XLZD@Boulby Workpackage 5: Computing

Daniela Bauer, Davide Costanzo

Imperial College London

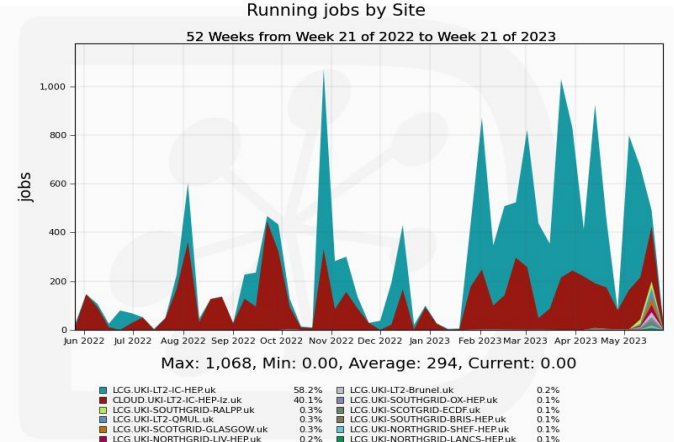
The University Of Sheffield.

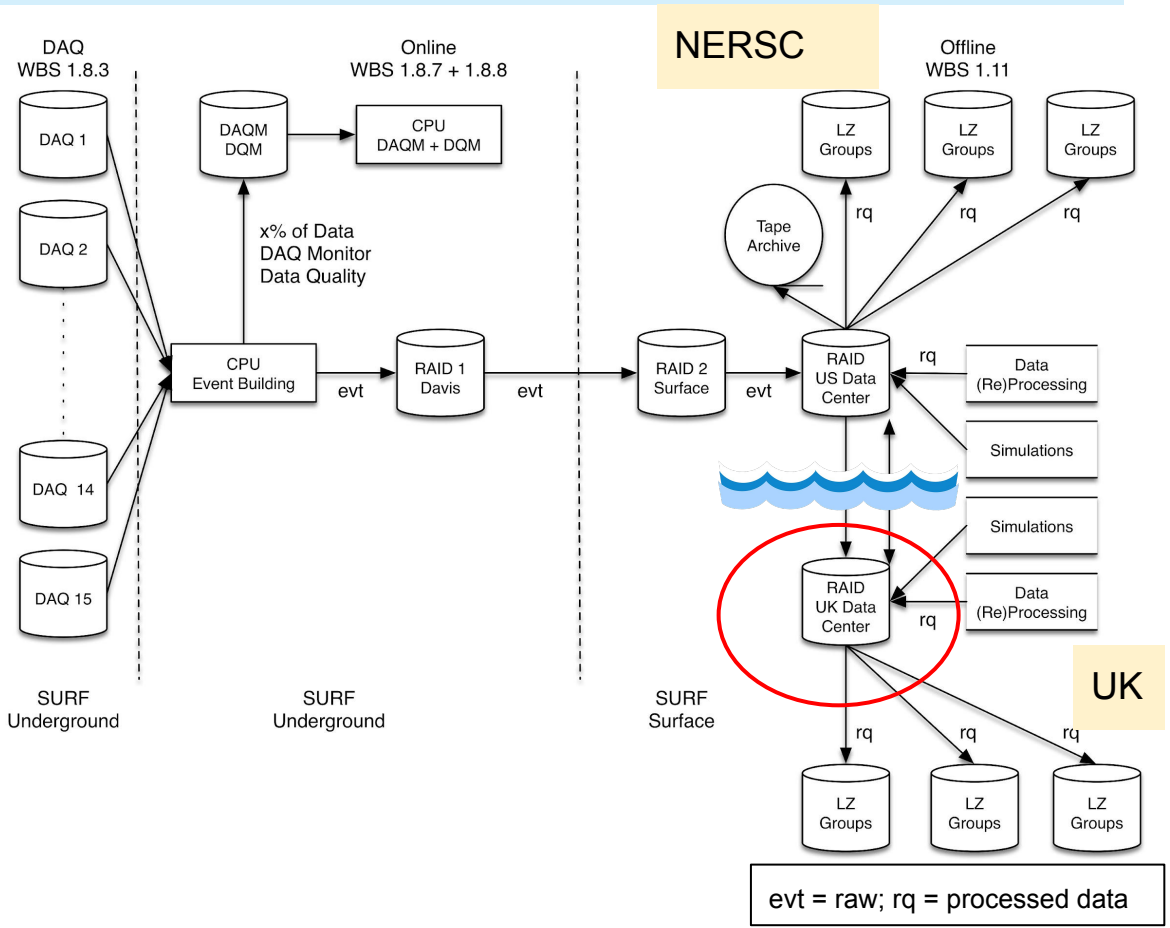
Introduction

- Dark matter computing in the UK: Where are we now ?
  - Context: Computing for physics (STFC) in the UK
- XLZD@Boulby Computing - defining the scope
  - Infrastructure: Data centres, Networking, Compute & Storage Hardware (above and underground)
  - Software Infrastructure: Where does computing stop and physics start ?

Dark matter computing in the UK: The LZ Computing Model

- LZ has two computing centres, each holding a complete copy of the data.
- The UK data centre is hosted at Imperial College.
  - 7 PB (and growing) of LZ storage
  - Compute and majority of storage provided through either GridPP or IRIS.





Context: Computing for STFC Science in the UK

