaboration on gravitational waves observation composed of LIGO[33], Virgo[34], Kagra[35] and future observatories. By design, the observatories are several thousands kilometers apart in order to improve the GW origin pinpointing capabilities, and data from multiple observatories must be used together to enhance space resolution. Each one of the observatories has its own custodial storage centres for both raw and reconstructed data. Such data must be distributed to the computing centres forming the IGWN computing infrastructure. Rucio is the tool of choice for the handling of

custodial storage replicas, and is being considered as a tool for the transfers of raw data from the experimental sites to the primary custodial storage arrays.

LIGO is already transitioning bulk data management duties from a legacy custom tool to Rucio, which is expected to be the default for the next IGWN observing run (O4), whereas Virgo is still in an evaluation stage, comparing it to the legacy solution before complete transition. In the LIGO instance, the http server and daemons are deployed through Kubernetes on the PRP/Nautilus HyperCluster, while the back-end PostgreSQL database is served from a virtual machine at Caltech. The deployment is provisioned through the continuous integration/deployment system and container repository in the LIGO Gitlab instance. Data transfers are brokered by the File Transfer Service (FTS) system hosted by the Mid-West Tier-2 center at the University of Chicago. The data transfers and basic host metrics are monitored via a Grafana instance which analyzes an internal Rucio ActiveMQ message queue stored in an Elasticsearch database.

The Virgo deployment will also be container-based and will run the Rucio components on a Kubernetes cluster, while the database infrastructure will run in VMs, currently on the INFN Torino Private Cloud, and have offsite live backups in other computing centers.

Within IGWN, the data distribution to computing facilities is currently handled by exporting the namespace via CVMFS and sharing the actual data through a StashCache hierarchy, using the “external data” CVMFS plugin. This is ideal for data aggregation, since each experiment can publish its data in its own StashCache origin, while on the computing nodes all repositories can be mounted alongside each other. The Rucio Fuse-POSIX module was initially developed to allow for the client-side aggregation of multiple Rucio deployments, enabling ready-made multi-experiment initiatives. In addition, the Rucio Fuse-POSIX module can be used to represent the Rucio catalog in a POSIX-like manner, to allow users access to data without requiring changes to their typical approach to storage representation and without the need for additional infrastructural components.

7 LDMX

The Light Dark Matter eXperiment (LDMX) [36] is a fixed-target missing momentum experiment currently being designed to conclusively probe the sub-GeV mass range of several models for thermal-relic dark matter. For the design of the LDMX apparatus, detailed simulations of a large number of events of incoming electrons are required. To simulate a sufficient sample of the most rare background processes, the total number of events needed is similar to the final number of expected data taking events. A distributed computing pilot project was started in late 2019 to evaluate the benefits of sharing computing and storage resources as well as allowing simulation data access across a multi-institute collaboration. The approach taken was to re-use, as much as possible, existing software and systems that were proven to work for other HEP experiments. This naturally led to the choice of Rucio as a distributed data management system for LDMX.

The initial setup of the pilot LDMX distributed computing system consists of 4 computing sites with local storage on shared file systems, which are not accessible outside the site, and a GridFTP storage at SLAC accessible by the collaboration. Simulations run at all 4 sites, the resulting output files are stored locally, and a copy is uploaded to the SLAC GridFTP storage. The use case for Rucio is primarily as a catalog of data locations rather than for data transfer. In addition, Rucio provides a metadata catalog of the produced simulation data. Because LDMX does not have its own metadata system it uses Rucio’s generic JSON metadata system. This system allows storing and querying arbitrary key-value pairs rather than using a fixed database schema, which for a new experiment, allows far greater flexibility in the information it stores.

The Rucio server and a selection of daemons required for accounting and deletion (abacus, reaper and undertaker) are installed from Docker containers on a single lightweight VM. A PostgreSQL database is installed on a second VM. Since Rucio is only used as a catalog, the resources required to run it are very small and no scalability issues have been observed so far with this setup. As of the end of 2020, Rucio is cataloging a total of almost 200TB of data in around 1.7 million files.

