

Analysis Facilities Monitoring Discussion

Rob Gardner, Liz Sexton-Kennedy

with contributions from Alessandra Forti, Markus Schulz, Thomas Hartmann, Yves Kemp, James Lett, Giordon Stark

WLCG/HSF Workshop @ DESY
16 May 2024

Plan for our discussion

- Offer questions, define a scope for today
- Show some examples of existing monitoring of AFs and grid-based analysis
- Open discussion, your ideas

Analysis Facilities Whitepaper

8 Monitoring and Metrics

To provide an extensive overview on how resources are used in order to guide infrastructure development and to allow users to make an informed decision about which infrastructure to use, key parameters should be published for each AF.

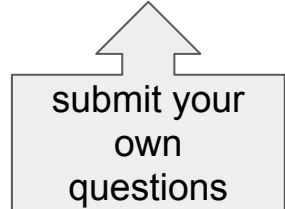
Metrics for AFs can be categorized into four areas

1. User experience metrics
2. User trend metrics
3. Performance metrics
4. Facility metrics

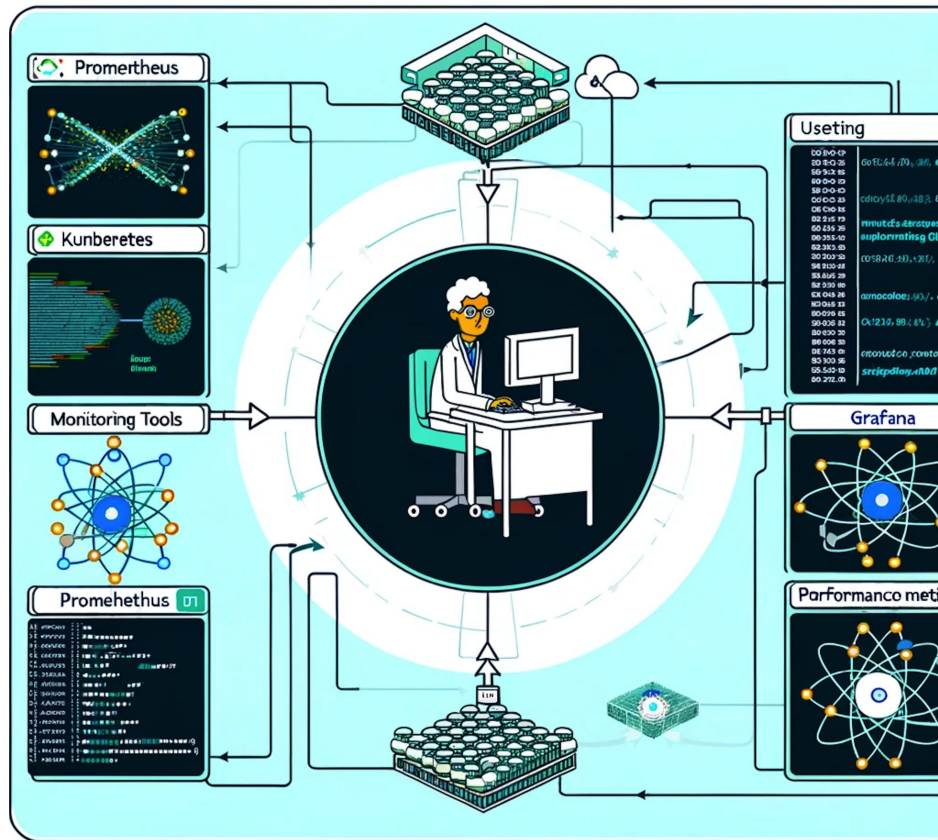
What are the questions, and from whom?

- Users (experience)¹
 - What resources are available?
 - How long will my jobs wait in queue? Why do they run so slow? Why is my notebook hanging? Why did my last few jobs not finish? Why are my jobs held? Why did my jobs fail? Why are they being held?
 - How do I access my data? Is it local? How do I get X software installed? How do I run with my container?
- Resource providers (trends, performance, facility metrics)²
 - What resources (cpu, disk-capacity, disk-fast, network, gpu) are under-provisioned?
 - What are the performance bottlenecks?
 - What are the (unexpressed) requirements?
 - Managing the storage - scratch, precious, freeing up space, group storage
 - Scheduling bursty workflows & precious resources (GPUs, fast storage)
- A fifth category: metrics for **framework & platform developers**³
 - Which data formats are physicists most often using and by which frameworks?
 - Are performance targets met? (e.g. X TB / Y minutes)
 - Where are the inefficiencies and user pain points?
 - What capabilities are missing?

[link](#)

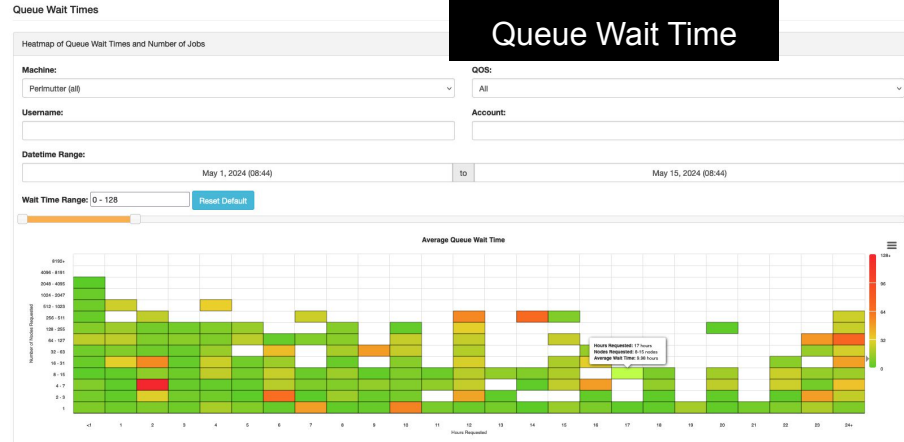


some
examples



NERSC (Perlmutter)

Queue Wait Time



User Queues

Dashboard

My Personal Disk Usage

Space Inodes

Unable to get your quota information for all machines. Likely causes are a machine outage or a SF API outage.