- **GridPP:** A collaboration of 19 institutions providing primarily computing for the LHC experiments. LZ computing is supported under the “other VO” remit, Imperial College as the LZ host facility is a member of GridPP: https://www.gridpp.ac.uk/
  - GridPP also provides opportunistic capacity at other GridPP sites (RAL, Sheffield, Bristol, etc) to LZ.
- **IRIS:** Established in 2018 to create and develop the digital research infrastructure needed to allow UKRI to continue to play a leading role in global projects, e.g. SKA or DUNE. LZ is an IRIS Science Partner, Imperial College is an IRIS provider: https://www.iris.ac.uk/
- **JISC:** Provides networking to UK education and research facilities: https://www.jisc.ac.uk/

Imperial College London

The University Of Sheffield.

WP5 XLZD@Boulby: Scope for “preliminary activity”

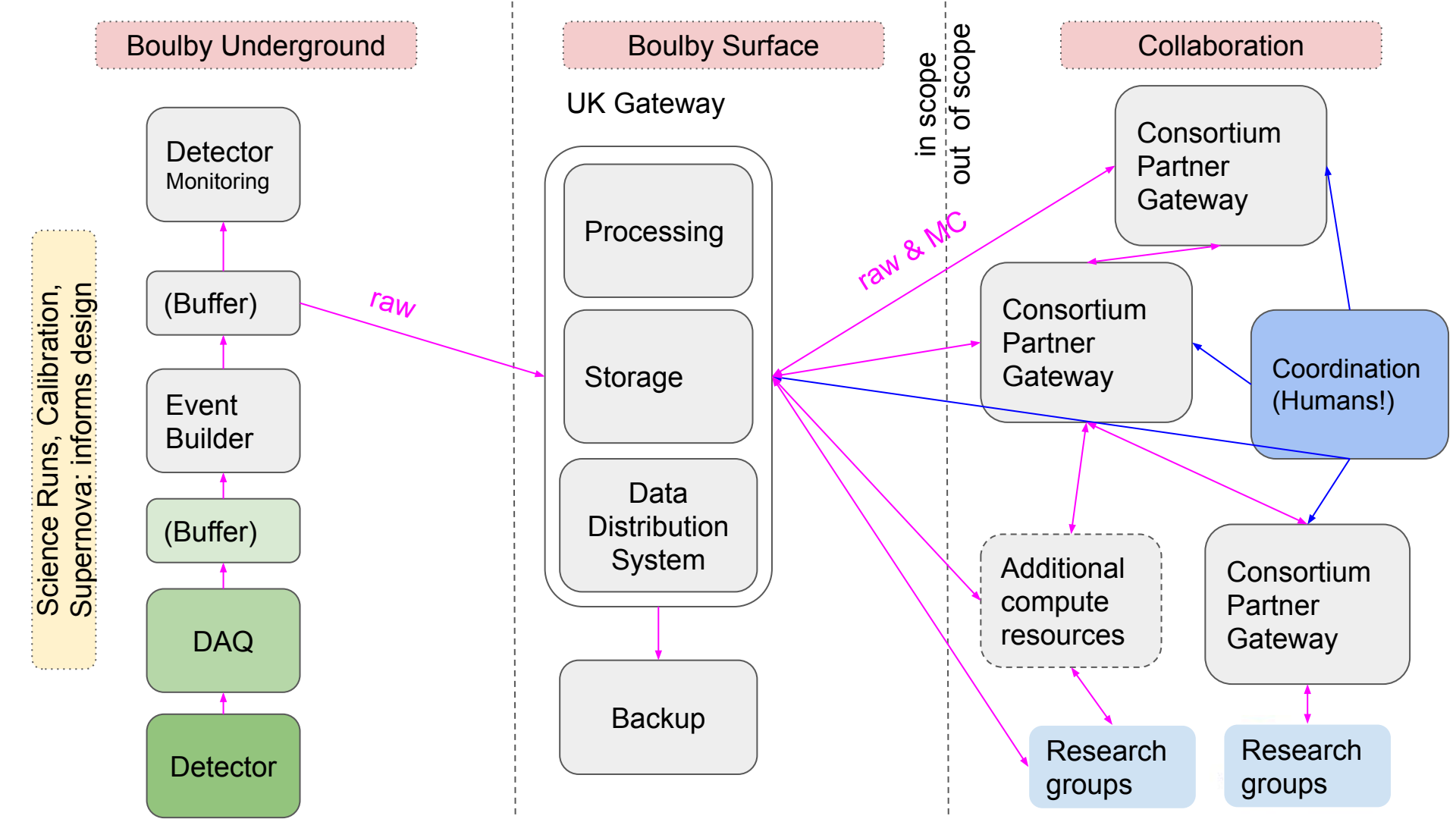
Immediate needs:

