

ATLAS Updates

Jyoti Prakash Biswal

(On behalf of the UK Cloud squad)

Rutherford Appleton Laboratory

GridPP51 & SWIFT-HEP07

Conferencing – The Edge, Sheffield

26 March 2024



Rutherford Appleton
Laboratory



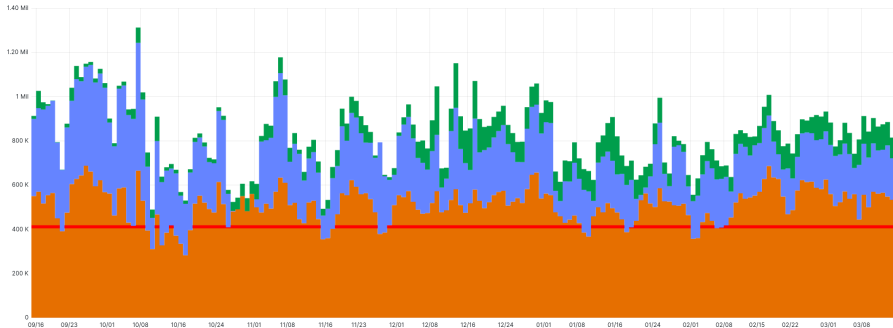
GridPP
UK Computing for Particle Physics



UK-specific

[Analysis of performance of the UK sites from various angles.]

UK Compute - I



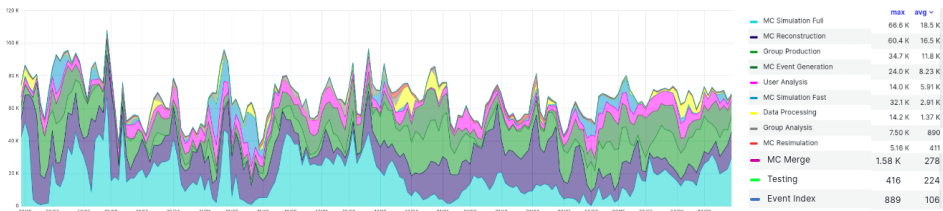
Slots of Running jobs (HS23) – last 6 months, until 15 March 2024 – [link](#).

- UK delivered an average of ~ 825 K HS23 (against the combined pledge of 411 K HS23) over the last six months.

| Type | Average HS23 |
|--------|--------------|
| RAL | 246 K |
| Tier2s | 511 K |
| Tier3s | 69.4 K |

Each colour is in line with the ones in the plot above.

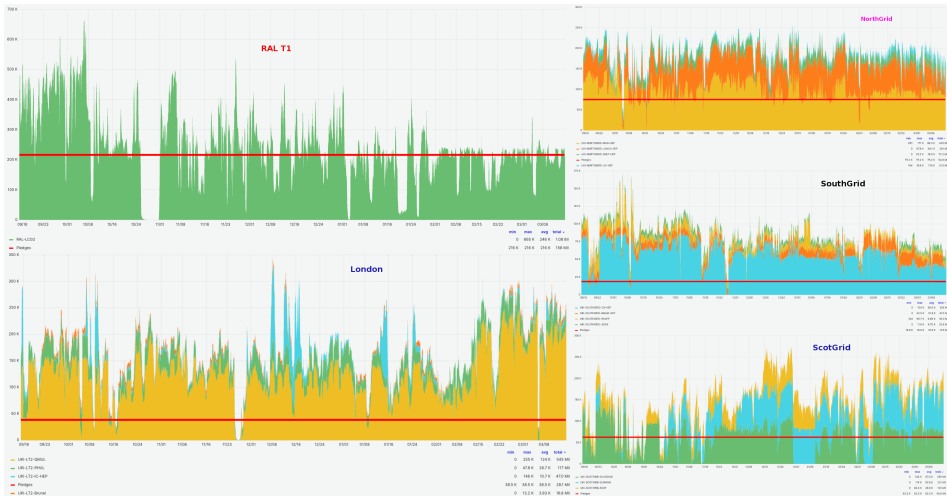
- Peaked at ~ 1.30 M HS23.



Slots of Running jobs by ATLAS production class – last 6 months, until 15 March 2024 – [link](#).

