



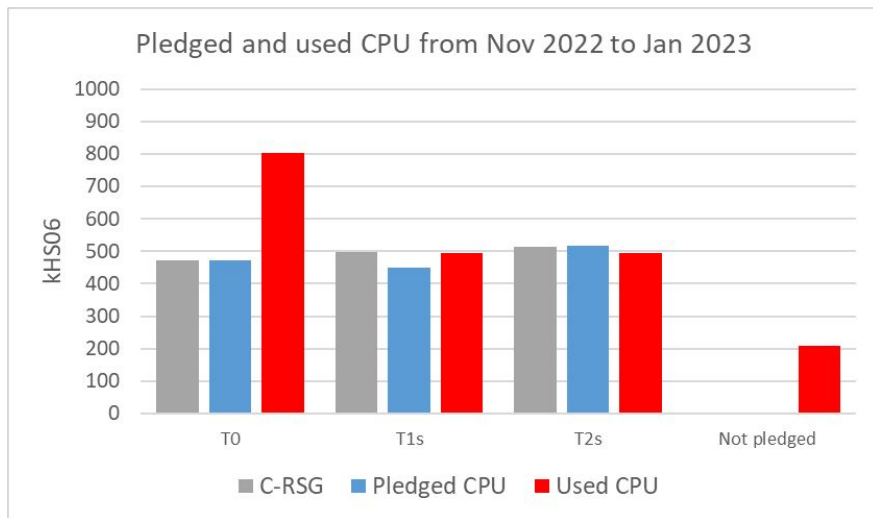
Computing Resources and processing plans for 2023

Latchezar Betev

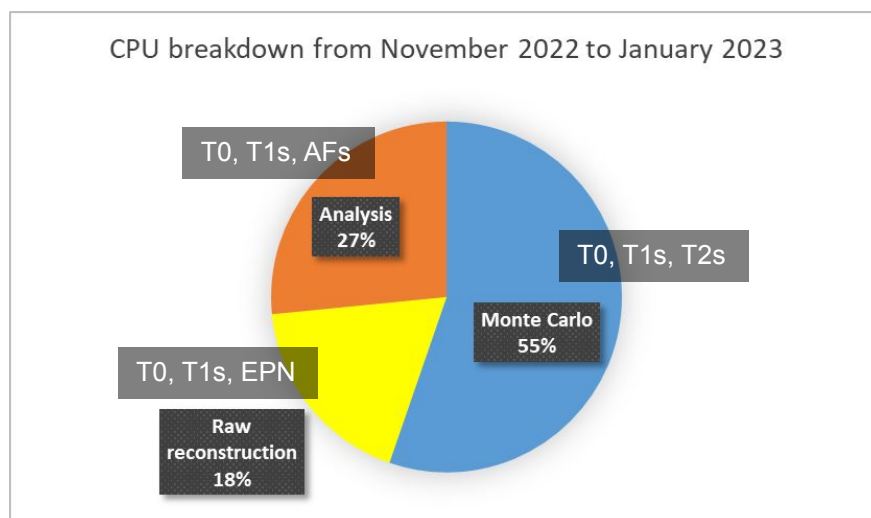


Resources usage

CPU utilization and breakdown by job types

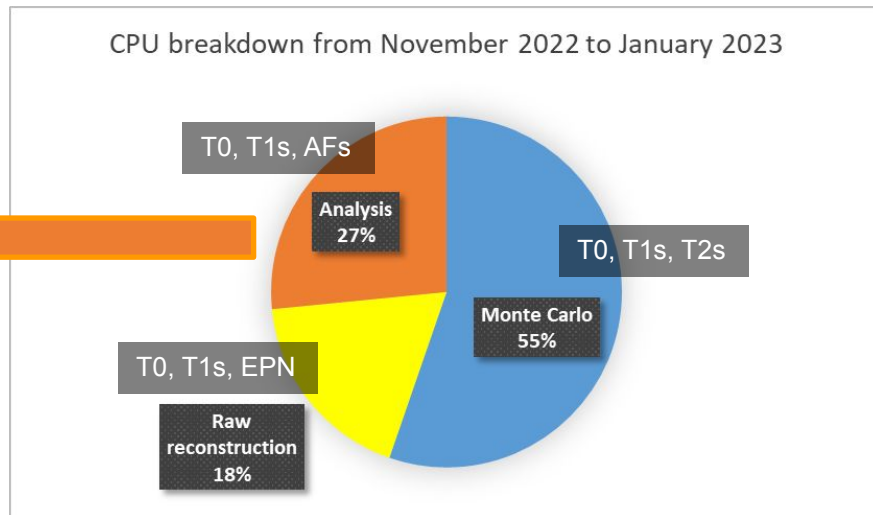
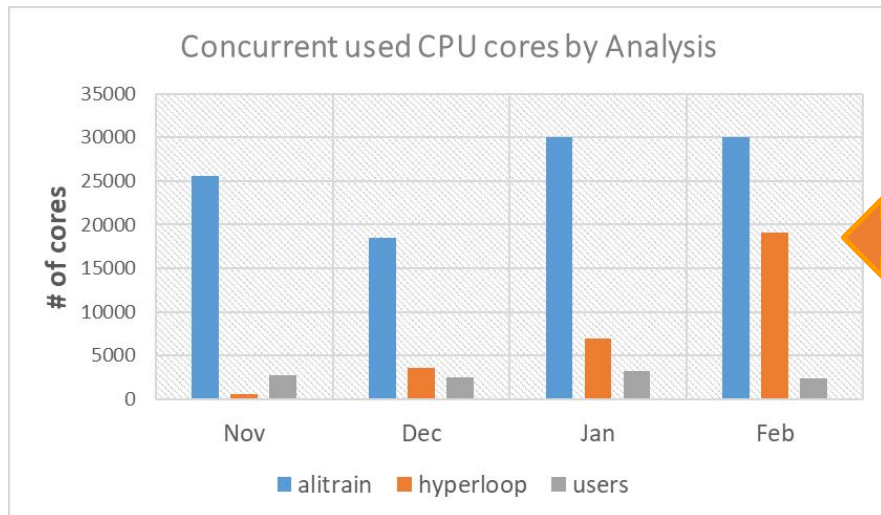


