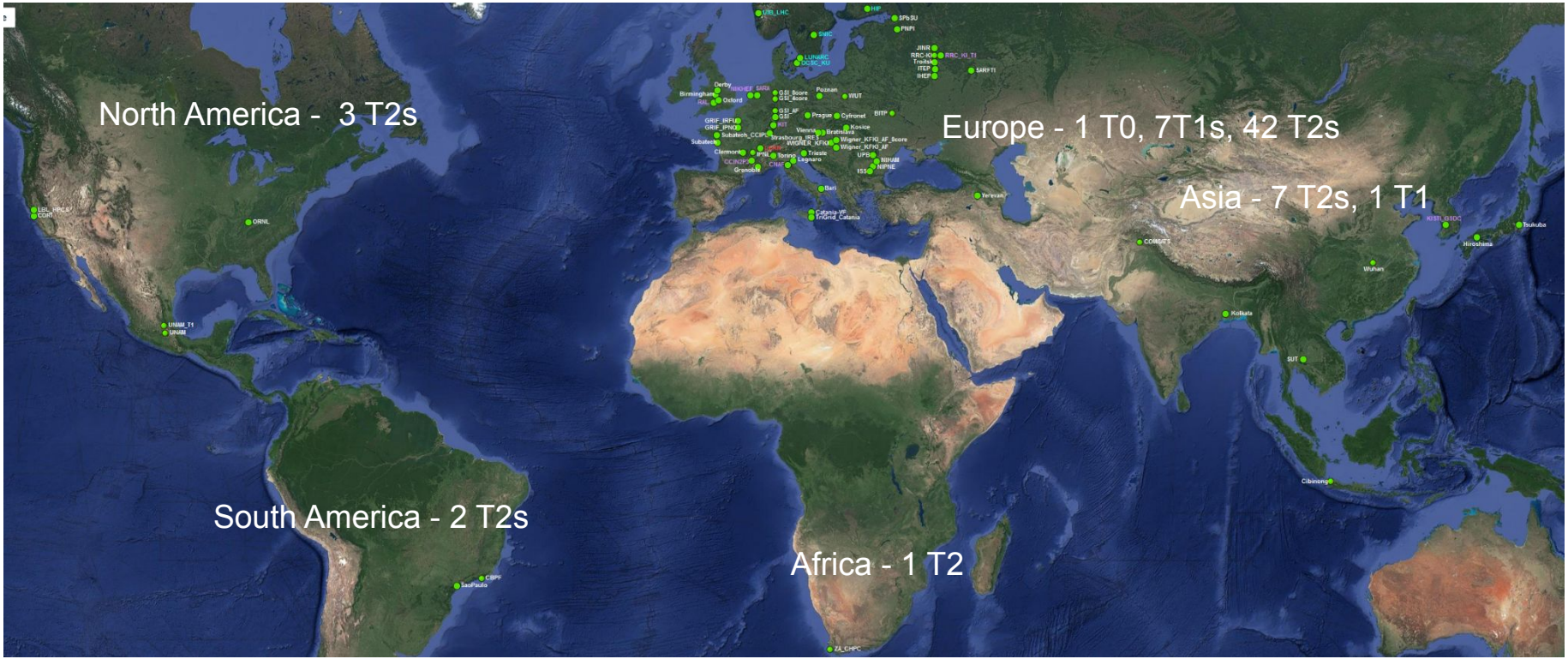
The ALICE Grid upgrade, methods and tools for LHC Run 3 and beyond