- Provide and coordinate computing needs for the design phase (simulation).
  - This can be achieved using the established LZ channels.
- Provide a software framework for the design phase.
  - Use LZ as a starting point, but with the understanding that this will not reflect the final design.
  - Note that this does not involve ‘physics’ frameworks.

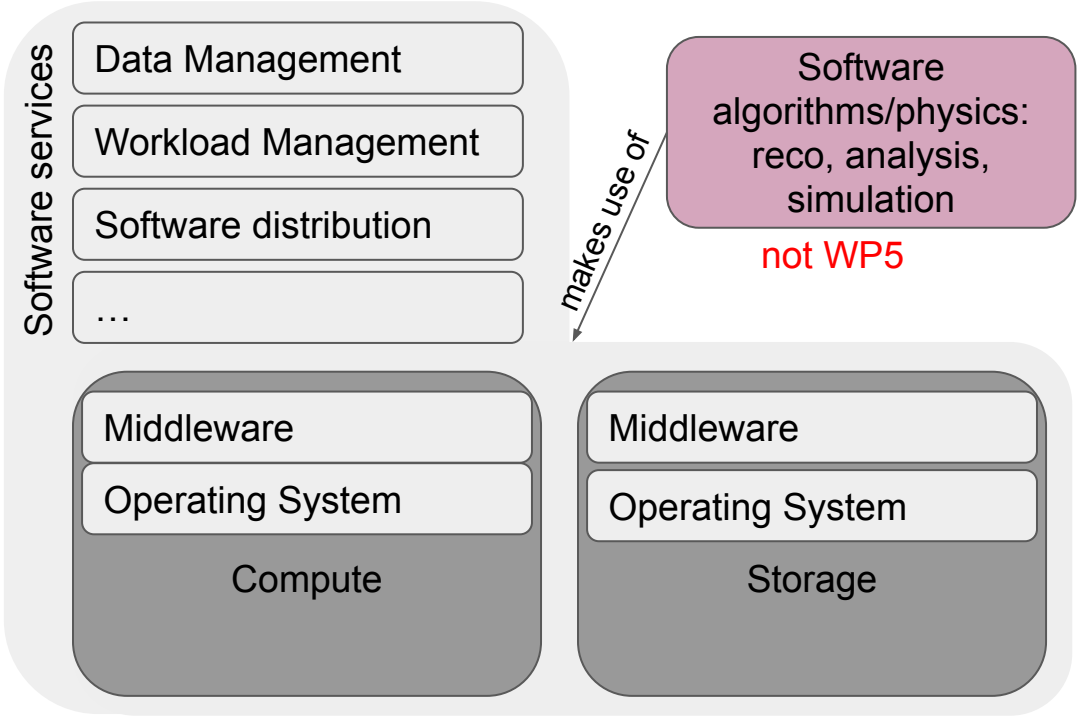
Design/Scoping/Costing:

- Design computing facilities needed:
  - Take into account that XLZD is an international collaboration.
  - In practice this means there will be an element of distributed computing.
- Data pipeline: Underground to Surface, Storage, Data distribution etc
- Data Centre Development @Boulby/North

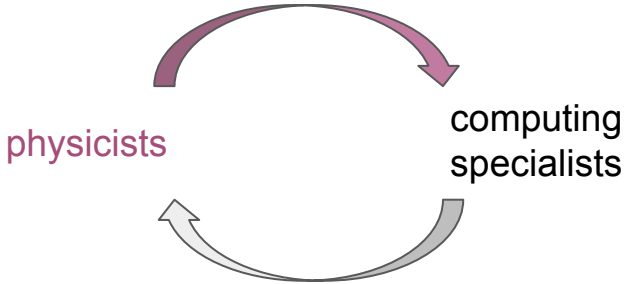
Computing is not an independent part of the experiment. It is very much led by the physics requirements.


Boulby Underground
Boulby Surface
Collaboration
UK Gateway
in scope
out of scope
Science Runs, Calibration, Supernova: informs design
Detector Monitoring
(Buffer)
Event Builder
(Buffer)
DAQ
Detector
raw
Processing
Storage
Data Distribution System
Backup
raw & MC
Consortium Partner Gateway
Consortium Partner Gateway
Coordination (Humans!)
Additional compute resources
Consortium Partner Gateway
Research groups
Research groups

Infrastructure: Software Level - an overview for physicists



Data rates (including MC).
Retention policies.
Turn around times.
Specific hardware requirements for algorithm (GPUs, memory).



Assist in determining requirements, taking into account external constraints (budget, security, maintainability, etc).
Design and deploy computing infrastructure.
Assist in interacting with infrastructure.

Imperial College London

WP5
Solutions are typically experiment agnostic

Networking

Good networking is crucial to any large scale computing activity.

Boulby (Surface) is well placed for this:

- Currently 1x 1 Gbps connection, upgrade to 1x 10 Gbps in progress, completion expected within ~3 to 6 months.
- Upgrade to additional 1x 10G connection for resilience: approved by JANET, awaiting groundworks around mine site, expected within ~12 to 24 months.
- Possibility for 1x 100G connection – being explored for a separate STFC project. Indication from JISC that this can be completed within ~36 months (i.e. by the end of the preliminary activity). WP5 would provide input.
- Additional 1x 100G link for resilience – will need further discussion with JANET to identify a resilient physical path – expected to be possible within the normal refresh cycle (~5 years) of the JANET network (i.e. well before XLZD would start data taking) WP5 would generate requirements and work with JISC.

Boulby (Mine):

- Upgrade of the networking into the mineshaft has just be completed:
  - 2x 24 pair fibres installed from the surface facility to the underground laboratory.
  - Currently 160 Gbps connection from the surface to the underground laboratory.
- Underground Networking/Computing:
  - Increased rack capacity (for e.g. storage) with 10 Gpbs connection to new experiment.
- WP5 would generate requirements and work with Boulby.

Questions ?

I have lots …..

Please ask !