- Full utilization of the available resources with significant amount of opportunistic CPU usage at the T0 and LBL_HPC, Japan, Winger and EPN
- CPU capacity covers the pledges @ T0, T1s and T2s but pledge / C-RSG 90% @ T1 in 2022



- Higher activity for raw calibration and reconstruction of Run 3 pp runs than previously reported
- Steady Run 2 analysis activity for conferences and publications (~30%)
- Expected lower MC share, priority to new data

CPU utilization and breakdown by job types

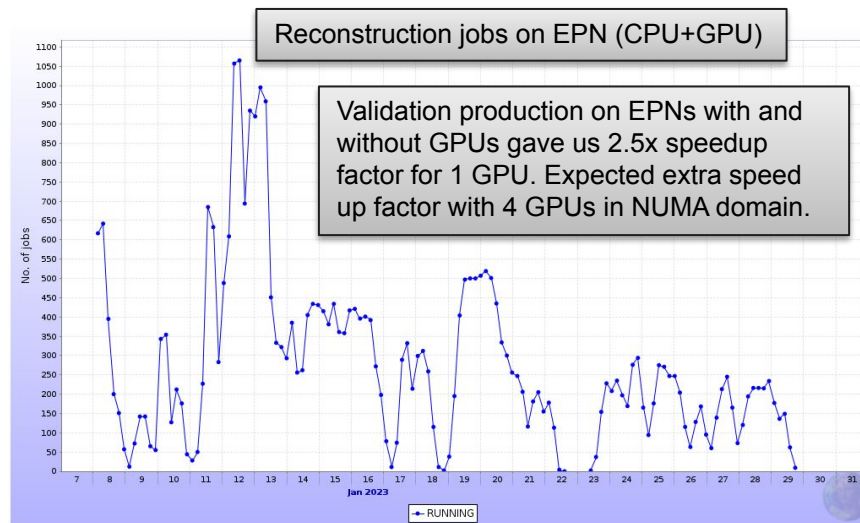
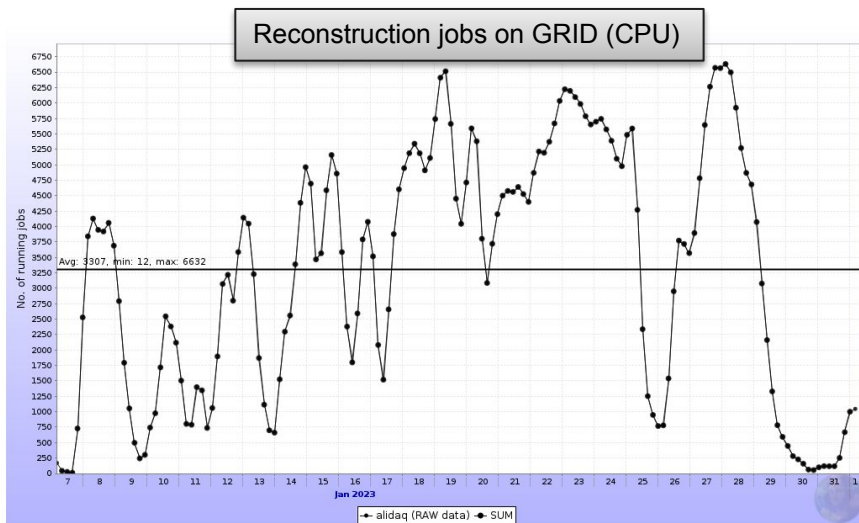


- Increasing Run 3 analysis share in the last months: hyperloop reached ~18% of CPU share
- While Run 2 analysis is steady at ~30%
- About 50% of our current CPU resources are used by analysis workflow

