UK Computing in DUNE

Andrew McNab

University of Manchester

Overview

- UK computing organisation in DUNE
- Computing capacity delivered to DUNE
- Computing Conceptual Design Report
- 2022 Data Challenge
- UK Development activities

DUNE UK Computing organisation

- Computing = hardware at sites and middleware which connects it
  - So not physics software. Not LArSoft etc.
- Development funded as part of WP1 of the UK construction project
  - At Edinburgh, Manchester, RAL-PPD
  - Non-funded institutes (Bristol, Imperial, Lancaster, QMUL, …) also provide computing capacity and support, attend ops meetings etc
- Monthly meetings to
  - Report status of development work
  - Review use of UK CPU and storage capacity by DUNE
- UK also contributes
  - DUNE Computing Contributions Board chair (Peter Clarke)
  - DUNE Computing International Technical Lead (Andrew McNab)

Leverage as a theme

- One of the themes of DUNE UK computing is the amount of leverage we are getting
  - Software (eg Rucio) and tools (eg GGUS tickets) from outside that we extend/adapt rather than write from scratch
  - CPU/storage capacity funded by IRIS that we apply for annually
  - Site-level operations that are funded by GridPP
- All this requires some DUNE UK effort to give them the necessary DUNE “flavour” and maintain that as things change
  - But we supply a lot more to DUNE globally than is funded by the computing part of the UK construction project
- LBNC noted the success of computing leverage in last March’s report, and this process was led by DUNE UK people