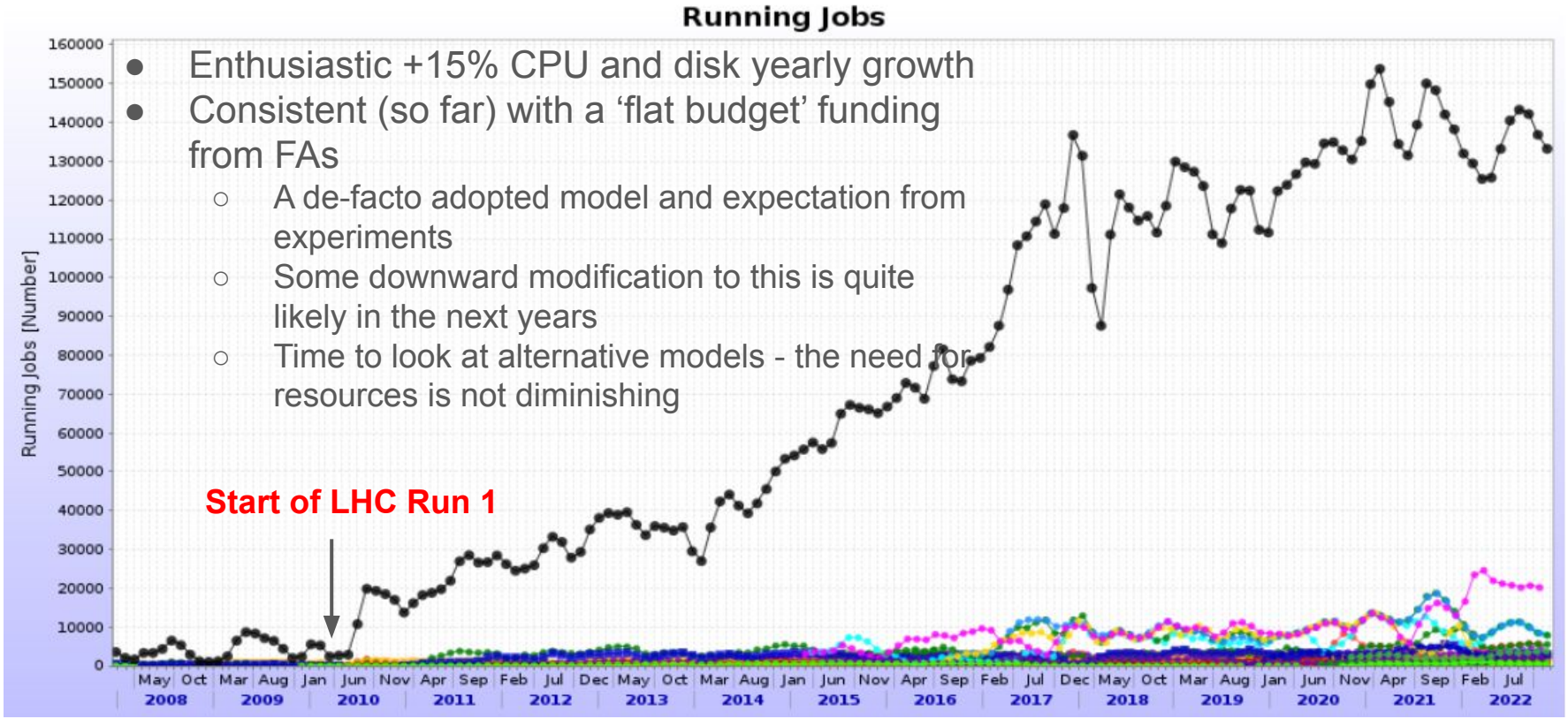
L. Betev

6th Asia Tier Center Forum, Krabi, Thailand, November 21-24 2022

The ALICE Grid - individual computing centres

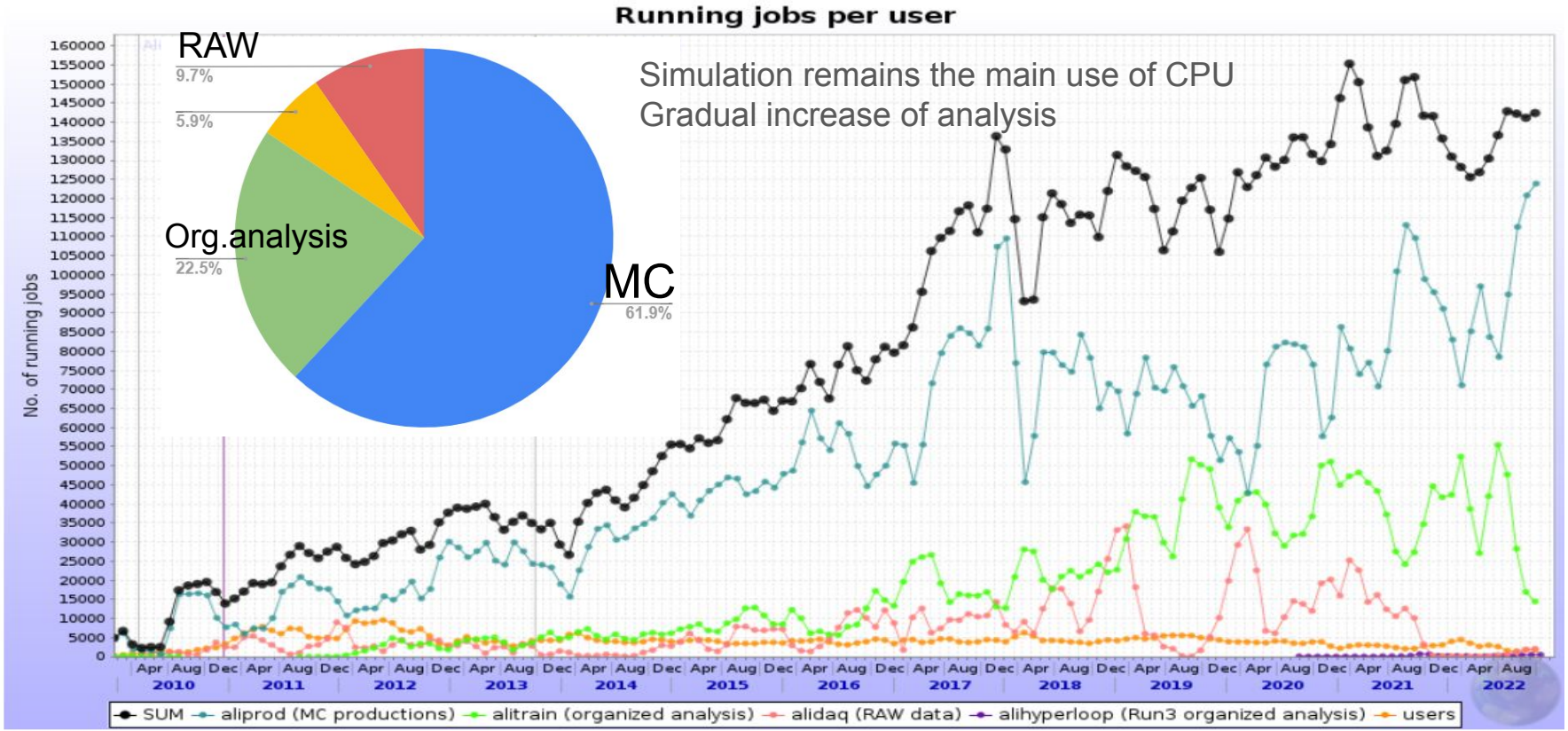


ALICE resources evolution



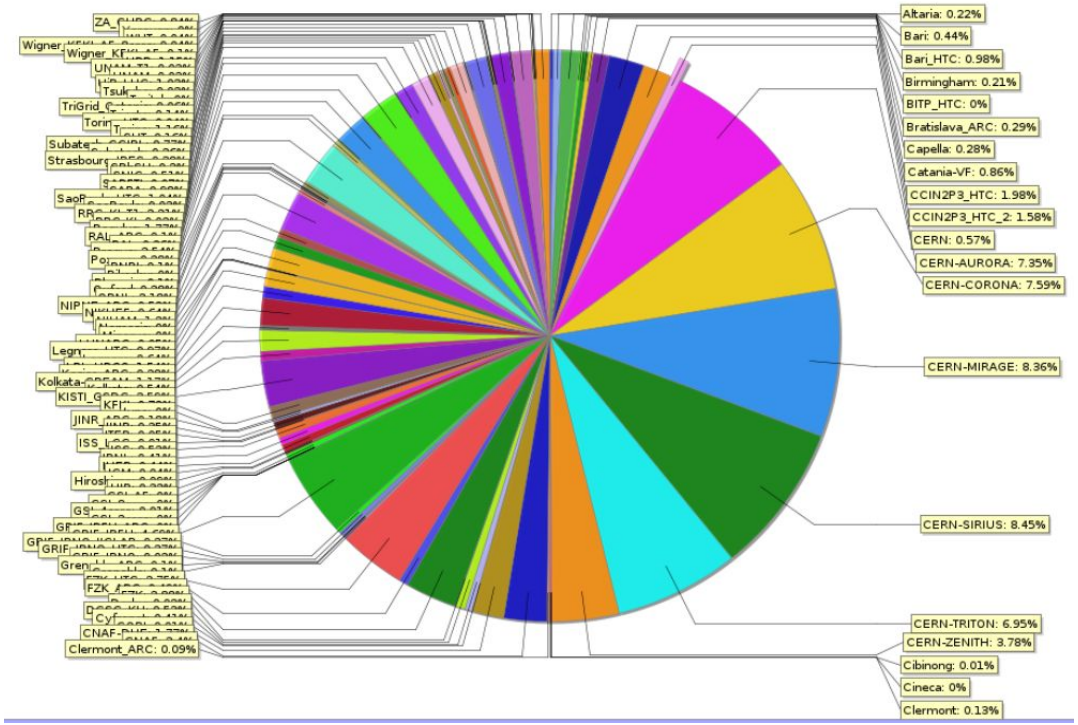
- Enthusiastic +15% CPU and disk yearly growth
- Consistent (so far) with a ‘flat budget’ funding from FAs
  - A de-facto adopted model and expectation from experiments
  - Some downward modification to this is quite likely in the next years
  - Time to look at alternative models - the need for resources is not diminishing