- Higher activity for raw calibration and reconstruction of Run 3 pp runs than previously reported
- Steady Run 2 analysis activity for conferences and publications (~30%)
- Expected lower MC share, priority to new data

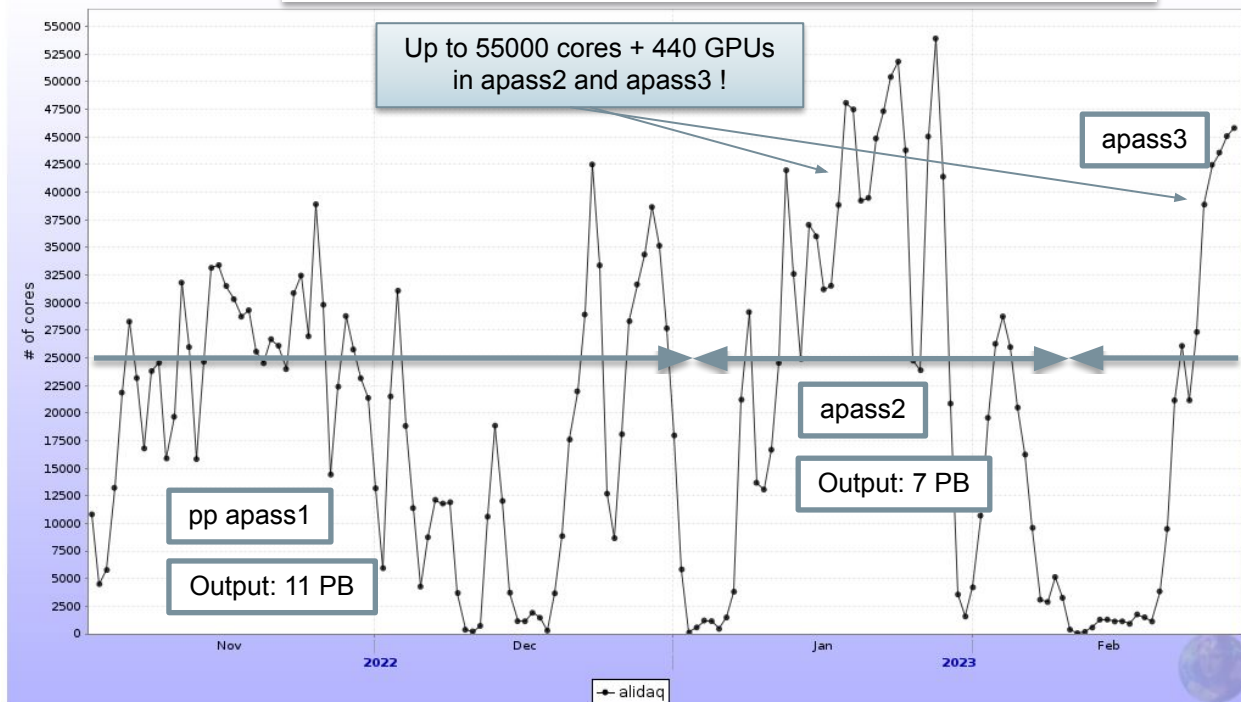
Asynch. reconstruction on GPU and CPU

- Asynch. reconstruction of 2022 pp data with CPU on GRID (8 cores per job) and with CPU+GPU (2 configurations has been tested in production: 16 virtual cores + 1 GPU per job and half EPN (64 virtual cores) in NUMA domain with 1 GPU per job)
- **Efficient reconstruction with GPU required tailored job tuning**, taking into account the GPU and CPU models and memory
- **EPN capacity was limited by outbound network connectivity, requires major network reconfiguration**
- The concurrent productions on GRID and on EPN allowed to speed up processing



2022 pp data processing

Concurrent CPU cores used for the async reconstruction (GRID+EPN)