My Active Jobs

No Active Jobs

My Completed Jobs

Job ID	Host	Completion Time	Wait Hours	CPU Hours
25284813	Perlmutter CPU	05/08/24 22:30	1,280	1,28
25284812	Perlmutter CPU	05/08/24 17:28	1,681	1,68
25284811	Perlmutter CPU	05/08/24 17:27	1,652	1,65
23132563	Perlmutter CPU	03/21/24 01:15	2,235	2,23
23093777	Perlmutter CPU	03/18/24 00:51	24,007	24,01
23093776	Perlmutter CPU	03/18/24 13:19	12,242	12,24
23093775	Perlmutter CPU	03/18/24 03:28	3,114	3,11
23017598	Perlmutter CPU	03/17/24 01:20	12,006	12,01

System Status

Compute Systems:

- Perlmutter Down

Global Filesystems:

- DNA Up
- Data Transfer Nodes Up
- Global Home Up
- Global Common Up
- Community File System (CFS) Up

Mass Storage Systems:

- HPS8 Request (Backup) Up
- HPS8 Archive (User) Up

Service Status:

Jobscript Generator

Job Information

This tool generates a batch script template which also realizes specific process and thread binding configurations.

Machine

Select the machine on which you want to submit your job.

Perlmutter - GPU

Application Name

Specify your application including the full path.

mctm

Job Name

Specify a name for your job.

nnlo

Email Address

Specify your email address to get notified when the job enters a certain state.

Quality of Service

Select the QoS you request for your job.

regular

Wallclock Time

Specify the duration of the job. The max walltime for the regular QoS is 24 hours.

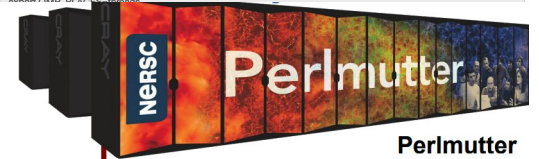
10 hours 30 minutes 0 seconds

Number of Nodes

Job Script Generator

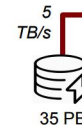
```
#!/bin/bash
#SBATCH -N 128
#SBATCH -C gpu
#SBATCH -G 512
#SBATCH -q regular
#SBATCH -j nnlo
#SBATCH -l 10:30:00
#SBATCH -A m1296

#OpenMP settings:
export OMP_NUM_THREADS=1
export MPILAB_PL_APSIS_threads=
```

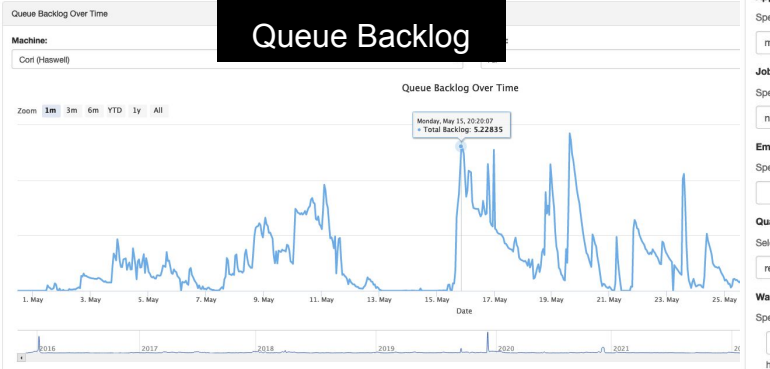


Perlmutter

1,500 NVIDIA A100x4 accelerated nodes
3,000 AMD dual-socket "Milan" CPU nodes
384 TB (CPU) + 240 TB (GPU) memory
HPE Cray Slingshot high speed interconnect
Debuted as World's 5th most powerful system
140 PF Peak



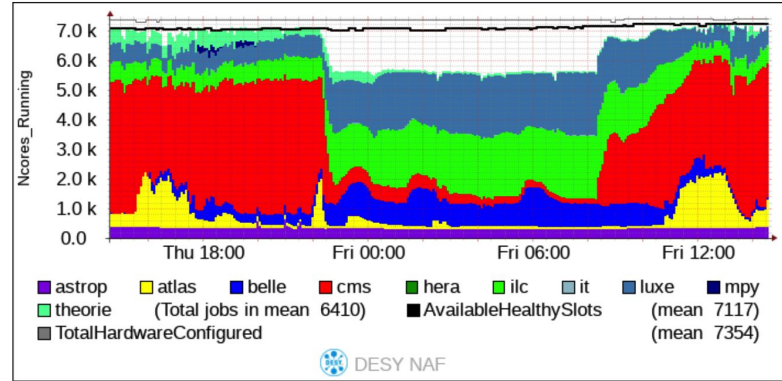
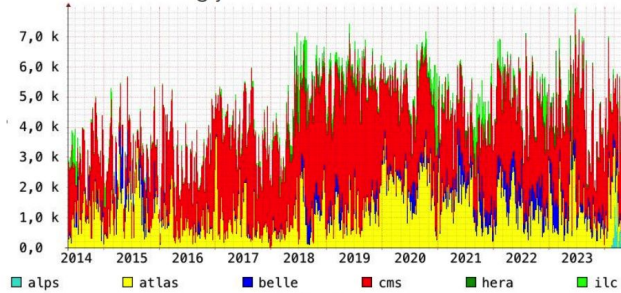
Queue Backlog