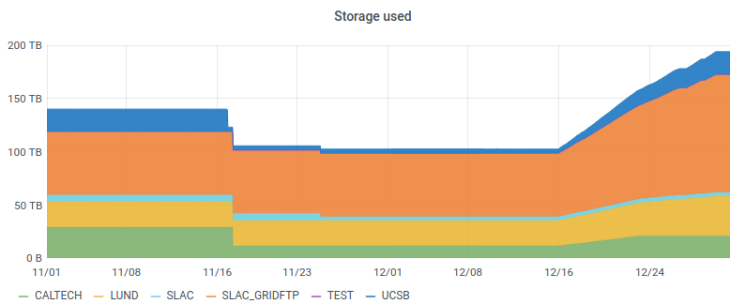


Figure 4. Volume of Rucio-managed storage at each LDMX site in November and December 2020.

Figure 4 is taken from an LDMX dashboard created using Prometheus/Grafana and shows the storage used in each site’s Rucio-managed storage over time in November and December 2020. In mid-November, some data was deleted and in the last two weeks of December a large simulation campaign was run which produced almost 100TB of new data.

8 Folding@Home

Folding@Home is a distributed computing project to study protein dynamics, with the objective to help medical and pharmaceutical professionals to develop effective drugs against a variety of diseases. In the course of the COVID-19 pandemic, a Rucio instance for Folding@Home was setup at CERN with the goal to deliver science data products from the Folding@Home servers to the Hartree High-Performance Computing Centre in the UK for end-user analysis.

The Rucio implementation proved surprisingly difficult due to the heterogeneous work servers from Folding@Home using different distributions of Linux, and the absence of a central authentication and authorisation entity, which is the typical case in the HEP world. The central deployment of Rucio itself was straightforward using the new Kubernetes-based installation mechanism. However, since there are dependencies which require deployment at the Folding@Home servers, a significant portion of work had to be invested to make them easily deployable in this environment, ranging from packaging to fixing bugs related to assumptions on the environment. One example includes the GFAL library, which at the necessary version, required either dependencies unavailable on the Debian distribution, or was not easily compiled. XrootD was deployed as the storage frontend, however, it was quickly discovered that several assumptions on the security infrastructure were made related to certificate expiration, which were not compatible with the newly instituted Folding@Home X.509 environment. This was eventually fixed in XrootD, but required custom compilation of packages.

Since the Hartree HPC is within a secure network area, transfers had to be tunneled through RAL’s Ceph object store (ECHO). A custom service running at Hartree HPC contin-

uously queries Rucio for the datasets available on the RAL storage, and securely moves data from RAL to Hartree.

Eventually, a working setup was demonstrated and that setup is now being moved into production for enduser analysis. Fermilab has also expressed interest to host the Rucio installation for Folding@Home in production for the long-term, as the CERN installation has only been deployed for demonstration purposes. That migration will be accomplished through a move of the Rucio database from CERN to Fermilab, as Fermilab is using the same Kubernetes-based installation mechanism.

9 Multi-VO Rucio

Rucio has shown its utility for VOs such as ATLAS, which have sufficient IT support to deploy and maintain the software. One of the main aims of the multi-VO enhancements to Rucio is to make adoption easier by new or smaller VOs. Such groups usually do not have the ICT effort available to make using Rucio an easy option. The UK's Science and Technology Facilities Council, a research council under UKRI (United Kingdom Research and Innovation), has a mandate to support VOs outside of the WLCG domain. To support VOs which might not otherwise consider installing Rucio on their own resources, STFC instigated the development of multi-VO enhancements to Rucio. The first deployment of these enhancements allows the Rucio service that STFC operate at RAL to support multiple VOs. The development work is an initiative of the UK's IRIS programme [37], which funds the creation of Digital Assets to support research activities in the UK.

The multi-VO design supports multiple VOs within a single Rucio instance. It provides isolation between the artefacts of the different VOs, such as user accounts, RSEs, rules, or DIDs. The technical constraints on the development work included the need to avoid major changes to the Rucio schema, otherwise the database migration process would have involved excessive amounts of data movement for large deployments, such as ATLAS.

The multi-VO enhancements are in place on the Rucio service at STFC RAL. The SKA VO, which was previously using the single-VO Rucio service at STFC RAL, has migrated to using the multi-VO Rucio transparently. A user from CMS is currently assisting with testing the isolation between VOs. Further work integrating the multi-VO DIRAC implementation with multi-VO Rucio is ongoing through collaboration with Imperial College. This will require relaxing constraints on the length of VO names to enable directly using the information obtained from the VOMS proxy certificate.

10 Summary & Conclusion

Over the last few years several scientific experiments have adopted Rucio for full production use. The transition of CMS to Rucio in 2020 was a crucial moment for Rucio. Also more recent migrations, such as the migration of Belle II to Rucio, and the different needs of ESCAPE sciences, pushed for further improvements and contributions to the project.