ALICE resources use per activity



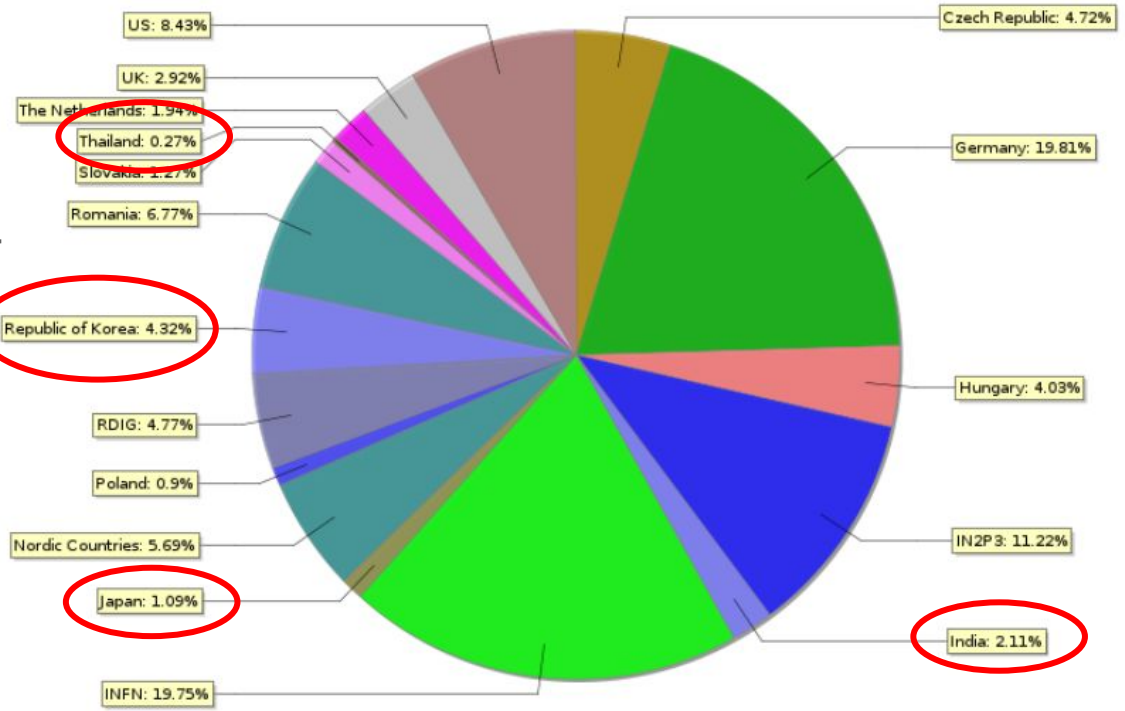
Role of Tiers

- T0 - RAW reco + MC + analysis
- T1s - RAW reco + MC + analysis
- T2s - MC + Analysis
- Differences between tiers - custodial storage + nominal services response time
- In practice - all tiers run effectively all types of workload (except RAW reco) and availability is ~same
- ALICE model can absorb any site size