- The top five are *MC Simulation Full*, *MC Reconstruction*, *Group Production*, *MC Event Generation*, and *User Analysis*.
- **Largely, CPU time has been spent on MC-related tasks!**

UK Compute - III/a



Job Accounting (UK Cloud): Slots of Running jobs (HS23) federation-wise – last 6 months, until 15 March 2024 – [link](#).

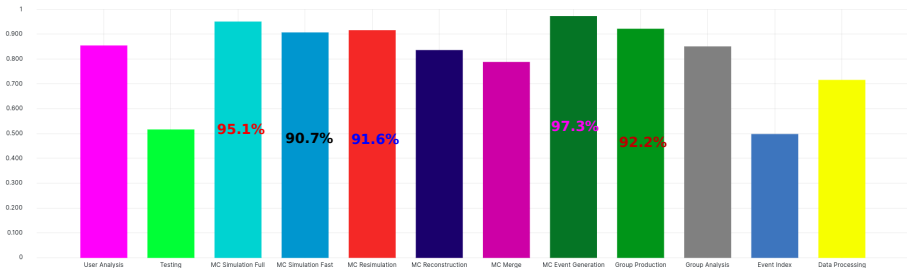
Note: Due to the scheduled downtime for electrical power testing at RAL and its aftermath, there are slumps during October/November 2023 – [GOCDB:34592](#).

| Site [Federation] | Average HS23 | Sum HS23 [Federation Pledge] |
|-------------------------|--------------|------------------------------|
| RAL [Tier1] | 246.00 K | 246.00 K [240 K] |
| Manchester [NorthGrid] | 98.30 K | 186.09 K [80 K] |
| Lancaster [NorthGrid] | 64.10 K | |
| Sheffield [NorthGrid] | 16.50 K | |
| Liverpool [NorthGrid] | 7.19 K | |
| Oxford [SouthGrid] | 58.00 K | 82.39 K [25 K] |
| Birmigham [SouthGrid] | 10.80 K | |
| RALPP [SouthGrid] | 8.89 K | |
| Sussex [SouthGrid] | 4.70 K | |
| Glasgow [ScotGrid] | 67.40 K | 146.60 K [63 K] |
| Durham [ScotGrid] | 50.60 K | |
| Edinburgh [ScotGrid] | 28.60 K | |
| Queen Mary [London] | 124.00 K | 165.23 K [35 K] |
| Royal Holloway [London] | 26.70 K | |
| Imperial [London] | 10.70 K | |
| Brunel [London] | 3.83 K | |

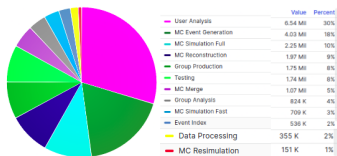
Job Accounting (UK Cloud): Slots of Running jobs (HS23) federation-wise – last 6 months, until 15 March 2024 – [link](#).

The individual sites from the corresponding federations are mostly above their pledge.

UK Compute - IV



Average CPU Efficiency Good jobs – last 6 months, until 15 March 2024 – [link](#).

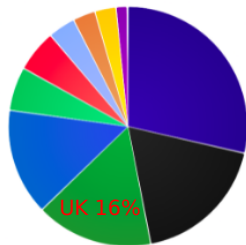


Pie Chart: WallClock Consumption of Successful and Failed Jobs – last 6 months, until 15 March 2024 – [link](#).

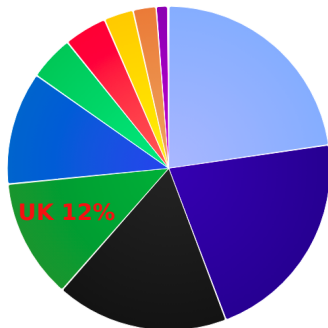
Completed jobs – last 6 months, until 15 March 2024 – [link](#).

The CPU efficiency is satisfactory across the range of activity types (top plot). User analysis shows $\sim \frac{1}{3}$ of the completed jobs (bottom-left plot). Moreover, $\sim 6\%$ of walltime is lost to failed jobs – problems with staging data to/from the worknodes, or direct access failures. There are also some cancelled jobs (bottom-right plot).