Computing capacity delivered

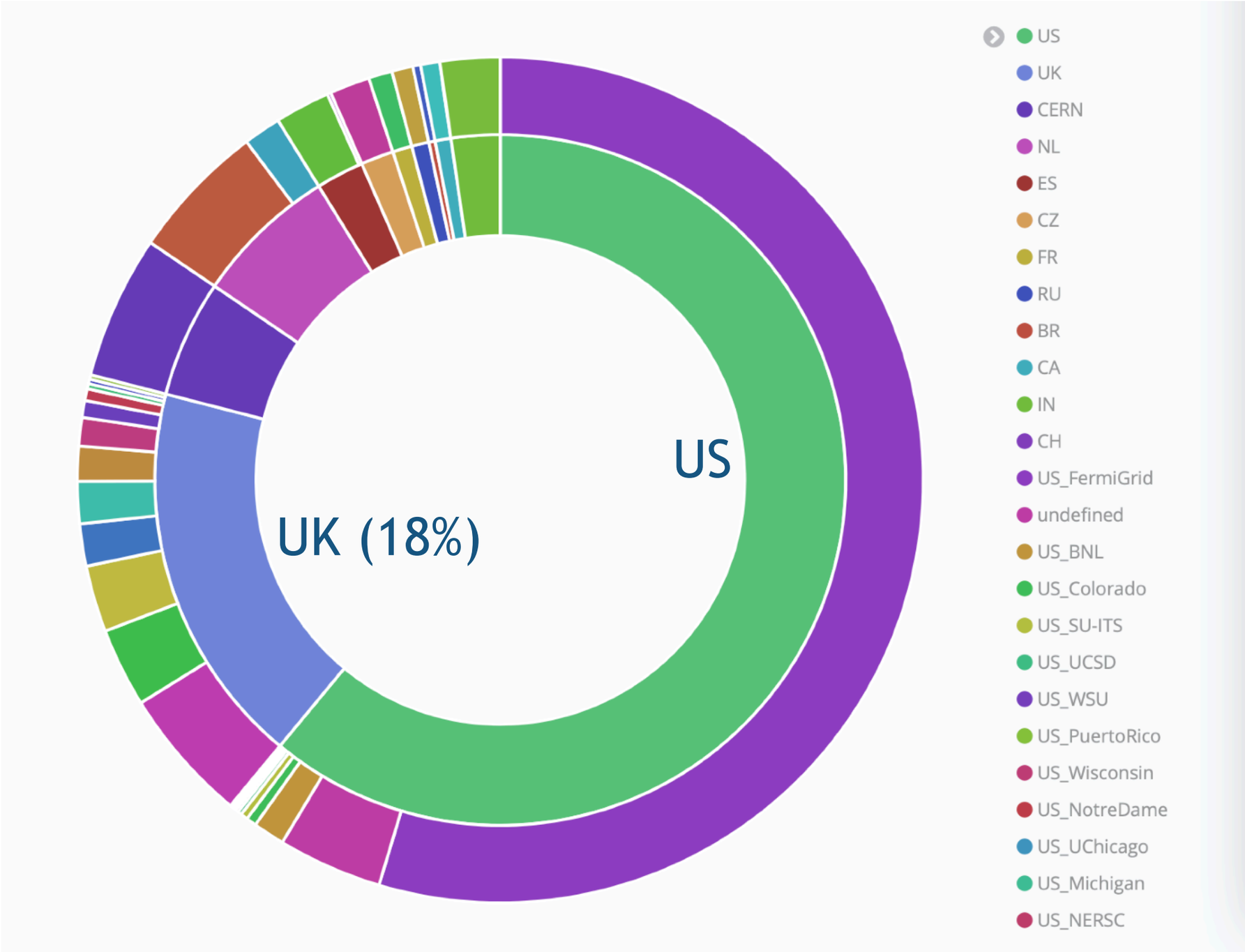
DUNE jobs for 2022

MARS jobs at FNAL are excluded

But all user jobs included

Still a strong bias towards running at Fermilab by default

But for production, the picture is changing as new tools are being used …

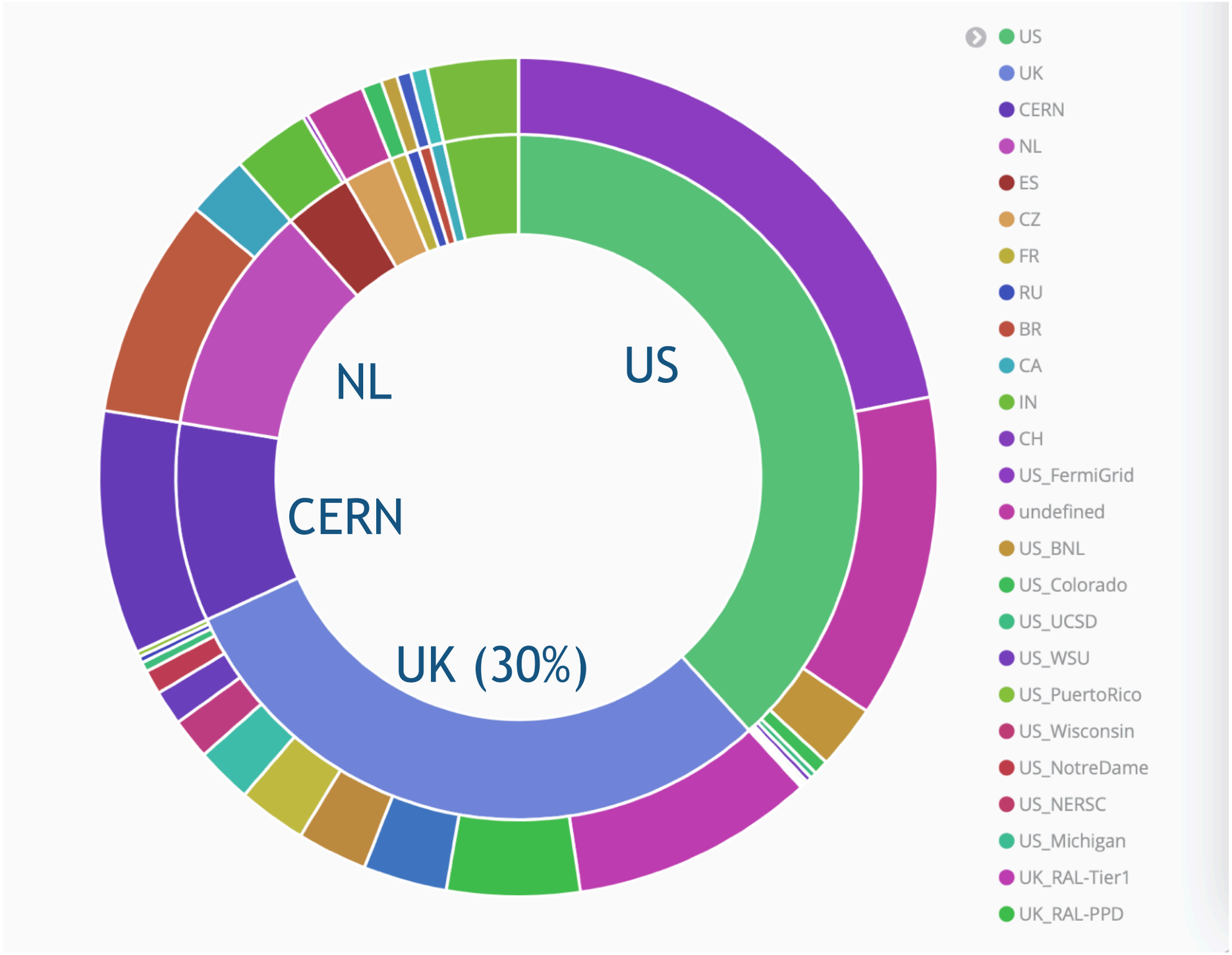


DUNE production jobs for 2022

Larger memory requirements enabled at UK sites to attract jobs