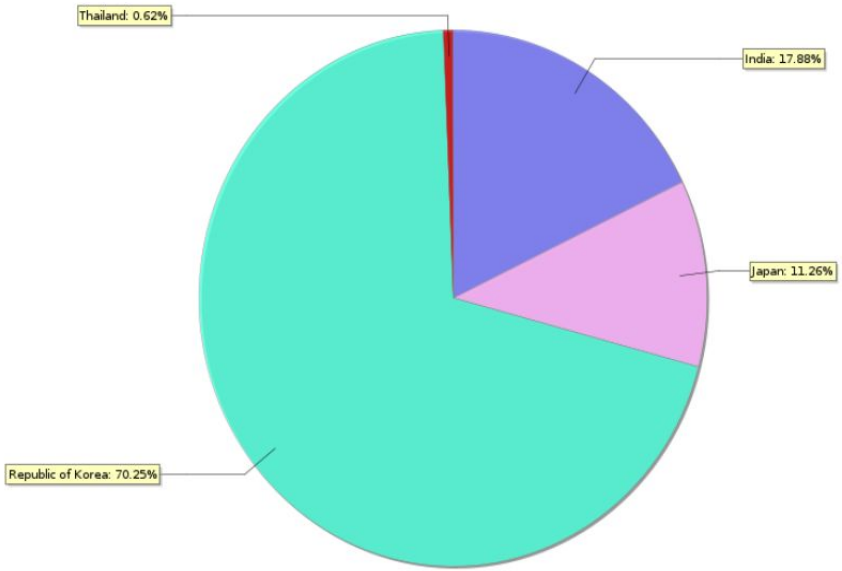
Regional contribution

- ~8% Asian contribution
- Last T1 remaining in Asia@KISTI
- The growth of Asian resources is similar to other regions
- The growth potential is there, as well as the expertise
- Diminishing role of smaller T2 centres - this is an unfortunate global trend

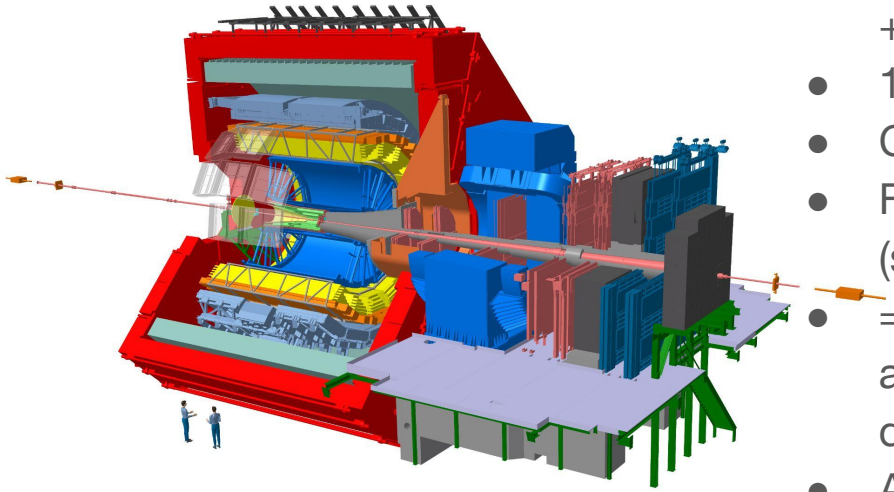


Role of Asian sites - resources and innovation

- KISTI remains the largest contributor
  - In production - first of its kind (worldwide) disk-based custodial storage element 12PB
- Hiroshima - site restarted with new storage and upgraded capacity
  - Another T2 in Nagasaki is being set up
- Kolkata - steady resources increase and normal operation
- SUT - a small T2, but consistent operation
  - Important as example for other Asian countries entering the Grid

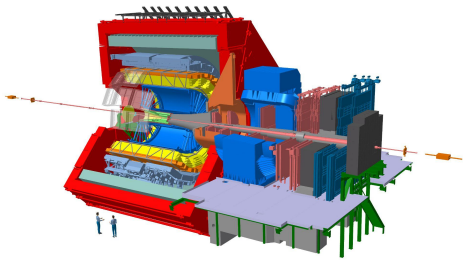


ALICE upgrade general



- p-p and HI physics
- 10x integrated luminosity L~10nb^{-1} (B=0.5T) + 3nb^{-1} (B=0.2T)
- 100x event rate of Run 1/2, 10x more data
- Continuous readout
- Focus on data compression and real time (synchronous) data reconstruction
- => Reasonable rates and data volumes after compression to storage and secondary data formats
- Adherence to ‘flat budget’ resources funding for data processing and analysis

Data flow and data rates



RAW data 3.5TB/s

First Level Processors (FLP)

STF - data from single FLP

600GB/s

Event Processing Nodes (EPN)

Compressed Time Frame (CTF) 10ms

100GB/s

Data buffer (80PB)