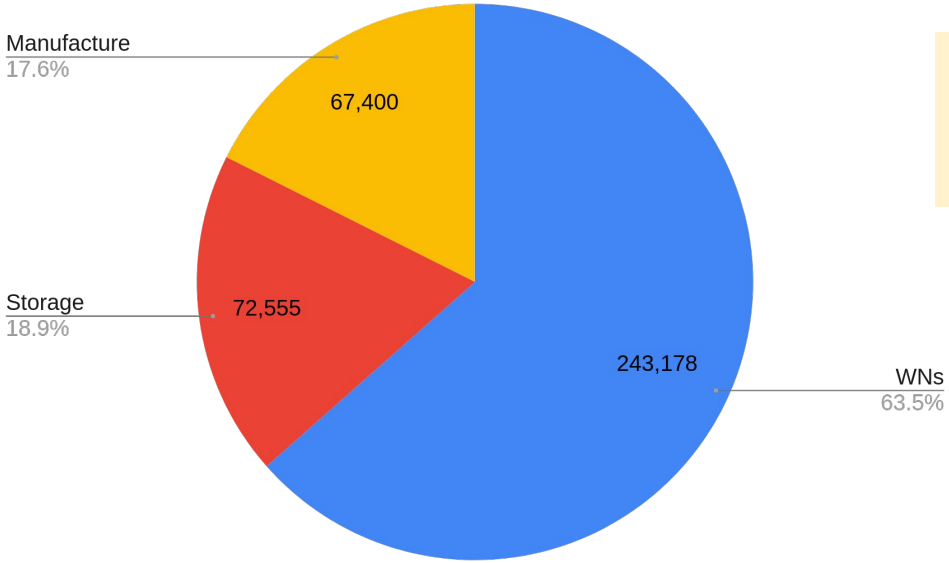
daniela.bauer@imperial.ac.uk

d.costanzo@sheffield.ac.uk

But in summary, we are well placed to get started and XLZD@Boulby presents a great opportunity to design the computing facilities XLZD wants and needs !

Supplemental information

Sustainability: Imperial College Tier 2 CO_2 equivalent per year



Imperial Tier 2 Site:

- 22 PB storage
- ~16000 job slots

Source:
CHEP 2023: Environmental Impact of a large Tier 2

Total: 383,000 kg CO_2e/year

Imperial College London

The University Of Sheffield.

IRIS & GridPP

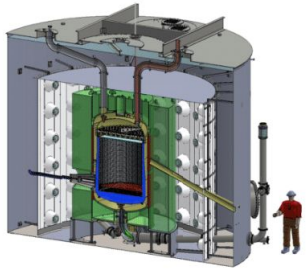
GridPP:	IRIS:
~110k cores	~50k cores
~110 PB disk	~22 PB disk
~125 PB tape	~3 PB tape

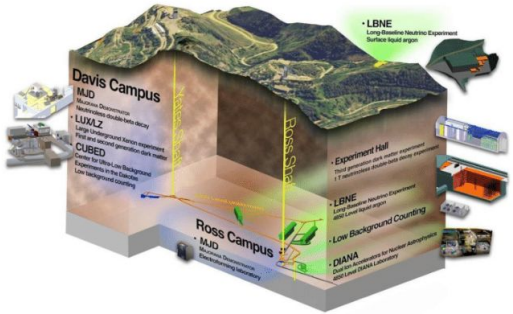
Imperial College London

The University Of Sheffield.

The LZ Dark Matter Experiment (experiment adjacent computing facilities)