UK/Other Clouds - I



| | Value | Percent |
|----|----------|---------|
| US | 271 Mil | 29% |
| DE | 174 Mil | 18% |
| UK | 150 Mil | 16% |
| FR | 136 Mil | 14% |
| IT | 55.4 Mil | 6% |
| CA | 53.4 Mil | 6% |
| ND | 34.5 Mil | 4% |
| NL | 28.2 Mil | 3% |
| ES | 26.6 Mil | 3% |
| RU | 13.3 Mil | 1% |
| TW | 449 K | 0% |

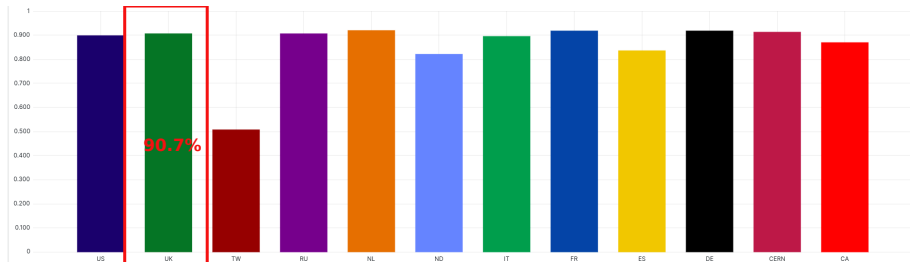


Slots of Running jobs (HS23): T1+T2+T3 (Grid Sites) – last 6 months, until 15 March 2024 – [link](#).

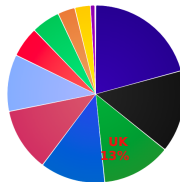
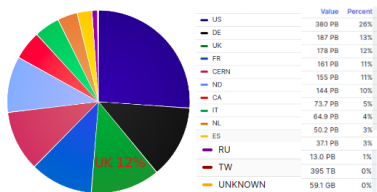
Slots of Running jobs (HS23): T1+T2+T3 (Grid Sites) + High-Performance Computing (HPC) – last 6 months, until 15 March 2024 – [link](#).

- The UK Cloud contributes $\sim 16\%$ of total ATLAS compute across the T1, T2, and T3 Grid sites (left plot). **UK excels in its performance – the assigned value is $\sim 15\%$.**
- Its value drops to $\sim 12\%$ once HPC (largely from opportunistic resources) is included (right plot).

UK/Other Clouds - II



Average CPU Efficiency Good jobs – last 6 months, until 15 March 2024 – [link](#). UK is fifth, could do better!

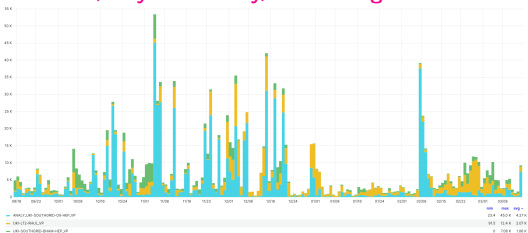


NBytes Processed (input) in %age – last 6 months, until 15 March 2024 – [link](#). **NBytes Produced (output) in %age** – last 6 months, until 15 March 2024 – [link](#).

Within the past six months or so, the UK processed ~ 175 PB data. [Over-estimated as counts only total file sizes, and not, e.g., direct-IO in User Analysis]

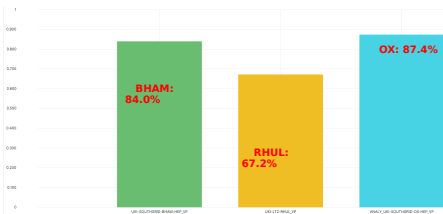
Virtual Placement/XCache - I

- Caching and data placement never fails to cause curiosity!
- **Virtual Placement (VP):**
 - Uses rucio to assign “virtual replicas” at sites ⇒ aims to improve cache hit rates.
 - Only transfers the bytes that are needed for the workload.
 - Exclusively configured for **Analysis jobs**.
- **XCache:**
 - Provides caching of data accessed using XRootD protocol.
 - Sits in between client and upstream XRootD servers and can cache/prefetch full files or only blocks already requested.
 - ATLAS monitoring available [@elasticsearch](#).
- **At the moment, – Oxford, Royal Holloway, and Birmingham have configured caches for VP.**



Slots of Running jobs (HS23) – last 6 months, until 15 March 2024 – [link](#).