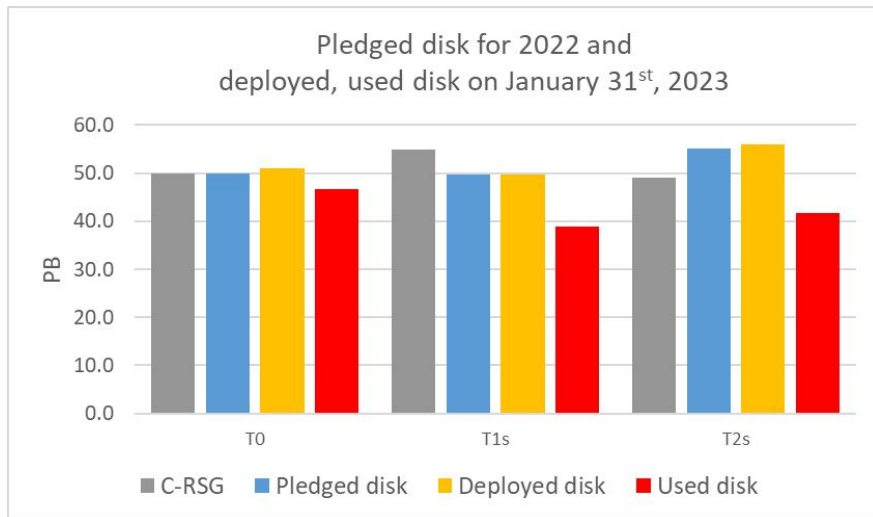


- Exploitation of GRID (CPU) and EPN (CPU+GPU) resources together allows to process a complete pass ($1 \cdot 10^{12}$ events) in **less than 25 days**: CPU + GPU
- Required 7-11 PB storage at T0, T1s (replicas) per pass
- Inbound connection from EOS02: average 30 GB/s but peak close to design value
- Outbound connection to T0, T1s SEs: average 6 GB/s but remote storage bandwidth is limited (10-100 Gb/s \rightarrow 1.2-12 GB/s) and shared with WLCG

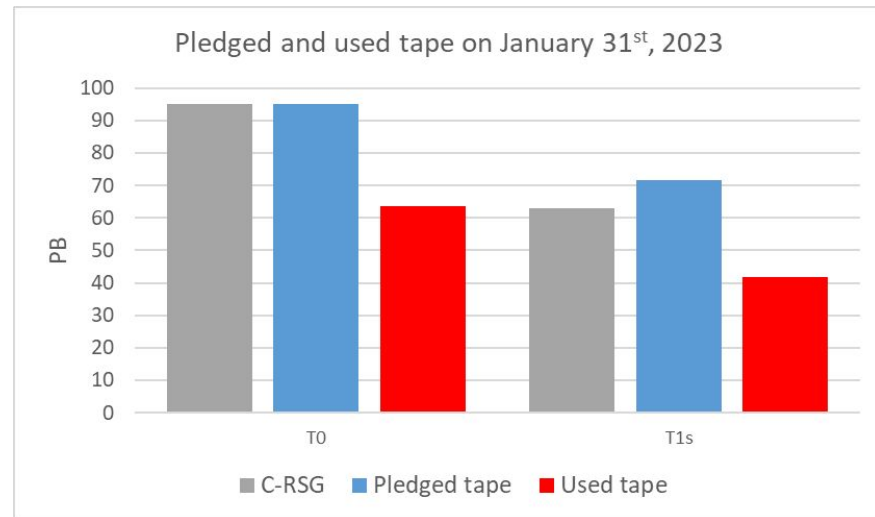
Run 3 MC productions and simulation estimates

- First Run 3 estimates of the simulation computing needs based on experience from Run 2 with Geant 3 (TDR and addendum)
- In 2022 ALICE performed extensive General-Purpose MB Run 3 simulations with Geant 4:
 - Unanchored simulation of pp at 500 kHz @ 13.6 TeV with realistic filling scheme
 - Unanchored simulation of Pb-Pb for the validation against 2018 Pb-Pb results
 - First productions anchored to 2022 periods: pp at injection and top energy + 2022 Pb-Pb tests
- As well as different signals, as prompt and non-prompt J/psi, Heavy Flavors, Multi-Strange and Jets have been injected into MB pp events for the selection studies
- Promising results on GRID with 8-core queues for pp and Pb-Pb GP MB:
 - Repeated the pp and Pb-Pb simulations on a reference machine
 - **The total improvement is about 4 times better than the estimates based on Run 2**
- Run 3 MC estimates has been updated accordingly:
 - Generate 4x more events with same CPU resources (ratio MC/data 8% instead of 2%, including embedding): increase of disk to store output (2.5 PB -> 10 PB for the whole Run 3)