Broader set of sites being targeted

2022 Data Challenge in Q4 was testing systems with automated matching of jobs to best site



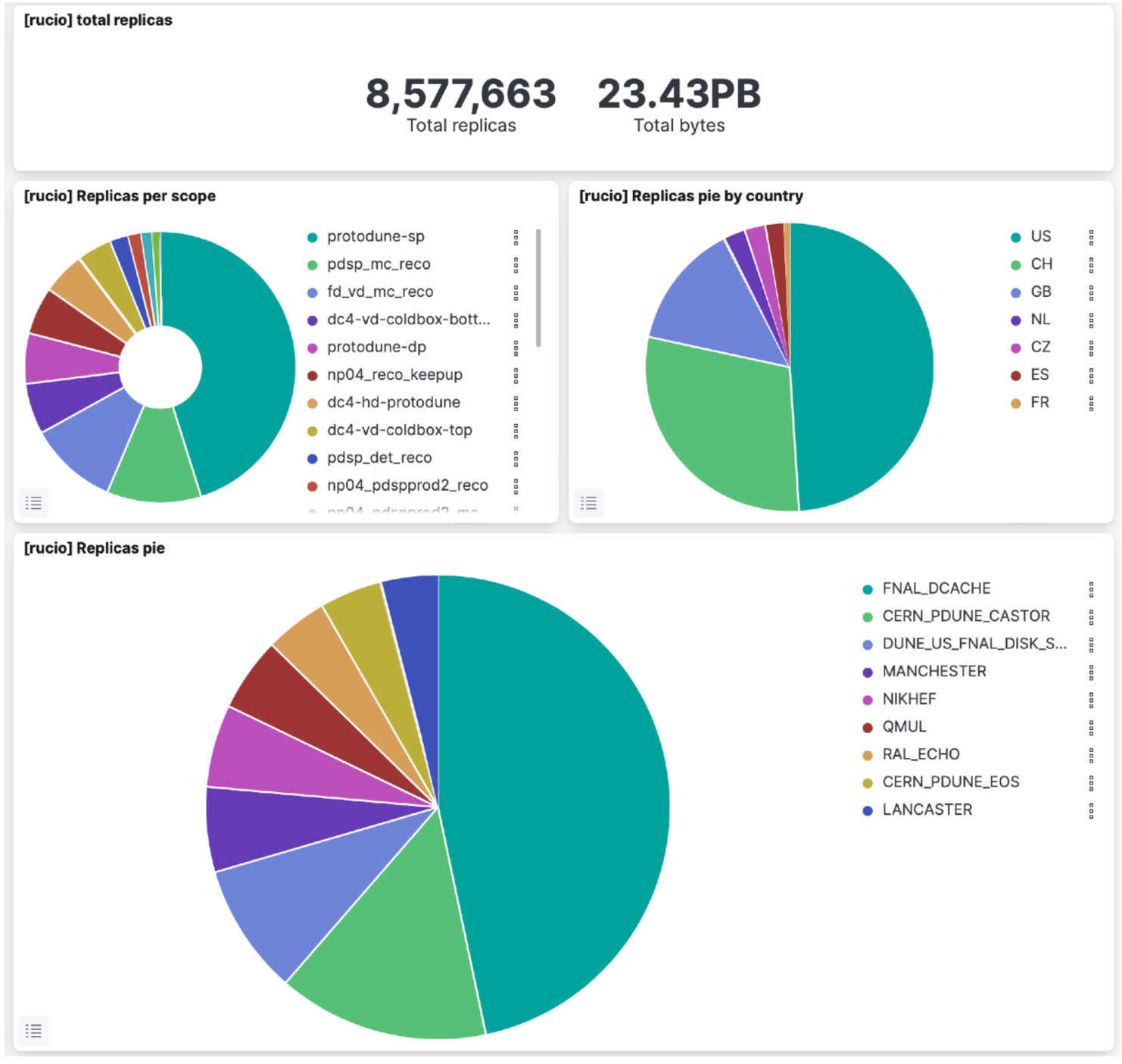
DUNE storage usage status

Screenshot of the storage dashboard developed by Edinburgh as part of DUNE UK

US and CERN dominate but a lot of that is CERN tape or tape-backed dCache at Fermilab