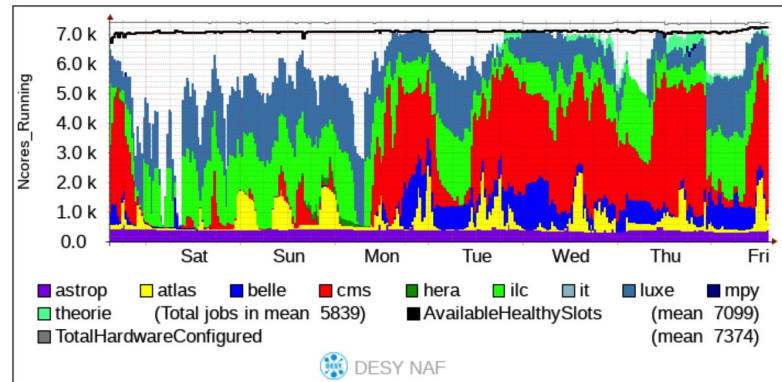
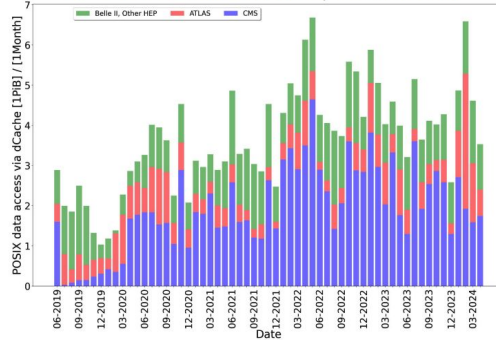
DESY NAF