DISK and TAPE utilization



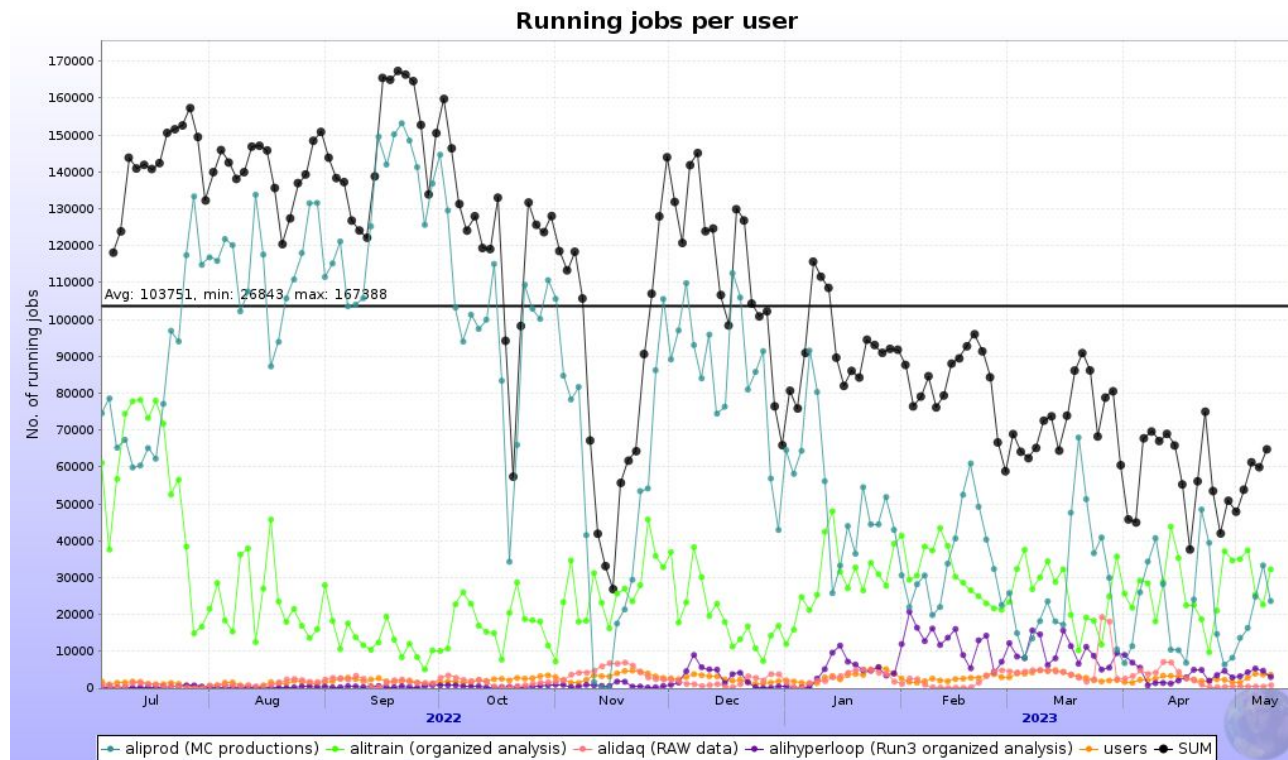
- Overall deployed / 2022 C-RSG ~100%:
 - Deficit at T1s, surplus at T2s
- Current disk use 93%, 78% and 76% of pledged capacity at T0, T1s and T2s respectively
- Deletion of apass1 ongoing (11 PB @ T0,T1s)



- Pledges / C-RSG: 100% at T0 and 114% at T1s (8.8 PB surplus)
- **Expected usage in line with the requested resources as for CPU and disk excluding Pb-Pb (about 52 PB available wrt C-RSG)**