Snapshot of usage NOT space provided. So skewed by any mass deletions

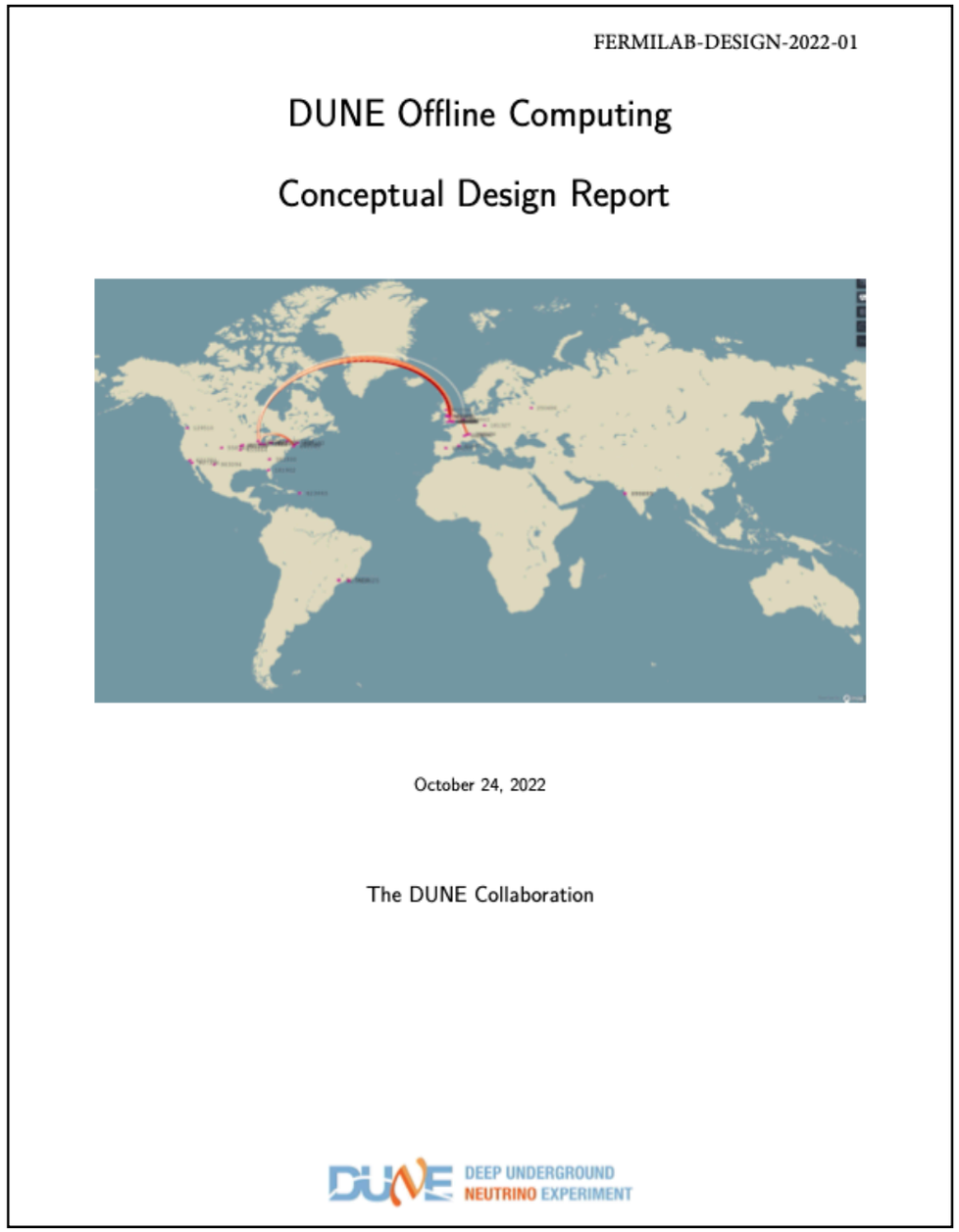
But UK sites are comparable to persistent disk at FNAL and CERN



Computing CDR

DUNE computing CDR

- Finally finalised in October
- Very useful process since it provoked discussions / meetings / workshops to help make decisions
- Consistent and credible design for DUNE computing now “in print”
- A lot of UK contributions
- It is very detailed in places
  - “It’s quite like a TDR”
- Clear path of evolution from where we are to the CDR picture
  - Like SAM to Rucio etc