aggregate resource delivery

Running job slots on the NAF since 2014

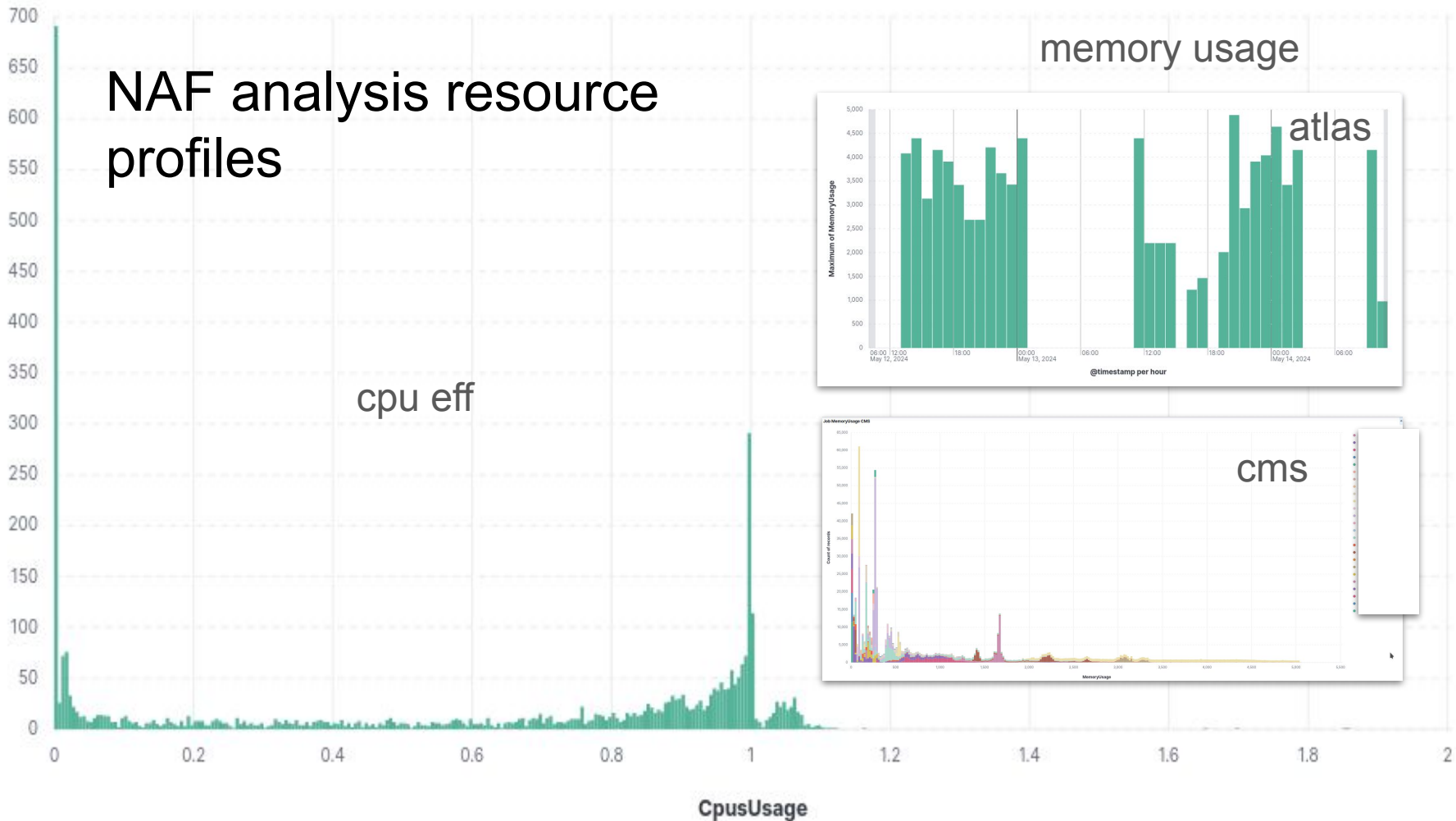


NFS traffic to/from NAF, PiB/month

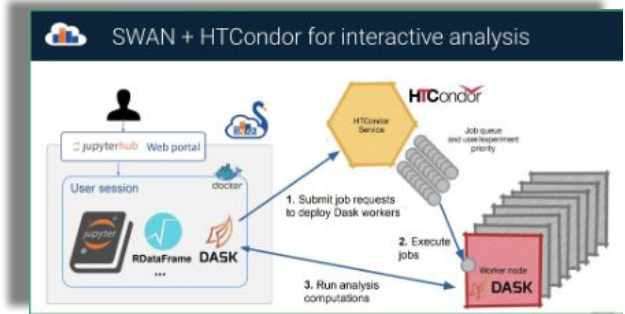


NAF analysis resource profiles

Count of records

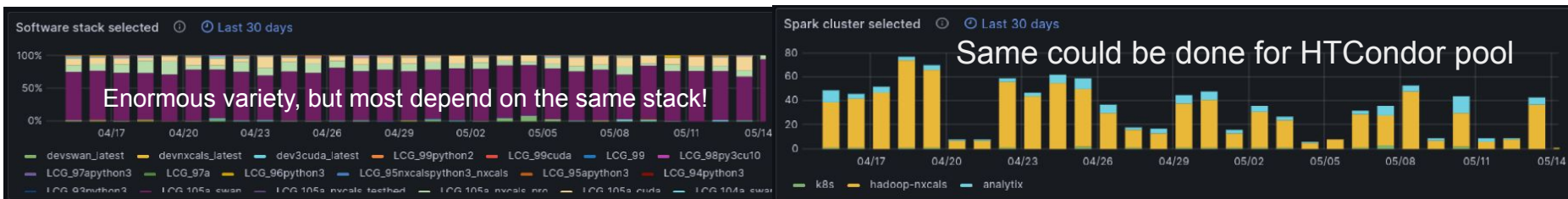
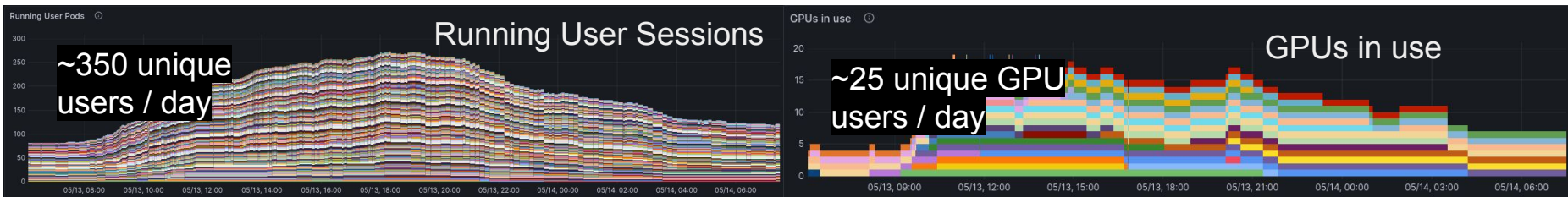


CERN AF-Pilot



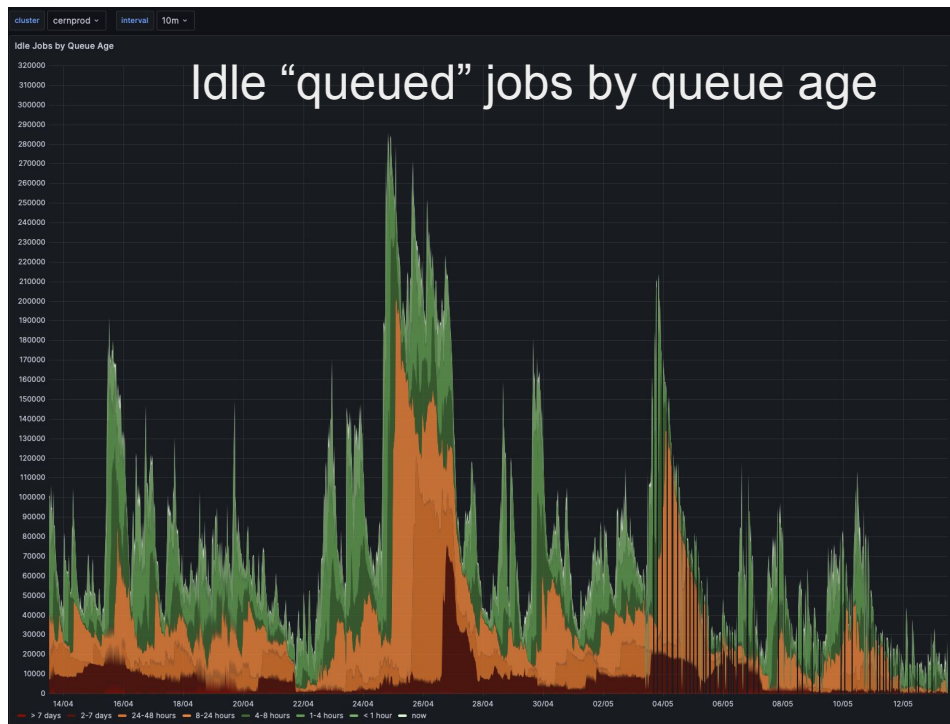
- The AF-pilot is linking and expanding existing services at CERN and connecting them to additional resource and load management systems
 - SWAN, CONDOR, Dask, RDataFrame, Coffea, CERN's shared GPUs, EOS...
 - There is quite detailed monitoring for all existing services and the infrastructure...
 - These have and will be combined to create a worthwhile AF specific monitoring
- The CERN AF-pilot is still in its infancy...
 - We have checked that in both systems we can track the use
 - But we haven't come to adding the specific attributes into our probes

CERN AF-Pilot Examples: SWAN

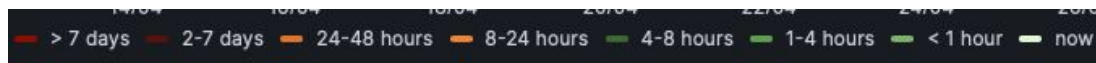


CERN AF-Pilot

Ideas for Condor:



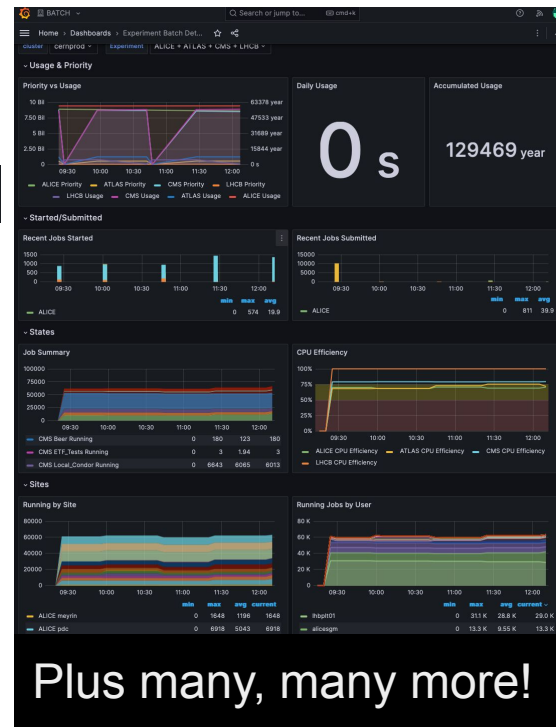
But with very different time ranges << 1h