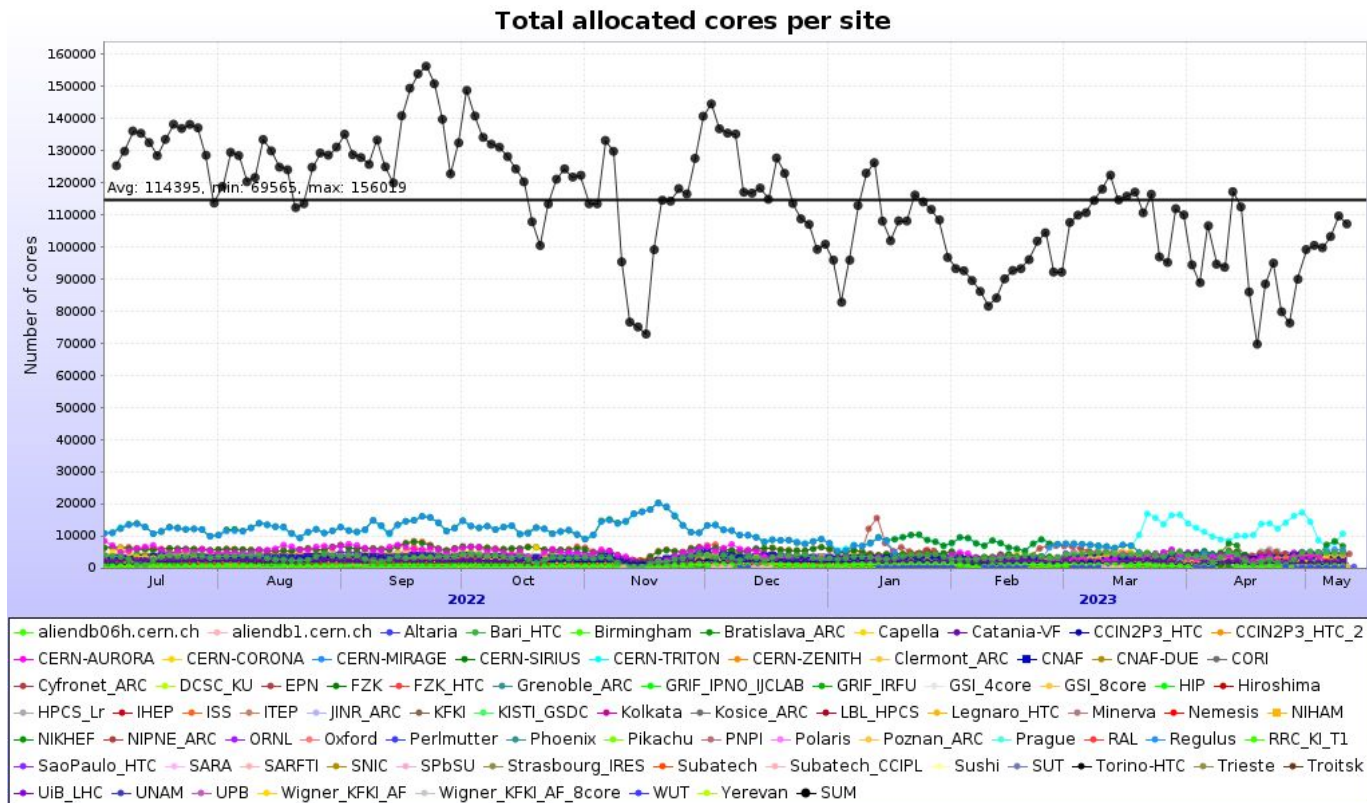
Job profile

Mix of single-core (alitrain), 1-2-4 core (hyperloop), 8-core (O2 MC and O2 RAW)



Cores profile

Profile still under-counts the number of running cores - to be fixed soon.





Computing resource and processing plans 2023 - 2024

Computing resources needs for 2023

- C-RSG endorsed our computing resource requests for 2023 in April 2022
- FAs entered their 2023 pledge input into CRIC in September 2022
- no HI at high rate in 2022 has reduced the planned resource usage in 2023 but
 - Measured larger CTF average event size during 2022 pp data taking wrt MC (+50% @ 650kHz)
 - Strategy under study to reduce the impact on tape:
 - The assumption for the computing requirements is +30% CTF average event size **as an upper limit** wrt our previous estimates
 - Adoption of more aggressive compression strategy B postponed to 2024 (after HI)
 - While the requests for 2023 were based on strategy B
 - Conservative strategy A impacts on the needed tape for 2023
 - The LHC schedule for 2023 and 2024 has been re-discussed and it will require more computing resources to cope with the new conditions, notably longer HI period in 2023 (5.5 w):
 - excluding setup: 27 days Pb-Pb data taking + 5 days of pp reference run
 - **Pb-Pb 2023 target 3.25 nb⁻¹ + pp ref runs target 3 pb⁻¹ + pp @ 13.6 TeV target 30 pb⁻¹**

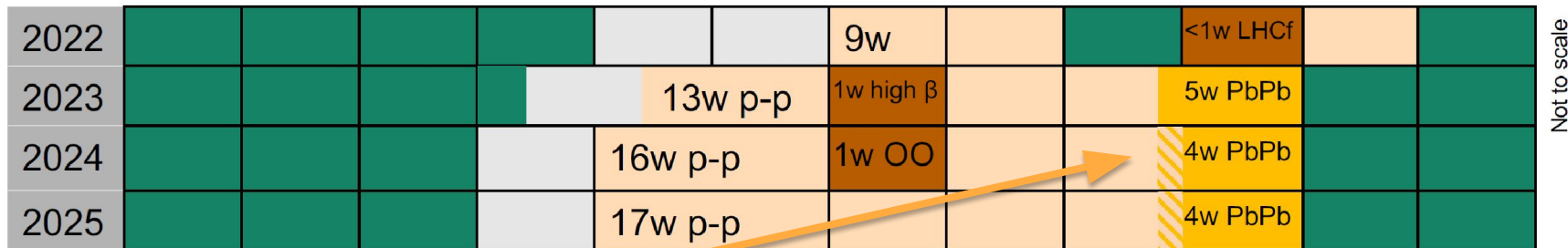
Projections of tape and disk needs for 2023

- Tape most affected by strategy A and larger number of clusters in TPC:
 - Considered for 2023 a scenario with CTF average event size at +30% **as an upper limit**
 - Available tape in 2023 **~109 PB** (2023 pledges - expected used tape in Apr 2023)
 - pp low-field events are about 30% larger with respect to the full-field due to the lower p_T cutoff
 - considering the short pp run in 2023 and the impact on tape requests:
 - **in 2023 only 10%** of Run 3 luminosity goal (**3 pb⁻¹**), while the remaining 90% in 2024

| Collision Type | pp | pp low field | pp ref run | Pb-Pb | Total Size |
|-----------------------|-------|--------------|------------|---------|------------|
| Integrated Luminosity | 30/pb | 0.3/pb | 3/pb | 3.25/nb | |
| Strategy B MC [PB] | 0.7 | 0.7 | 4.1 | 47.7 | 53.2 |
| Strategy A MC [PB] | 1.3 | 1.2 | 7.6 | 71.5 | 81.6 |
| Strategy A +30% [PB] | 1.7 | 1.5 | 9.9 | 92.9 | 106.0 |