FERMILAB-DESIGN-2022-01

DUNE Offline Computing

Conceptual Design Report

October 24, 2022

The DUNE Collaboration

DUNE DEEP UNDERGROUND NEUTRINO EXPERIMENT

2022 Data Challenge

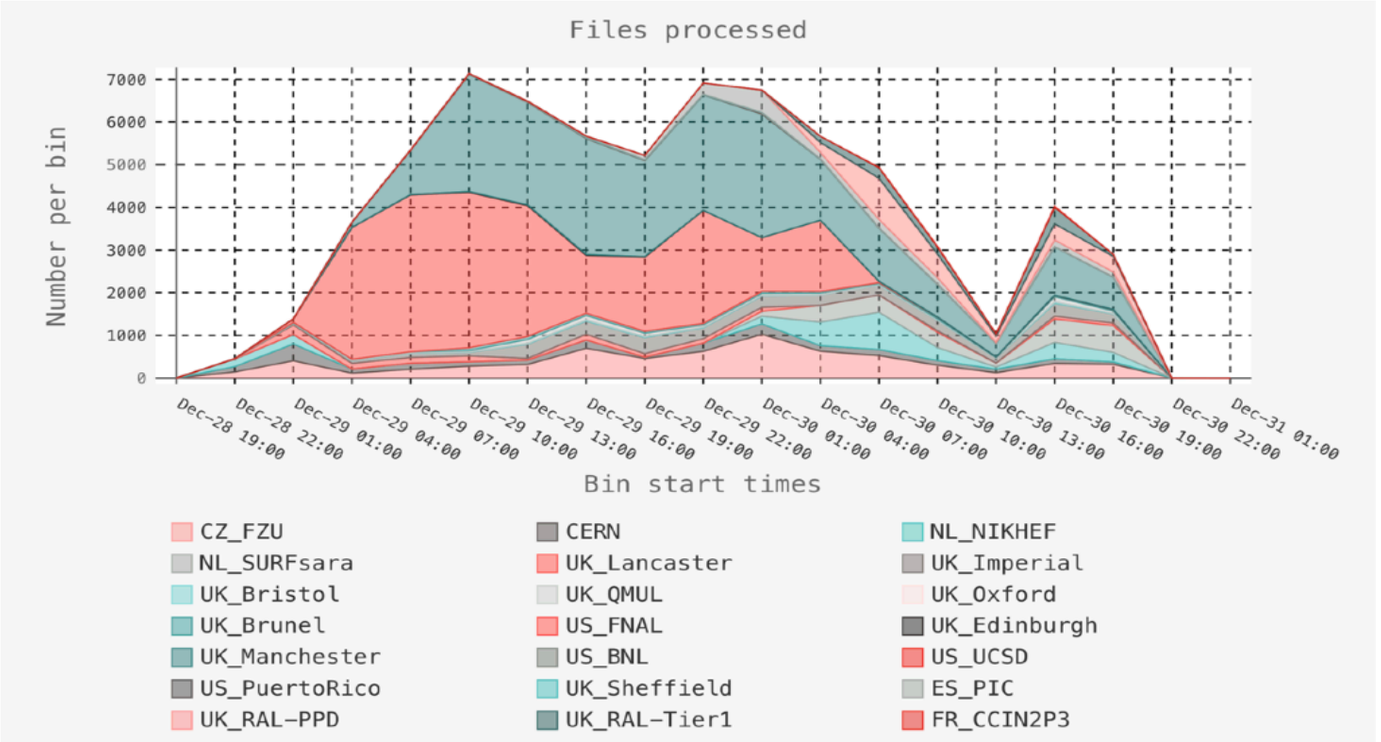
2022 Data Challenge

- A two phase data challenge was done last year to check readiness of new components for ProtoDUNE II data taking
- Phase I, Data Transfers: summer
  - Registering and copying files from CERN to FNAL
  - Relied on Rucio file catalogue rather than SAM
  - Demonstrated ability to reliably transfer data from ProtoDUNE II from CERN to Fermilab at sufficient rate
- Phase II, Data Processing: autumn/winter
  - Processed data at multiple sites, using Rucio file catalogue to discover where input files were
  - The Workflow System from DUNE UK was used to run jobs “near” unprocessed files to maximise efficiency

2022 DC: Phase 2

- One of the charts from the Workflow System dashboard
  - Two other productions running at the same time as this
- Nearest unprocessed file determined which production ran at each site, moment by moment
- Total rate sufficient for jobs processing ProtoDUNE II data