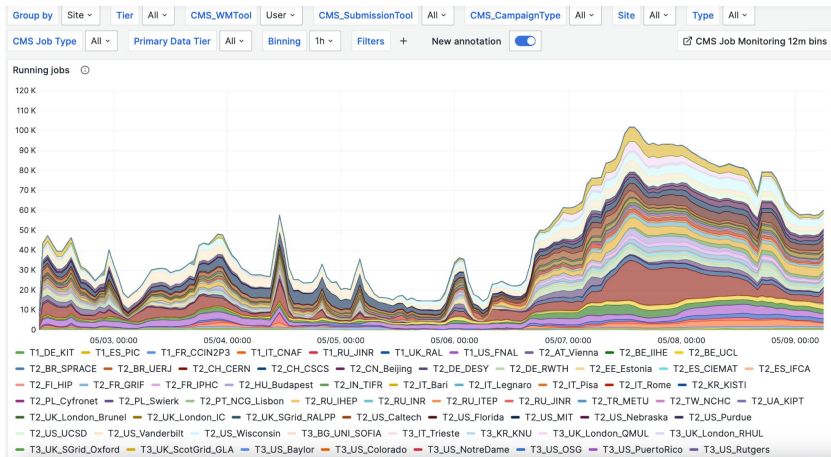
CERN AF-Pilot Ideas for Condor-2

Home > Dashboards > Experiment Batch Details

- We have detailed and complex monitors for the experiments' activities
 - 50+ diagrams...
- To this we consider to add similar plots covering the “interactive” usage patterns
- The accounting data at CERN can be queried to extract the AF specific resource usage
 - We haven't looked at this in detail
 - There hasn't been a lot of usage yet