Recent developments from ATLAS and CMS drove the migration of Rucio from Puppet to Kubernetes-based cloud-like deployments. This allows automatic scaling of the system to operational needs and the builtin monitoring possibilities with Prometheus and Grafana will further simplify the deployment of a Rucio instance. Belle II recently finished their migration to Rucio now managing over 100 million files with the system. Since Belle II uses Dirac as a Workload Management system, a Rucio file catalog plugin was developed and integrated. CMS transitioned to Rucio during 2020 with a step-wise migration. After adapting numerous systems, such as the CMS production system and other parts of the CMS

offline software stack, as well as careful testing, the final migration steps were accomplished in October of 2020. The transition was seamless and accomplished without any downtime. The integration of Rucio in the ESCAPE Data Lake has been proven successful for flagship communities in Astro-particle Physics, Electromagnetic and Gravitational-Wave Astronomy, Particle Physics, and Nuclear Physics that are together pursuing the findable, accessible, interoperable, and reusable, as well as open-access, data principles. The IGWN is starting its transition from legacy solutions to Rucio, letting it handle the whole custodial storage via multiple deployments, federated at client level. Studies on how to handle metadata is still in active progress, as well as the possibility to use Rucio for handling analysis byproducts such as calibrated or weighted data. The LDMX experiment started their Rucio evaluation in 2019 and now hosts over one million files on Rucio. In contrast to ATLAS and CMS which manage object metadata external to Rucio, LDMX is making good use of Rucio's generic metadata capabilities. During the COVID-19 pandemic a Rucio demonstrator instance was setup for the Folding@Home project, to cope with their increasing data handling needs. This demonstrator is soon to be moved into production. The Multi-VO Rucio project successfully evolved Rucio to be able to host multiple virtual organizations, independently, within one Rucio instance. This development is especially aimed at small experiments/communities, who would not have the technical resources to host their own Rucio installation.

In conclusion, the adoption and evaluation of Rucio by several scientific projects has driven the development of several features and contributions, now successfully being utilised by the community. Several new experiments migrated to Rucio in the past year, further expanding the Rucio user base.

References

- [1] M. Barisits *et al.*, “Rucio: Scientific Data Management“, Computing and Software for Big Science, **3**, Springer International Publishing,
- [2] ATLAS Collaboration, “The ATLAS Experiment at the CERN Large Hadron Collider“, JINST **3**, S08003 (2008)
- [3] L Evans and P Bryant, “LHC Machine“, JINST, **3**:S08001,
- [4] T. Abe *et al.*, “Belle II Technical Design Report“, KEK-REPORT-2010-1, arXiv:1011.0352 (2010)
- [5] K. Akai *et al.*, “SuperKEKB Collider“, Nucl. Instrum. Meth. **A 907**, 188-199 (2018)
- [6] Federico Stagni, Andrei Tsaregorodtsev, André Sailer and Christophe Haen, “The DIRAC interware: current, upcoming and planned capabilities and technologies“, EPJ Web Conf. **245** 03035,
- [7] Cédric Serfon *et al.*, “The Rucio File Catalog in Dirac“, these proceedings CHEP2021
- [8] Cédric Serfon *et al.*, “The Integration of Rucio in Belle II“, these proceedings CHEP2021
- [9] J. Rehn, *et al.* “PhEDEx high-throughput data transfer management system“, ISBN: 0230630160 (2006).
- [10] Y. Iiyama, C. Paus, M. Goncharov, “Dynamo — The dynamic data management system for the distributed CMS computing system“, Journal of Physics: Conference Series **898**, 9 (2017)
- [11] ESCAPE, <https://projectescape.eu/> (2021), accessed: 2021-02-20
- [12] R. Bolton, S. Campana, A. Ceccanti, X. Espinal, A. Fkiaras, P. Fuhrmann, Y. Grange, EPJ Web Conf. **245**, “ESCAPE prototypes a data infrastructure for open science“, 04019 (2020)
- [13] EOSC Portal, <https://eosc-portal.eu/> (2021), accessed: 2021-02-20