Custodial storage T0 ⅔ T1s ⅓

CTFs 8GB/s

CTF 20GB/s

T0/T1 Async. Processing, Analysis

Local storage

AOSs, MCAODs ~10%

Analysis Facilities AODs, MCAODs

T2s MC production, Analysis

Local storage

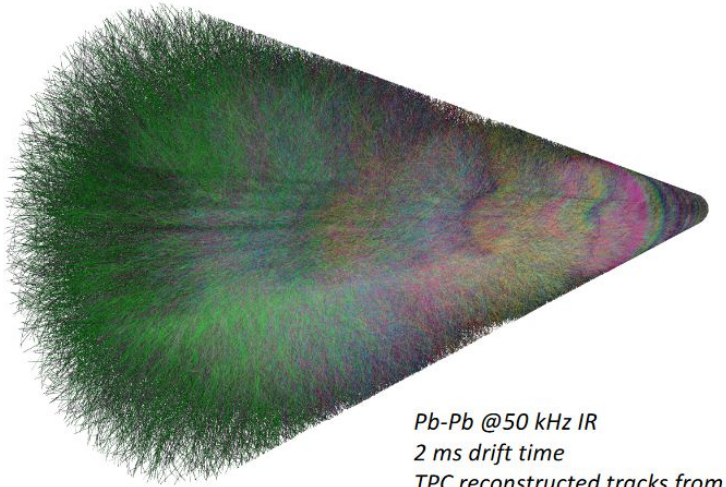
AOD - Analysis Object Data
MCAOD - AOD from MC

The O2 facility (EPNs)

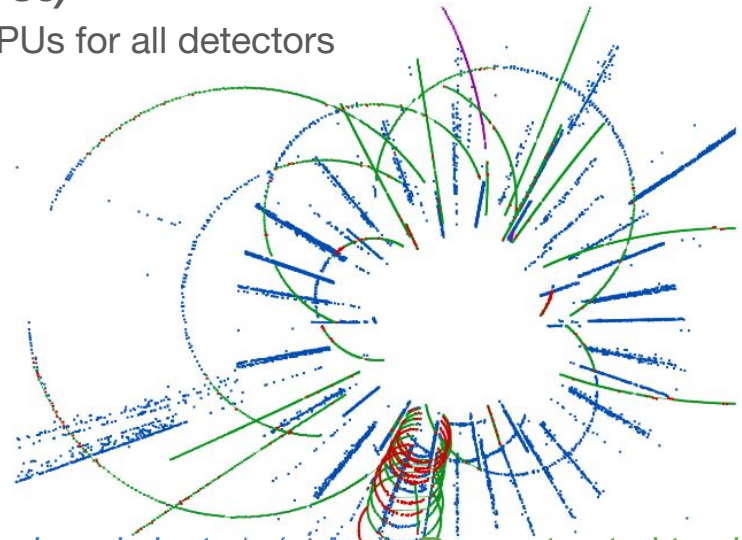
- Container-hosted computing facility located at the ALICE site, PUE<1.07
- High-throughput system, heterogeneous computing platform (CPU+GPU)
- 250 dual CPU nodes (ROME, 64 cores, 512GB RAM) with 8 AMD (MI50, 32GB) GPUs/node
- Functions
  - Data aggregation (Detector STFs to global CTF)
  - Synchronous global reconstruction
  - Calibration and data volume reduction
  - Quality control
  - Asynchronous (offline) reconstruction
- Containers house a backup EOS storage in case of network interruption to CC

Synchronous data processing

- Goal - to compress the RAW data by about factor 35 (3.5TB/s -> 100GB/s)
- Through zero suppression, clusterization, tracking, optimized data format
  - ***Mandatory use of GPUs (40x faster than CPUs)***
  - All synchronous level software is written for GPUs for all detectors