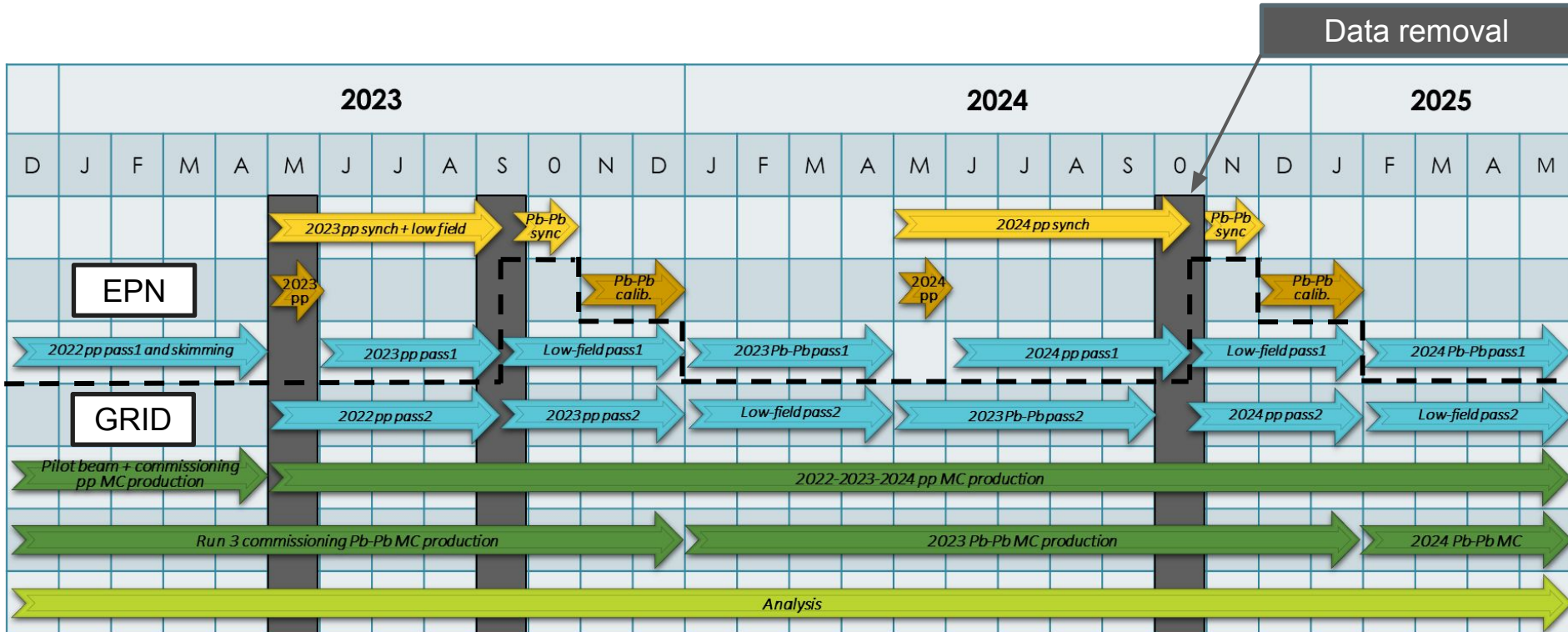
- Being the 2022 HI postponed to 2023, disk used in 2022 was less than requested:
 - AOD average event size not affected by larger number of clusters in TPC
 - 29 PB of available disk in 2023 – 18 PB needed in 2023: 11 PB surplus to be carried over