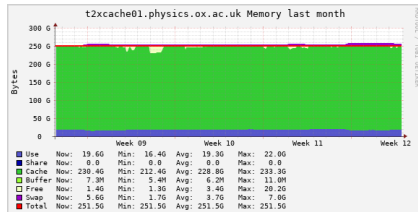
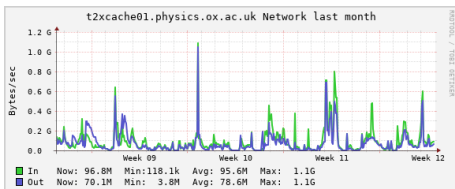
Virtual Placement/XCache - II



| | Value | Percent |
|----------------|-------|---------|
| User Analysis | 553 K | 60% |
| Group Analysis | 219 K | 24% |
| Testing | 147 K | 16% |

Average CPU Efficiency Good jobs – last 6 months, until 15 March 2024 – [link](#). The corresponding non-VP numbers are – 83.3% (BHAM), 92.5% (RHUL) and 92.3% (OX).

Completed jobs – last 6 months, until 15 March 2024 – [link](#).



Oxford-only – XCache **network** last month.

Oxford-only – XCache **memory** last month.

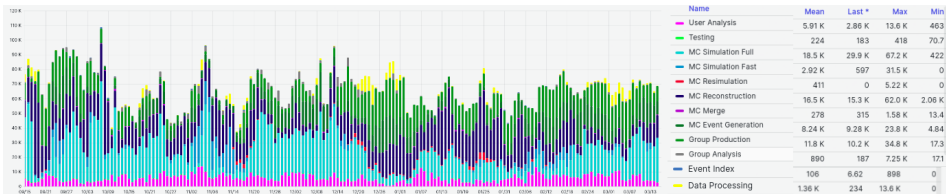
Discussions:

- Both VP and XCache have been performing well.
- Edinburgh was interested (some time ago) in VP. Is there any further news?
- There are notification emails when VPs are excluded from HammerCloud, but no emails after they come back online!
- Both VP (analysis) and production queues use the XCache.
- To do:
 - Demonstrate quantitatively the benefits of VP/XCache.
 - Test/move the production queues to bypass the cache.

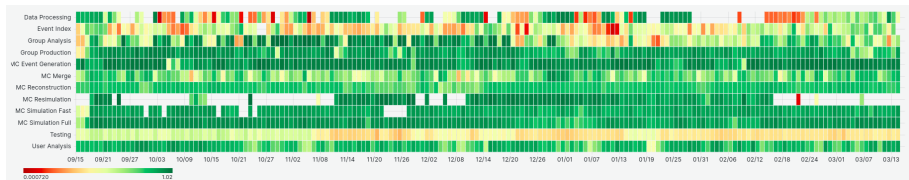
More trials and scrutiny are required before making the recommendation. These sectors are expected to see significant growth.

UK Sites – Job Mix

- The influence of job mix is frequently underestimated and easily overlooked when analysing the job efficiencies and success rates at UK sites.
- It is **crucial** to comprehend the scope of this.



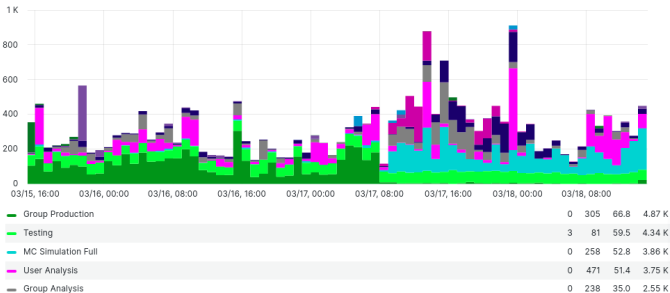
Slots of Running jobs – last 6 months, until 15 March 2024; [link](#).



CPU Efficiency Good jobs – last 6 months, until 15 March 2024; [link](#). The darker the green colour, the better the CPU Efficiency.

UK Sites – Job Mix (contd)

- A site-specific example is Lancaster.
 - Its storage was a lot happier around the time the job mix changed! (Behind the scenes qualitative observation.)
 - *i.e.*, when Group Production is replaced with MC Simulation Full, things are better!
 - Was it a coincidence?
 - Was the old job mix causing stress on the storage?



Completed jobs – over a few days in March 2024; [link](#).

Action: The Tier1 liaison(s) will try to keep a close eye on job mix in future.