- For reconstruction the location isn't that significant
- But for analysis can increase job efficiency by a factor of 10 by running jobs near input files



UK development activities

UK development activities

- We have two broad areas of development responsibility within DUNE Computing Consortium
- First, Data Management (Edinburgh)
  - This is centred on Rucio which DUNE uses to keep track of where copies of files are on the grid
    - Rucio is a well-established standalone project, but we add core functionality DUNE needs and DUNE plugins
- Second, Workflow System (Manchester and RAL-PPD)
  - Discovering lists of files to process, how to process them, and making sure the files are processed at the “right” site
- Plus other smaller projects: eg RAL-PPD extended ETF for DUNE

Rucio development for DUNE

- (Text from James Perry, Edinburgh)
- Core Rucio work needed by DUNE
  - Simplifying S3 protocol dependencies, and debugging and fixing a problem with S3 Rucio Storage Elements in Multi VO mode
  - Enforcing version check on policy packages to prevent breakages when Rucio API changes
- Improvements to DUNE policy package:
  - Updating to work with newer Rucio versions
  - Using MetaCat Python API instead of direct REST calls (with help from Igor Mandrichenko)
  - New SURL algorithm that queries MetaCat, for use with tape storage sites
- Further improvements to Rucio requested by DUNE are being designed, some as a result of the data challenge experience

DUNE Data Management

- (Text from Wenlong Yuan, Edinburgh)
- Rucio monitoring @ Edinburgh is “in production”. It is useful to trace storage allocation and usage, Rucio file replicas status, Rucio file transfers status, etc.
- Stashcache @ Edinburgh works well, a talk about it on CVMFS workshop (https://indico.cern.ch/event/1079490/timetable/#27-a-new-uk-stashcache-at-edin)
- Commission https third-party-copy transfer is completed. Now all RSEs can use https TPC in Rucio transfers. Commission and tests on Antares tape at RAL Tier-1 have been done, Antares is ready to be used. Now waiting for DUNE transfers data to Antares
- Tested and commission DUNE RSEs. Brought SURFSara, ES_PIC, IN2P3 dCache online, and Antares tape at RAL Tier-1 is ready to be used.
- Traced ‘rucio upload’ transfers failures from some RSEs for data challenge phase II. Figured out xrootd url prefix issue for DPM, StoRM storage element, and CERN EOS authenticate issue with DUNE VOMS server
- Investigated and applied WLCG SRR (Storage Resource Reporting) json files for all DUNE RSEs, which is good for Storage allocation and usage monitoring
- Plans: make SAM/workflow system monitoring work; explore how to make XCache + Rucio work (could be the first DUNE XCache Site)