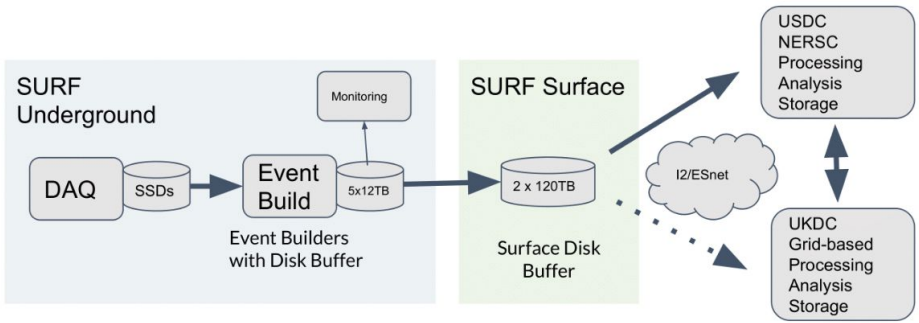
LZ detector





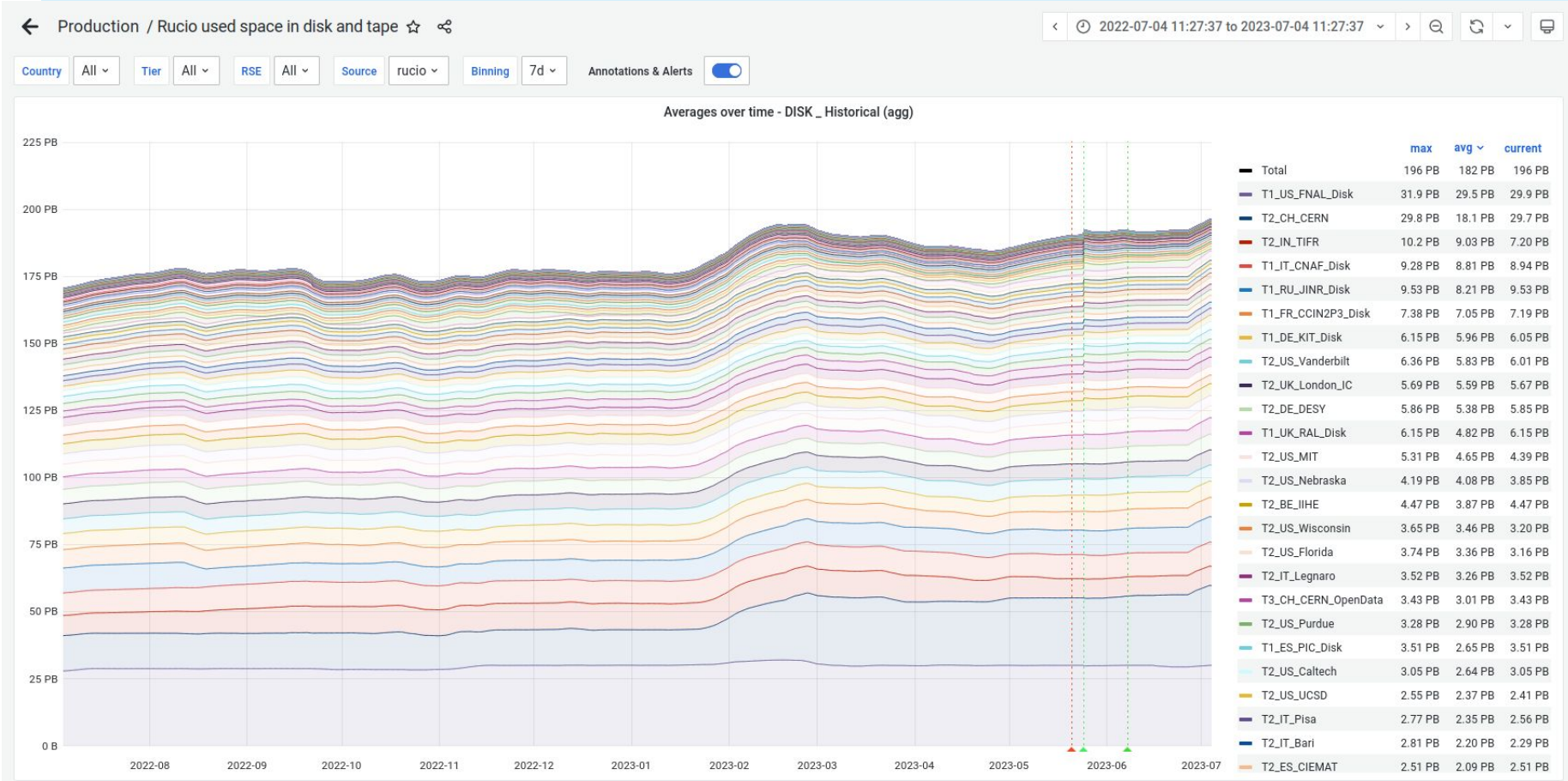
NIM A 163047 (2020)

- Liquid Xenon Time Projection Chamber (TPC)
  - 7 tonnes active LXe
  - Located 4850ft (~1mile) underground at Sanford Underground Research Facility (SURF) in Lead, South Dakota, USA
  - Currently operating and taking data
- Data rates
  - up to ~ 40MByte/s during routine Dark Matter search
  - up to ~ 350 MByte/s during calibration
- Dataflow deeply buffered
  - hold 2 months of Dark Matter search data at SURF
- ~200 active user accounts in system (shifters, experts, etc.)