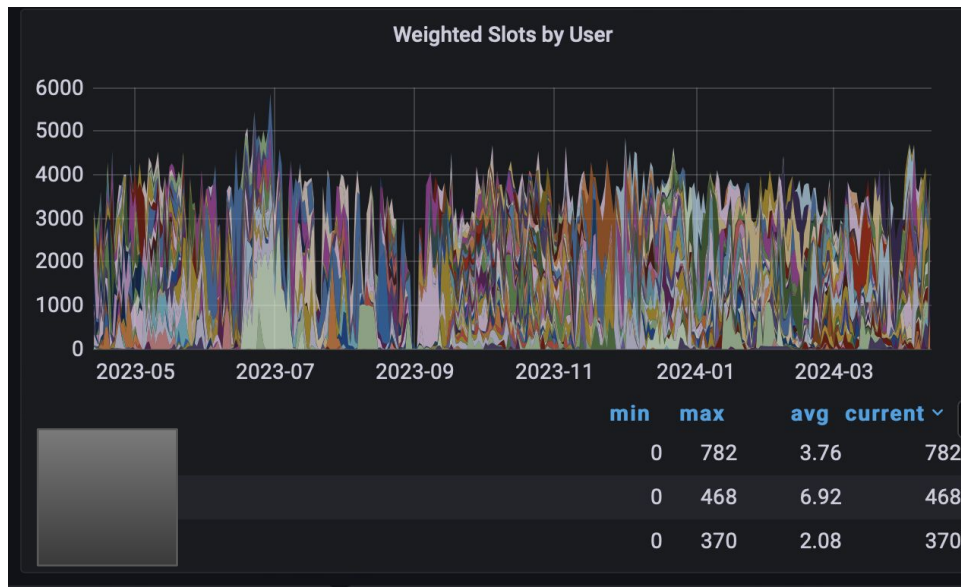


CMS - grid and local

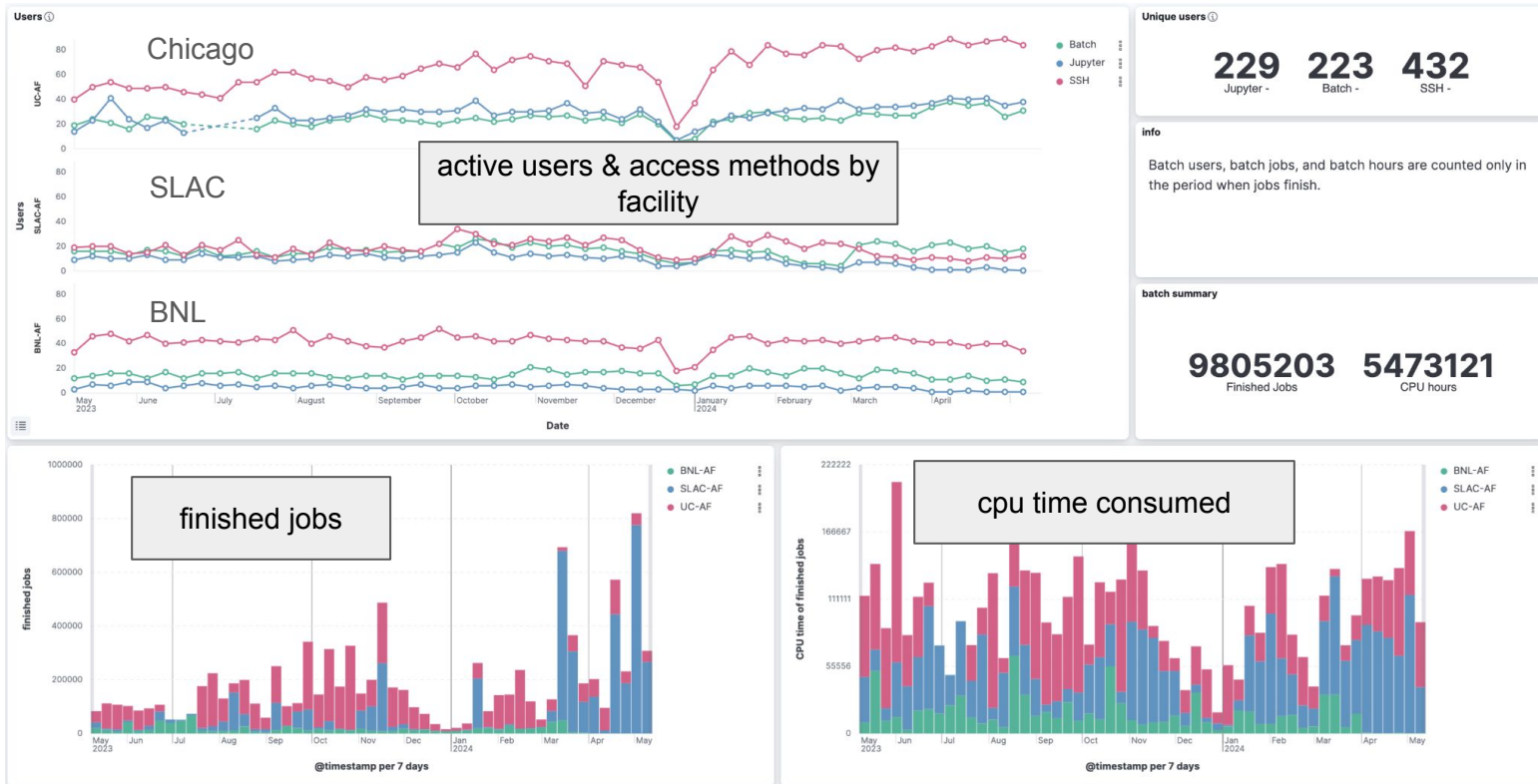


CMS Users in the Global Pool of WLCG sites

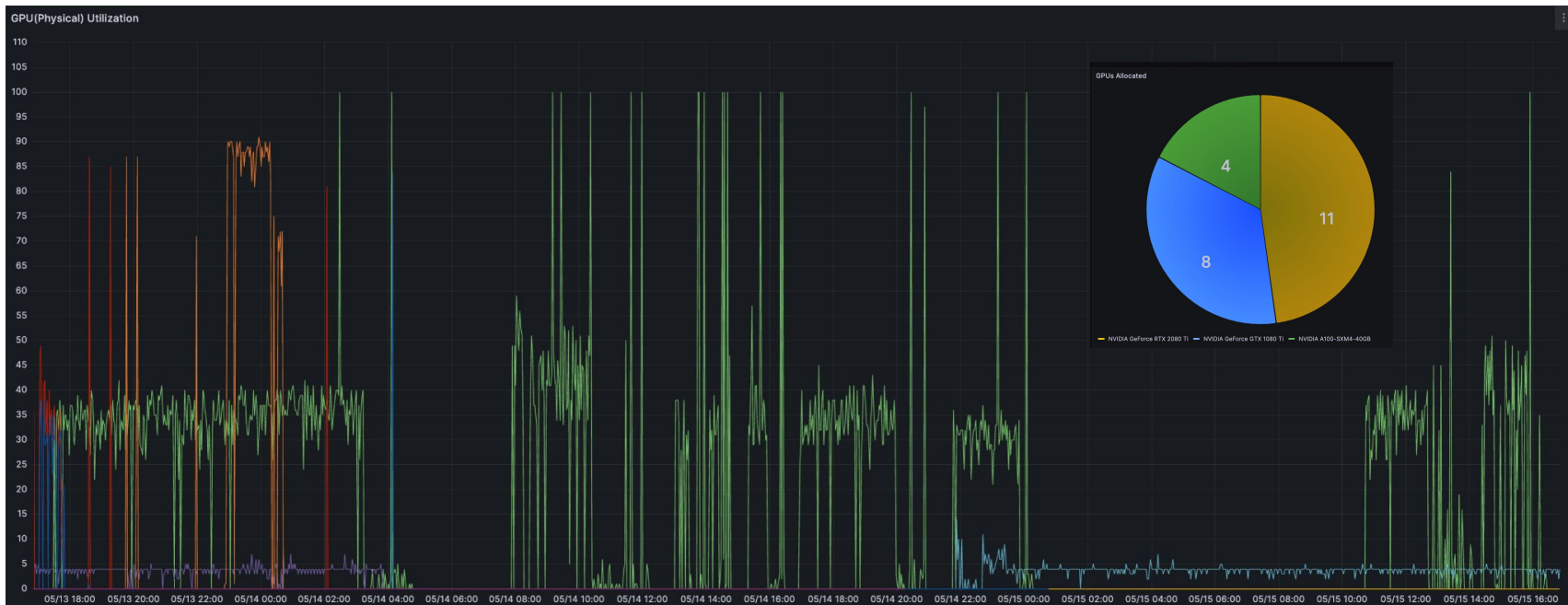
LPC users monitor over a year



ATLAS US Shared Tier3s ([link](#))

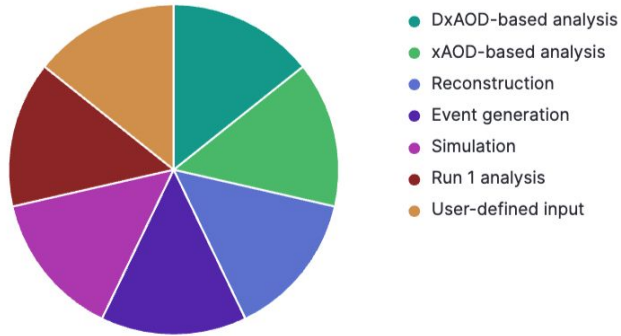


GPU utilization

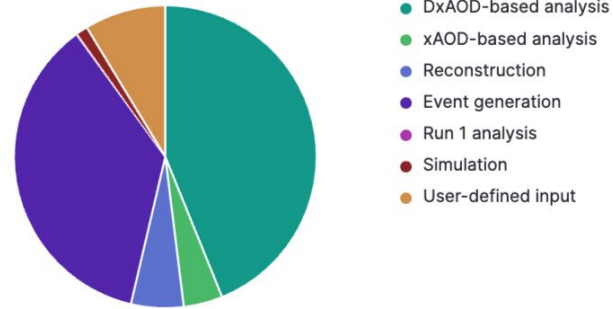


ATLAS grid analysis ([link](#))

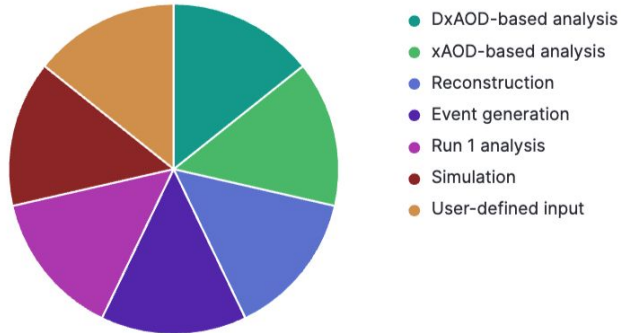
MK Analysis jobs per job type (users)