Baseline scenario for 2024

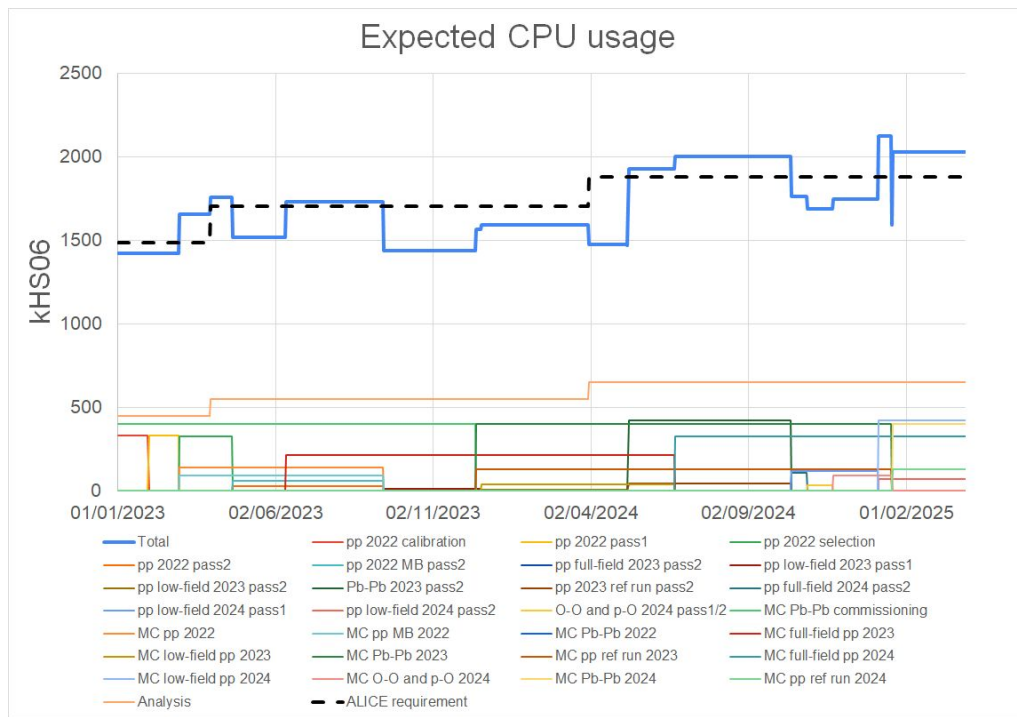


- Assumed that the HI run in 2024 could be extended to 5 weeks
- Same luminosity goals of 2023 for Pb-Pb and pp ref runs:
 - **3.25 nb⁻¹ of Pb-Pb collisions (strategy B aggressive)**
 - **3 pb⁻¹ of pp ref run**
- Such an assumption accommodates with some margin, all the different possible scenarios for the HI period in 2024.
- Considered as **upperlimit**:
 - **112 days of pp in 2024:**
 - ~42 pb⁻¹ of pp full-field
 - ~2.8 pb⁻¹ of pp low-field
 - **Short O-O and p-O run:**
 - 1 nb⁻¹ and 5 nb⁻¹, respectively

Updated 2023-2024 processing timeline




CPU needs for 2023 - 2024



- The blue line is the minimum CPU capacity needed to process all the planned productions on GRID, while the dashed line shows the ALICE requests
- The achieved performances of the asynch. reconstruction on EPN with GPUs allow to lower our 2024 CPU request from 1960 kHS06 (previous estimate) to 1880 kHS06

Disk and tape needs for 2024

- Disk: AOD average event sizes are unchanged with respect to 2022 and 2023 requests
- Tape: considered the adoption of compression strategy B (aggressive) in 2024:
 - But with **CTF average event size at +30%** as an upper limit for strategy B as well

| ALICE | | 2024 | | | | | | | | | | Total | Total - carry over from 2023 |
|--------------|--------------|--|----------------|---------------|----------------------|------------|----------------|---------------|---------------------|----------------------|-------------|-------------|---------------------------------------|
| | | pp 2023 | pp ref 2023 | Pb-Pb 2023 | pp low field 2023 | pp 2024 | pp ref 2024 | Pb-Pb 2024 | O-O and p-O 2024 | pp low field 2024 | | | |
| | | <i>To be processed in 2024</i>  | | | | | | | | | | | |
| Disk [PB] | Tier-0 | 0.0 | 1.4 | 4.9 | 0.2 | 1.6 | 0.7 | 2.3 | 0.2 | 1.5 | 12.8 | 9.3 | |
| | Tier-1 | 0.0 | 1.3 | 4.7 | 0.2 | 0.8 | 0.7 | 2.3 | 0.2 | 1.5 | 11.7 | 8.2 | |
| | Tier-2 | 0.0 | 1.4 | 5.1 | 0.2 | 0.5 | 0.7 | 2.4 | 0.2 | 1.6 | 12.1 | 8.2 | |
| | Total | 0.1 | 4.1 | 14.7 | 0.5 | 2.9 | 2.0 | 7.0 | 0.7 | 4.6 | 36.7 | 25.7 | |
| Tape [PB] | Tier-0 | 0.0 | 0.0 | 0.0 | 0.0 | 1.6 | 3.7 | 41.3 | 0.4 | 5.4 | 52.4 | 55.0 | |
| | Tier-1 | 0.0 | 0.0 | 0.0 | 0.0 | 0.8 | 1.9 | 20.6 | 0.2 | 2.7 | 26.2 | 19.9 | |
| | Total | 0.0 | 0.0 | 0.0 | 0.0 | 2.4 | 5.6 | 61.9 | 0.6 | 8.1 | 78.7 | 74.9 | |

Summary

- **Computing resource utilization:**
 - Full utilization of CPU resources
 - EPN CPU and GPU resources successfully exploited for the processing of pp data
 - Disk and tape expected usage in line with the requested resources excluding Pb-Pb
- **Computing resources needs for 2023 with the updated Run 3 schedule:**
 - The postponed 2022 HI data taking lowers our CPU and disk needs in 2022-2023
 - Re-assessed tape needs with strategy A with larger average event size (+30%)
 - and with longer HI period in 2023
- **Resource requests for 2024:**
 - Considered the carryover from 2023, step for tape (+75 PB)
 - CPU and disk in 2024 compatible with flat budget considering our 2023 requests
- **Sizeable impact of the war in Ukraine: RU resources needed to be replaced by 2024**