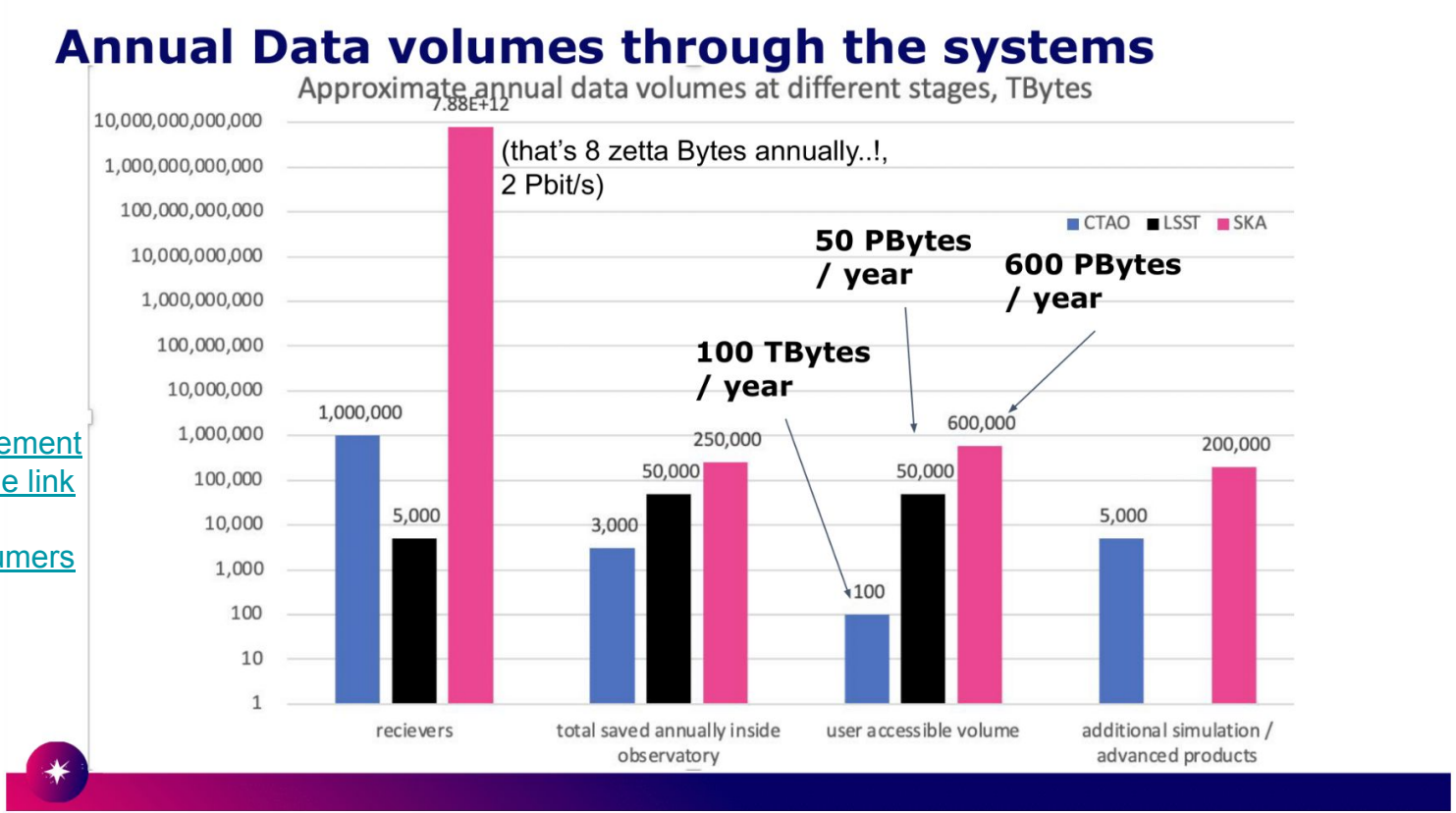


S. Luitz, CHEP 2023, https://indico.jlab.org/event/459/contributions/11644/

For scale: CMS data (last 12 months)



For scale: Astronomy



Source:
Rosie Bolton
[Global Data Management in Astronomy and the link to HEP Are we collaborators, consumers or competitors?]