General Topics

[Gathered from the most recent [ADC @ATLAS Software & Computing Week.](#)]

The reason(s)

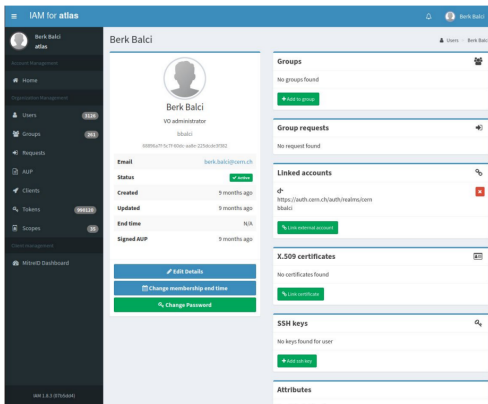
- The VOMS-Admin is currently running on CERN CentOS7 (CC7).
- Not supported on newer OS versions!
- The end of life (EOL) for CC7 is 30 June 2024.
- The successor, INDIGO IAM, has already been deployed and is in use.
 - A token-based Identity and Access Manager (IAM) with a component for VOMS backward compatibility, **VOMS-AA** (Attribute Authority).

VOMS-AA

- **Setup IAM:**
 - Import existing users from VOMS-Admin.
 - Add IAM-VOMS endpoint to Live Screensaver Creator (LSC) files on all sites supporting the VO.
 - Add IAM-VOMS endpoint to VOMS client configuration of the VO.
 - **IAM setup is complete for all four LHC VOs.**
- **Retire VOMS:**
 - Remove VOMS endpoint from VOMS client configuration of the VO.
 - Remove VOMS endpoint from LSC files.
 - **This part is pending/ongoing for the LHC VOs.**

VO Administration:

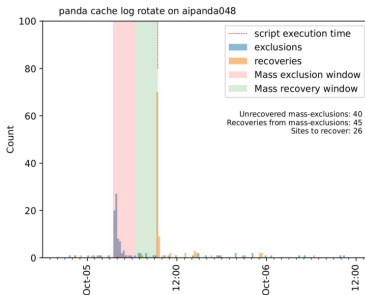
- INDIGO IAM web interface is used for Administration.
- Instinctive and simple to use.
- Training will be made available for the VO administrators.

**Timeline:**

- ~~DC24: 12-23 February.~~
- IAM and VOMS-AA update to 1.8.4 (after DC24).
- Announcement of the new deployment on Kubernetes and publishing .isc rpms (March).
- Training for VO Administration via IAM web interface.
- The timing of the switch from VOMS-Admin to IAM to be decided per VO → later could be better...

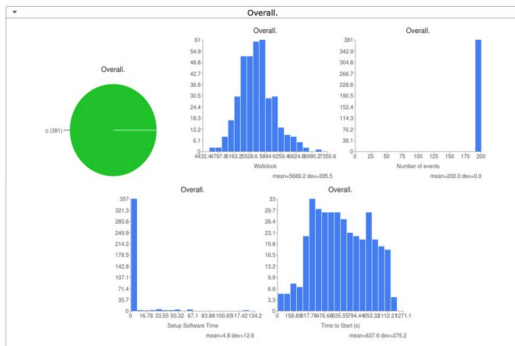
Automatic Massive Recovery:

- A new way to recover sites automatically after a massive auto-exclusion.
- **When queues start to recover from a mass exclusion event, recover all remaining queues automatically, without waiting on test results.**
- If, within 3 hours, there are 18 or more queues excluded which are not recovered yet \Rightarrow the alarm is sent.
- Additionally, if $> 40\%$ of the excluded queues have recovered \rightarrow all queues excluded during mass-exclusion are recovered.
- The exact rules are on the [Twiki](#).



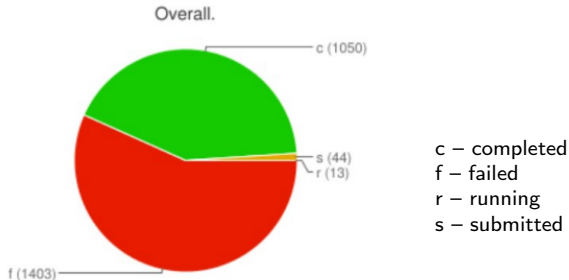
Stress Tests:

- Manual start/end of tests \Rightarrow Capable of many jobs in parallel.
- Examples – analysis, simulation, HEPSCORE.
- Results are available as plots:
 - [Test view](#).
 - [Table Site Overview](#).
 - Jobs could be downloaded from BigPanDA, e.g., [tid1258](#).