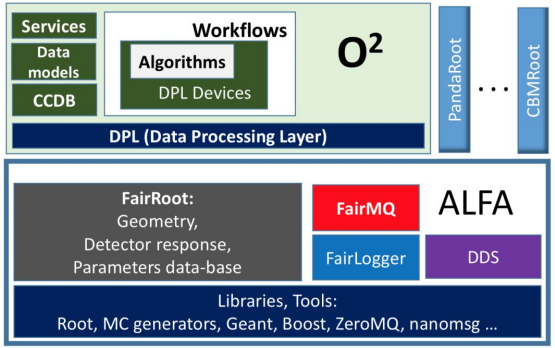


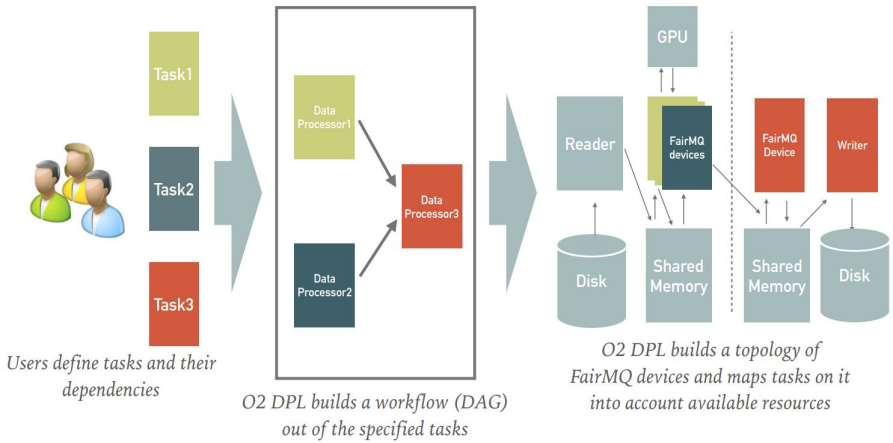
Pb-Pb @50 kHz IR
2 ms drift time
TPC reconstructed tracks from different colour-coded events



Unassigned clusters (noise) Reconstructed tracks
Removed clusters Failed fits

O2 Software framework





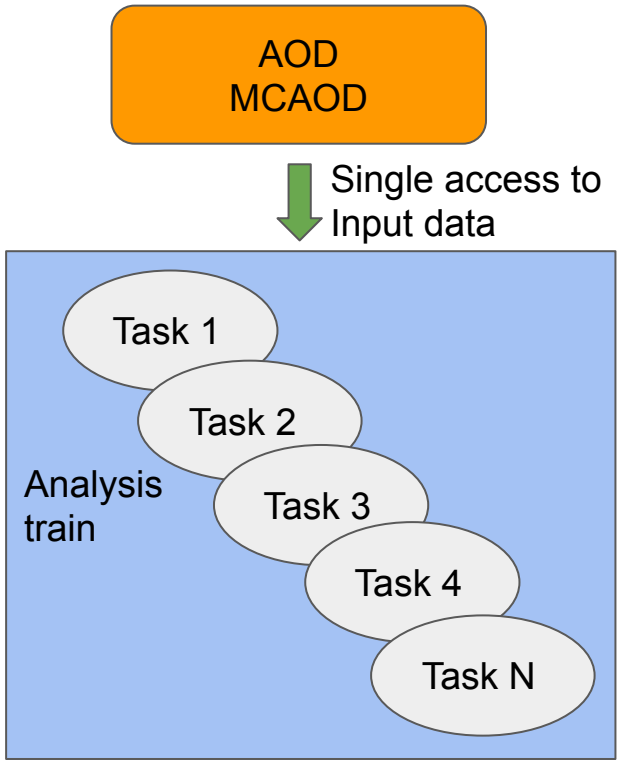
Users define tasks and their dependencies

O2 DPL builds a workflow (DAG) out of the specified tasks

O2 DPL builds a topology of FairMQ devices and maps tasks on it into account available resources

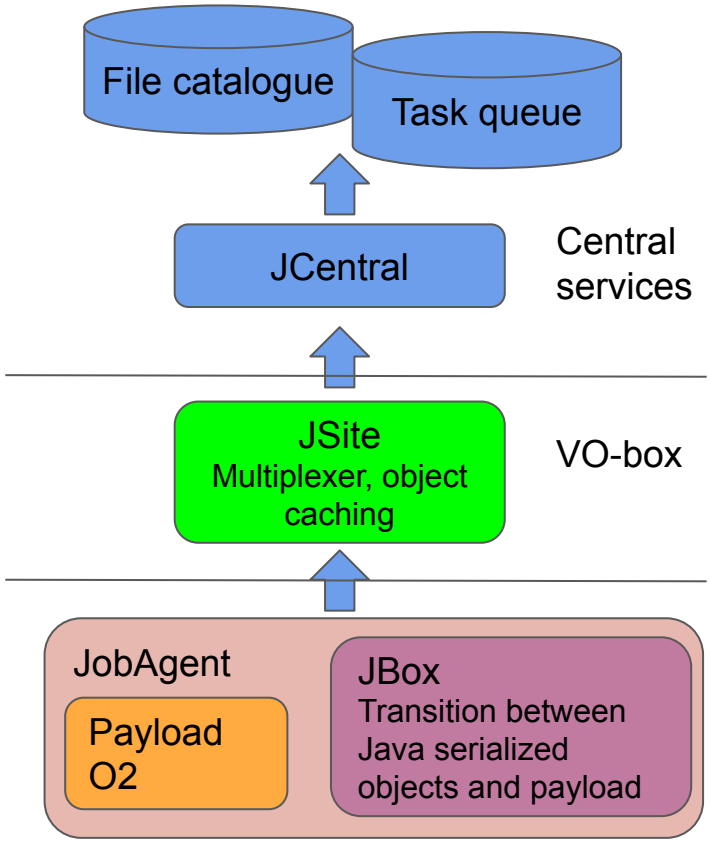
- Developed by ALICE in collaboration with the FAIR group at GSI Darmstadt
- Three major parts
  - Transport layer, based on FairMQ - message passing toolkit
  - Data model - ALICE-specific object description and content
  - Data processing layer - set of data processors implicitly organized in a logical dataflow for data transformation
- Trivially parallel and integrates tools for GPU offloading
- Natural use of multicore processing and shared memory