- [14] P. Fuhrmann, V. Gülzow, *dCache, Storage System for the Future*, in *Euro-Par 2006 Parallel Processing*, edited by W.E. Nagel, W.V. Walter, W. Lehner (Springer Berlin Heidelberg, Berlin, Heidelberg, 2006), pp. 1106–1113
- [15] A. Alvarez, A. Beche, F. Furano, M. Hellmich, O. Keeble, R. Rocha, “DPM: Future Proof Storage“, *Journal of Physics: Conference Series* **396**, 032015 (2012)
- [16] *EOS*, <http://eos.web.cern.ch/> (2021), accessed: 2021-02-20
- [17] A.J. Peters, L. Janyst, “EOS as the present and future solution for data storage at CERN“, *Journal of Physics: Conference Series* **331**, 052015 (2011)
- [18] A. Carbone, L. dell’Agnello, A. Forti, A. Ghiselli, E. Lanciotti, L. Magnoni, M. Mazzucato, R. Santinelli, V. Sapunenko, V. Vagnoni et al., *Performance studies of the StoRM Storage Resource Manager*, in *Third International Conference on e-Science and Grid Computing, e-Science 2007, 10-13 December 2007, Bangalore, India* (IEEE Computer Society, 2007), pp. 423–430
- [19] *XRootD Website*, <https://xrootd.slac.stanford.edu/> (2021), accessed: 2021-02-20
- [20] *GridFTP*, <https://en.wikipedia.org/wiki/GridFTP/> (2021), accessed: 2021-02-20
- [21] M. Actis *et al.* [CTA Consortium], “Design concepts for the Cherenkov Telescope Array CTA: An advanced facility for ground-based high-energy gamma-ray astronomy“, *Exper. Astron.* **32** (2011), 193-316
- [22] *EGO-Virgo*, <http://public.virgo-gw.eu/the-virgo-collaboration/> (2021), accessed: 2021-02-23
- [23] P. Spiller and G. Franchetti, “The FAIR accelerator project at GSI“, *Nucl. Instrum. Meth. A* **561** (2006), 305-309
- [24] H. J. A. Rottgering, R. Braun, P. D. Barthel, M. P. van Haarlem, G. K. Miley, R. Morganti and I. Snellen, “LOFAR - Opening up a new window on the Universe“, [arXiv:astro-ph/0610596](https://arxiv.org/abs/astro-ph/0610596)
- [25] J. A. Tyson [LSST], “Large synoptic survey telescope: Overview” *Proc. SPIE Int. Soc. Opt. Eng.* **4836** (2002), 10-20
- [26] E. Lorenz, M. Martinez, “The MAGIC telescope“, *Astronomy & Geophysics* **46**, 6.21 (2005)
- [27] *SKA*, <https://www.skatelescope.org/> (2021), accessed: 2021-02-23
- [28] *FTS*, <https://fts.web.cern.ch/> (2021), accessed: 2021-02-20
- [29] A.A. Ayllon, M. Salichos, M.K. Simon, O. Keeble, “FTS3: New Data Movement Service For WLCG“, *Journal of Physics: Conference Series* **513**, 032081 (2014)
- [30] *GFAL2*, <https://dmc-docs.web.cern.ch/dmc-docs/gfal2/gfal2.html/> (2021), accessed: 2021-02-20
- [31] S. Bagnasco, “Virgo and Gravitational-Wave Computing in Europe” *EPJ Web Conf.* **245** 07050 (2020)
- [32] B. P. Abbott *et al.* [KAGRA, LIGO Scientific and VIRGO], “Prospects for Observing and Localizing Gravitational-Wave Transients with Advanced LIGO, Advanced Virgo and KAGRA,” *Living Rev. Rel.* **21** (2018) no.1, 3
- [33] J. Aasi *et al.* [LIGO Scientific], “Advanced LIGO,” *Class. Quant. Grav.* **32** (2015), 074001
- [34] F. Acernese *et al.* [VIRGO], “Advanced Virgo: a second-generation interferometric gravitational wave detector,” *Class. Quant. Grav.* **32** (2015) no.2, 024001
- [35] T. Akutsu *et al.* [KAGRA], “Overview of KAGRA: Detector design and construction history,” [arXiv:2005.05574](https://arxiv.org/abs/2005.05574)
- [36] T. Åkesson *et al.*, “Light Dark Matter eXperiment (LDMX)“, *FERMILAB-PUB-18-324-A*, [arXiv:1808.05219](https://arxiv.org/abs/1808.05219) (2018)

[37] “IRIS - Supporting Digital Research Infrastructure“, URL: <https://iris.ac.uk/> (2021)