Rucio Monitoring for DUNE

[rucio] Mock _Total dids

5,064,868 DIDs

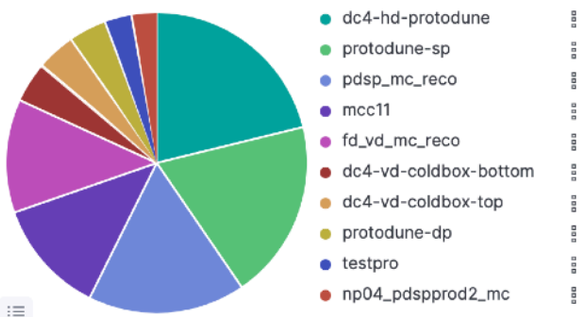
10.2PB Total bytes

[rucio] total replicas

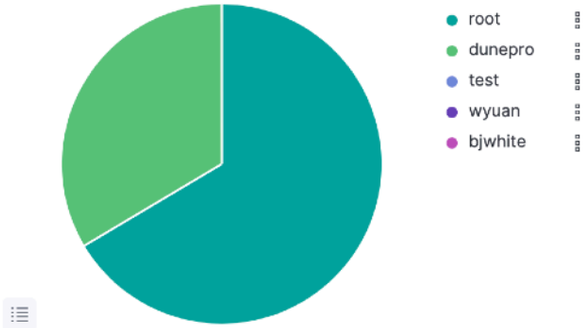
8,577,663 Total replicas

23.43PB Total bytes

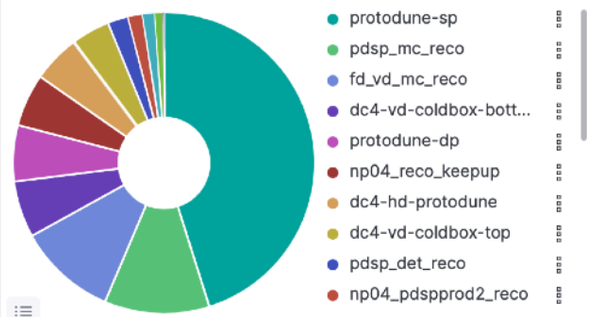
[rucio] DIDs per scope



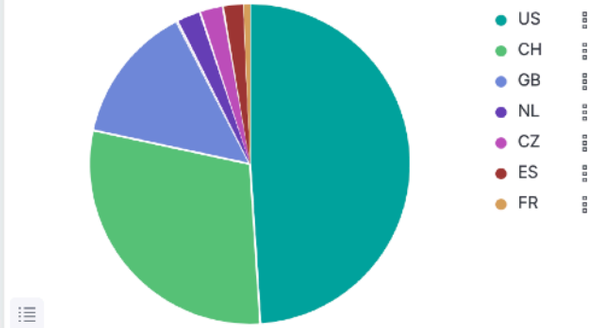
[rucio] DIDs per account



[rucio] Replicas per scope



[rucio] Replicas pie by country



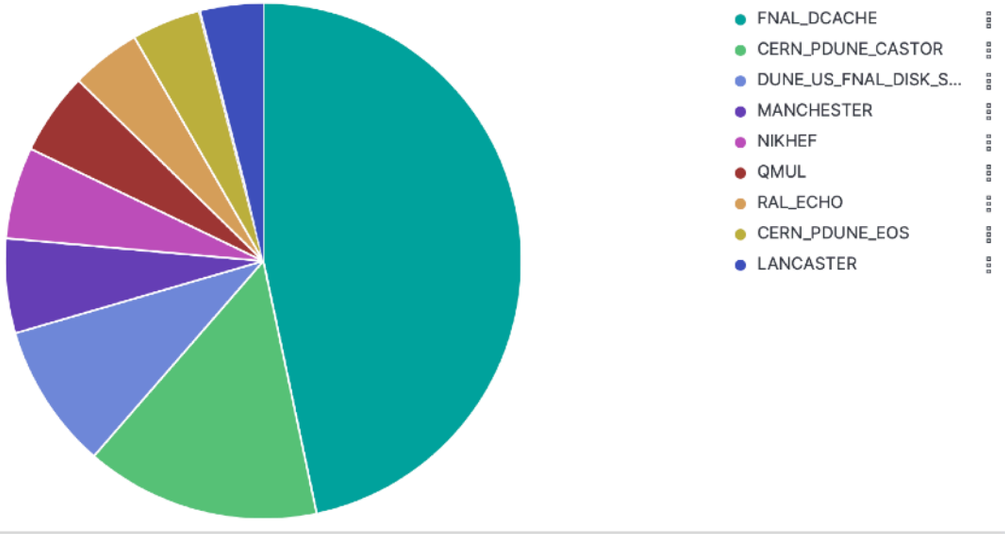
[rucio] Replicas per site

Export

RSE	Replicas	Total bytes
FNAL_DCACHE	3,478,029	10.17PB
CERN_PDUNE_CASTOR	1,094,266	5.15PB
DUNE_US_FNAL_DISK_STAGE	678,718	800.8TB
MANCHESTER	440,876	993.79TB
NIKHEF	428,323	542.3TB
QMUL	384,115	806.92TB
RAL_ECHO	325,541	846.06TB
CERN_PDUNE_EOS	320,371	1.1PB