IPv6 Test:

- Last year, [the WLCG management board](#) proposed an IPv6 deployment campaign on compute services.
- HammerCloud helps monitor progress with functional tests – [HC-1385](#).
- One job per hour; **not used for auto-exclusion**.
- Around half of the jobs are failing at the moment!
 - Production Functional Test (PFT): [1256 template](#); [Recent Test](#); [Recent PanDA jobs](#).
 - Analysis Functional Test (AFT): [1253 template](#); [Recent Test](#); [Recent PanDA jobs](#).



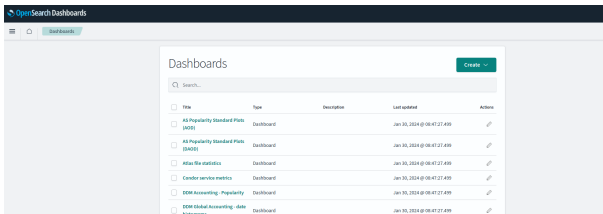
- CPU architecture: x86_64, GPU, aarch64, NVIDIA, or none.
- CPU type: CPU or GPU.
- Accounting Portal: Sum CPU Time, Sum CPU Work HS23, Sum Wallclock Time, ...
- ETF/SAM+HC:
 - ETF: Experiments Test Framework // SAM: Service Availability Monitoring // HC: HammerCloud.
 - Panels in HC are *static*, whereas in ETF+SAM are changeable.
- Suspicious Sites:
 - Based on ATLAS_PANDA.JOBS_STATUSLOG table, which contains all information about job statuses transitions (failed, closed, cancelled, ...), that allows very accurate measurements of the number of jobs in all statuses.
 - Metrics are calculated *hourly* – number of jobs in all statuses; average duration of jobs in each status; average cumulative duration of jobs in each status, ...

~ Suspicious Sites

Running Jobs Diff

| site | Average: h | Average: 24h | Average: week | Average: month | Diff [h/24h] | Diff [24h/week] | Diff [24h/month] + | Diff [week/month] |
|------------------------|------------|--------------|---------------|----------------|--------------|-----------------|--------------------|-------------------|
| CESN-T0 | 11.72 K | 11.28 K | 16.89 K | 19.93 K | 240.79 | -6.51 K | -6.55 K | -3.04 K |
| GoGrid | 1.82 K | 2.65 K | 4.17 K | 9.55 K | -823.29 | -1.33 K | -5.90 K | -4.37 K |
| MWT2 | 19.83 K | 20.89 K | 22.41 K | 26.31 K | -1.00 K | -1.51 K | -5.42 K | -3.91 K |
| UKI-NORTHGRID-MAN-H... | 2.01 K | 1.80 K | 4.29 K | 5.89 K | 209.83 | -2.48 K | -4.09 K | -1.61 K |
| CSCS-LC02 | 599.00 | 730.88 | 2.30 K | 3.70 K | -131.87 | -1.57 K | -2.97 K | -1.39 K |
| SWTZ_DOOGLE | 4.00 | 129.88 | 3.27 K | 2.90 K | -125.87 | -3.14 K | -2.83 K | 312.24 |

- [Documentation](#).
- OpenDistro to OpenSearch migration – [link to slides](#).
- Cluster: <https://os-atlas.cern.ch> [OpenSearch 2.11.1]
 - Archived data.
 - Plots, dashboards.
 - Security settings (users, roles).
- A few interesting items from the migration list:
 - PanDA, JEDI, BigPanDA, ProdSys, IDD, Pilot logs processing, and dashboards in ATLAS Kibana.
 - Harvester monitoring in MONIT Grafana and ATLAS Kibana + Harvester Service Monitoring.
- Other relevant links: [PanDA logger page](#) in BigPanDA; [Tasks status log/Kafka](#).



- UK performed very well w.r.t. other Clouds in the past six months.
- UK holds significant expertise through its personpower.
- Other specialised talks will cover the (un)covered areas (e.g., DC24, tokens, network, ...) in details.
- Hopefully, this presentation has touched upon some common/essential themes.
 - Sites could take note of the items they need.
 - Suggestions for improvement are welcome!

New appointment: *Brij Kishor Jashal* is set to start as the full-time ATLAS Tier1/Tier2 liaison in the early weeks of April 2024.

- Also connecting to this meeting remotely (before the starting period)!