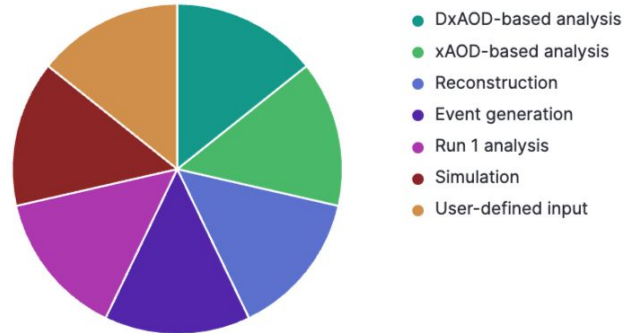
FL Analysis jobs per job type (walltime)



MK Analysis jobs per job type (counts)

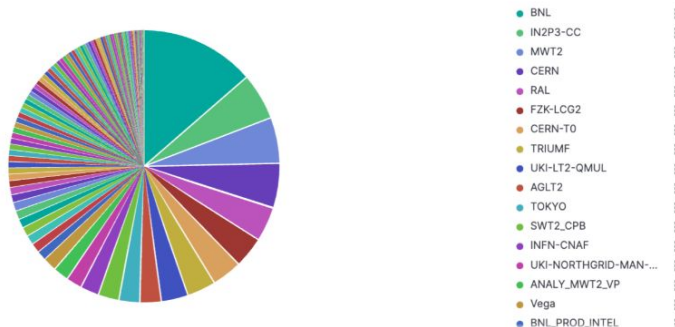


MK Analysis jobs per job type (events)

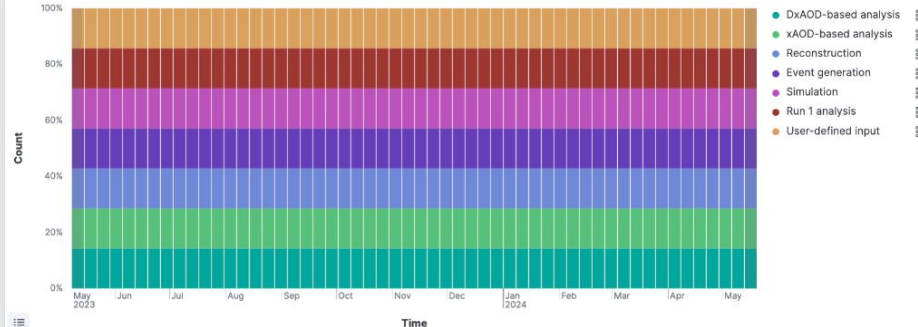


ATLAS grid analysis ([link](#))