1 2 3

[rucio] Replicas pie



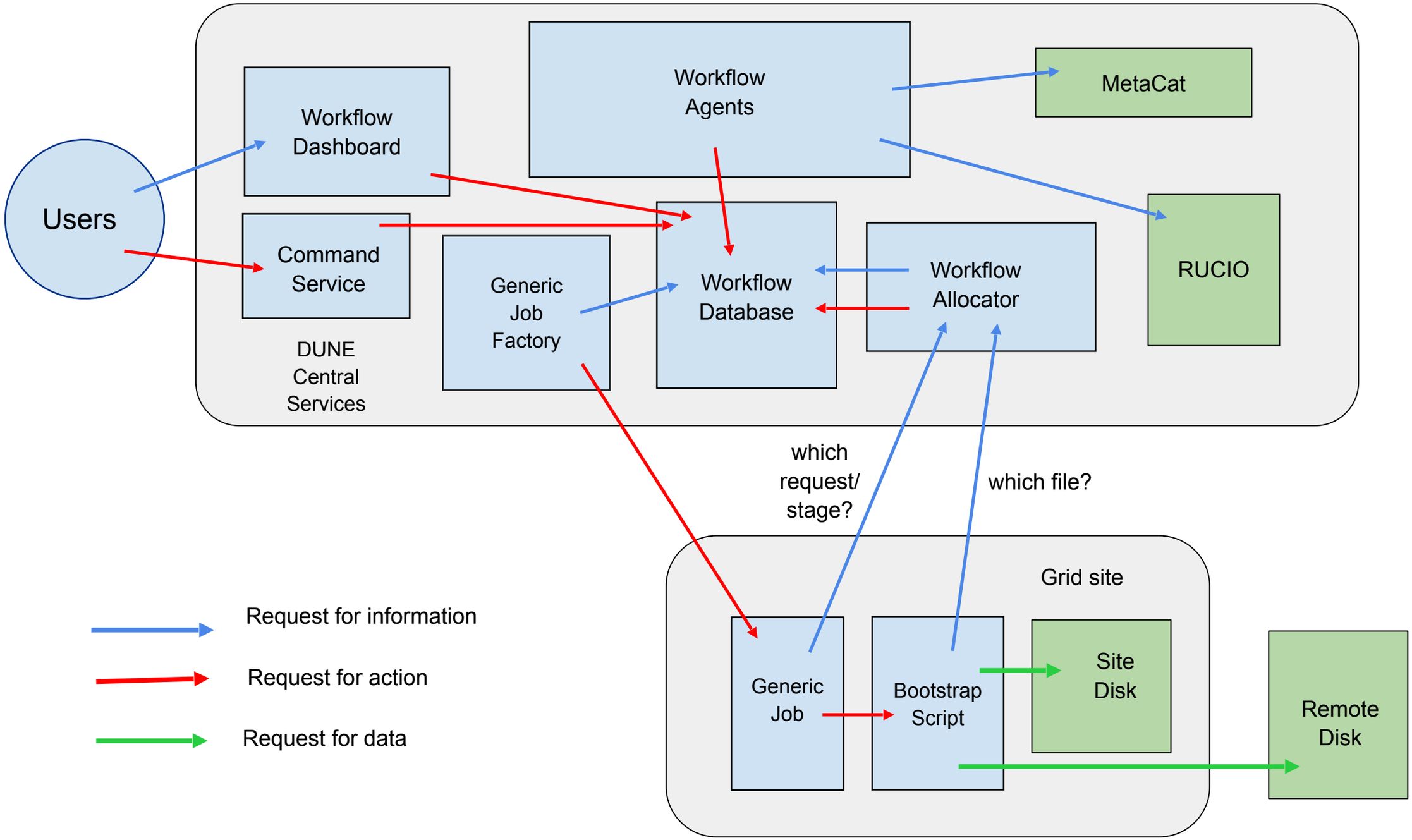
Workflow System

- Hosted on RAL cloud by RAL-PPD
  - Have begun process of containerising the services and agents
- Manchester developed the system validated in the 2022 Data Challenge
- Uses a “just in time” philosophy throughout
  - Defer decisions as late as possible so can respond to downtimes etc
- Matches jobs from workflows to sites
  - using locations of unprocessed files
  - for multiple productions / user requests in parallel
- Relies on Rucio to find file replicas and uses MetaCat queries to find which files to process:
  - `files from dc4:dc4 where core.run_type='dc4-hd-protodune' and core.runs>507141256 limit 20631`

Workflow System

- Workflow System gathers information about the DUNE sites and storages from multiple sources
  - Mainly OSG pilot factories which know about sites
  - and Rucio which knows about storages
- Agents run in the background to prepare workflow requests for processing
  - Gathering file lists and looking up their replicas
  - Looking for stalled jobs and retrying the files they were working on
- The system also submits tests jobs to all sites twice a day
  - Check site are working, and tests reading/writing to all storages from all sites
  - Automated Workflow Testing dashboard presents this to ops team

Workflow System components



Summary

- Approach of identifying key areas where we can have lots of impact via leverage is working out
- The UK is making a significant and visible contribution to the CPU and storage supplied to DUNE
  - Underpinned by STFC funding for hardware and staff via IRIS and GridPP
- We are active members of the Computing Consortium, providing a Technical Lead (Andrew McNab) and the Chair of the Computing Contributions Board (Peter Clarke)
- As part of the construction project, we are providing development effort and leadership in Data Management and in Workflow
- The resulting systems were key components of the successful 2022 Data Challenge in advance of ProtoDUNE II data taking