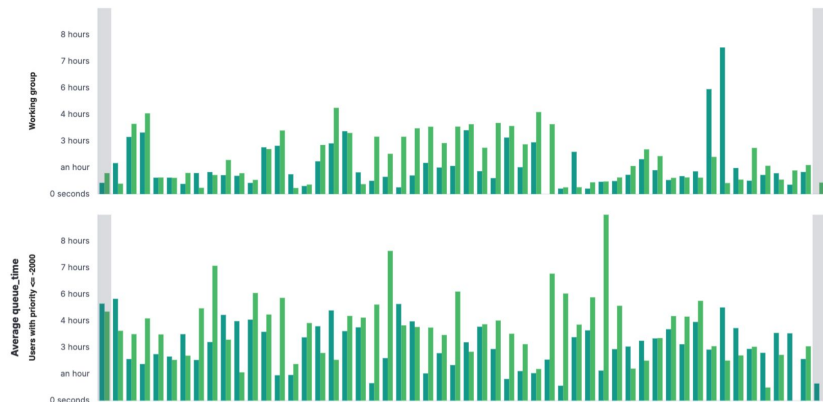
FL Analysis jobs per site



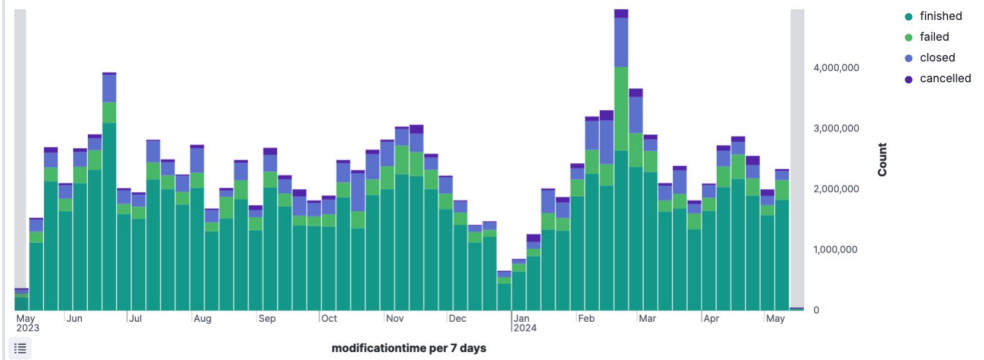
MK analysis vs input types



MK queue time vs time vs working group



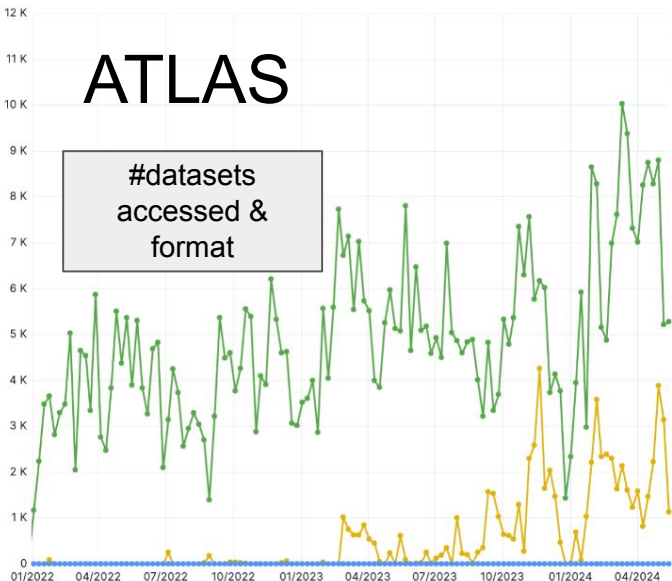
FL job status vs time



N datasets

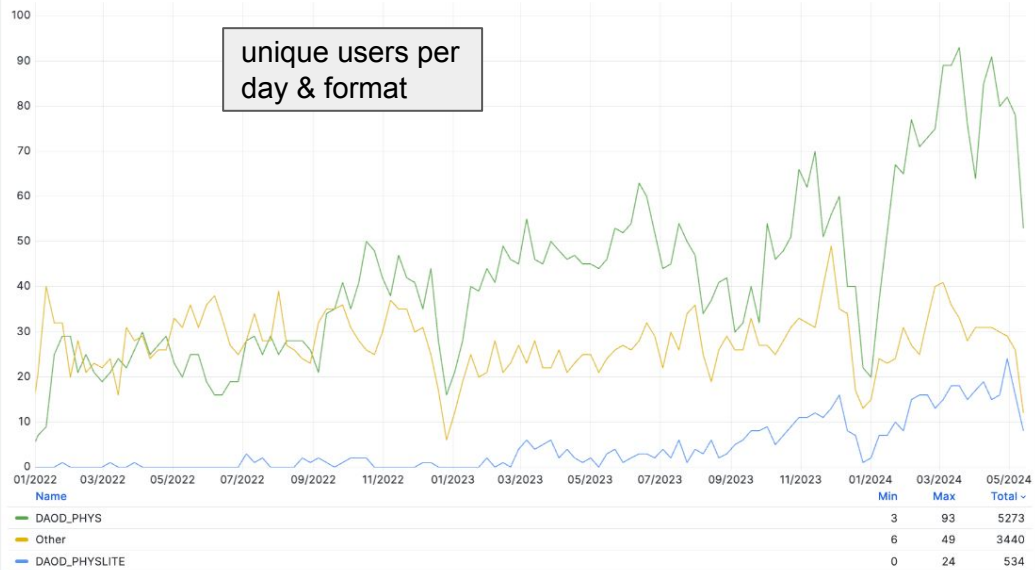
ATLAS

#datasets accessed & format



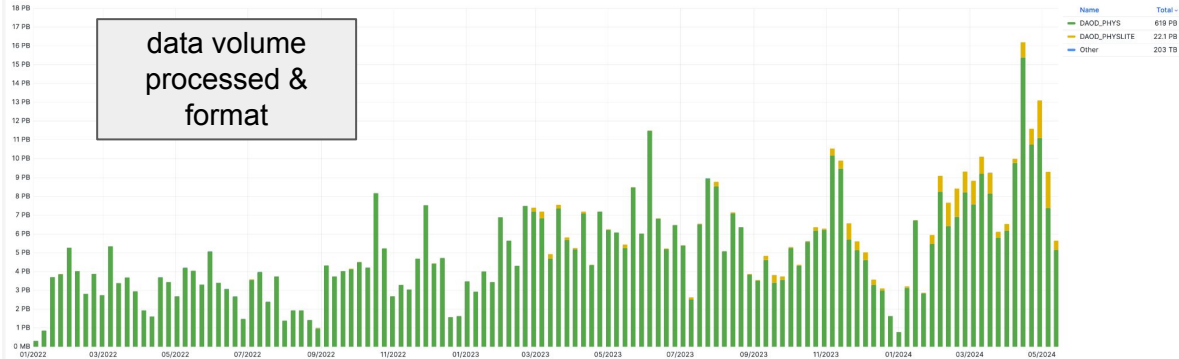
N Unique Users

unique users per day & format

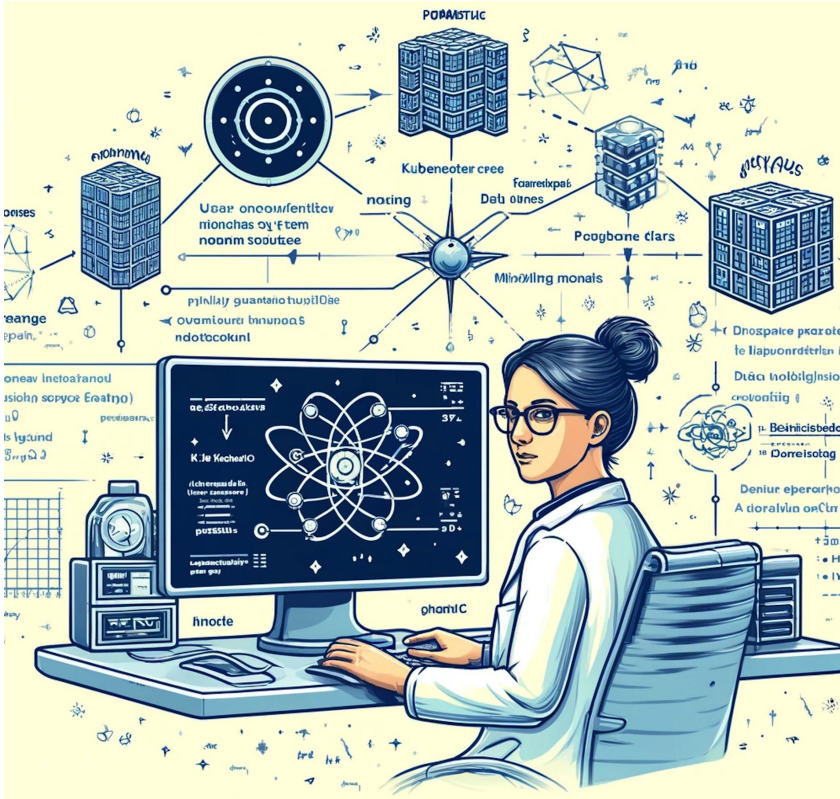


Data volume

data volume processed & format



discussion time!



[link](#)