Analysis facilities (AFs)



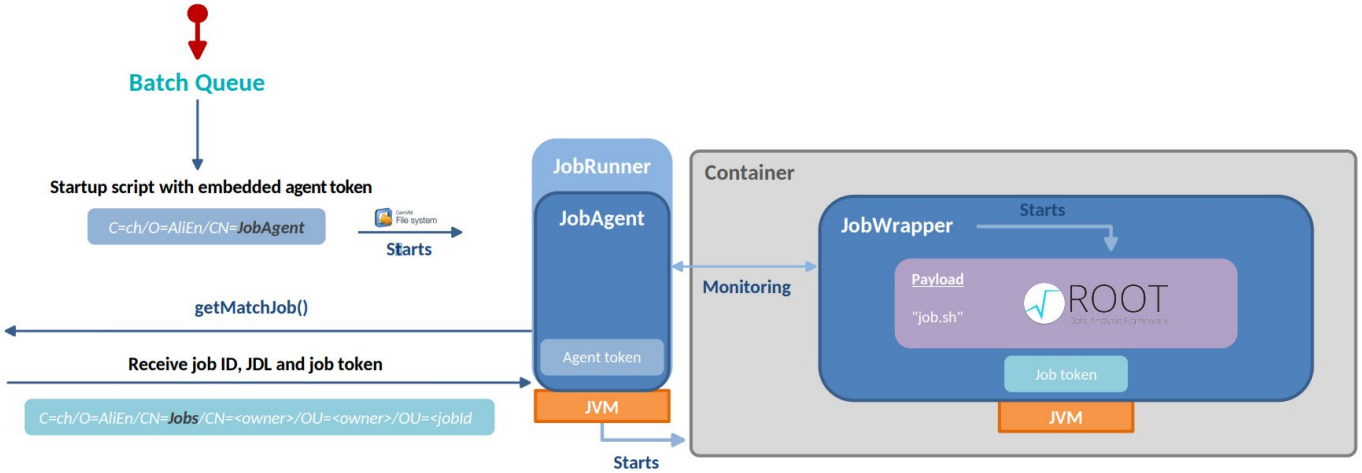
- New element of the computing model
- Data transferred to AF from T0/T1s/T2s
- Goals
  - Provide a location with comprehensive data samples from asynchronous and MC data processing at ~10% statistics
  - Fast tuning of analysis algorithms - once ready, run on full sample on the Grid
  - First data and low statistics analysis (if compatible)
- Incorporated in the Grid framework
- Sites tuned for fast I/O between storage and CPU
  - Approximate total size 6-8k cores, 10PB storage
  - ~15MB/s/core throughput
- As of today - GSI Darmstadt and KFKI Budapest (⅔ of the AF target, looking for more suitable sites)

Grid middleware development - JAliEn



- Evolution of the AliEn middleware
  - Refactored and rewritten in Java
- Highly efficient and scalable communications infrastructure
- Persistent, compressed, SSL channels
- Multiplexing and object caching
- Use of Java serialized objects
- Platform independent
- Multi-core enabled, HPC ready
- Deployed gradually on the existing infrastructure - no interference with operations

JobRunner, JobAgent and JobWrapper



- Entirely new method for both resources and job control
- Fully containerized workload
- Ability to run multiple jobs within the control of the same JobRunner
  - Effective control of any set of resources provided

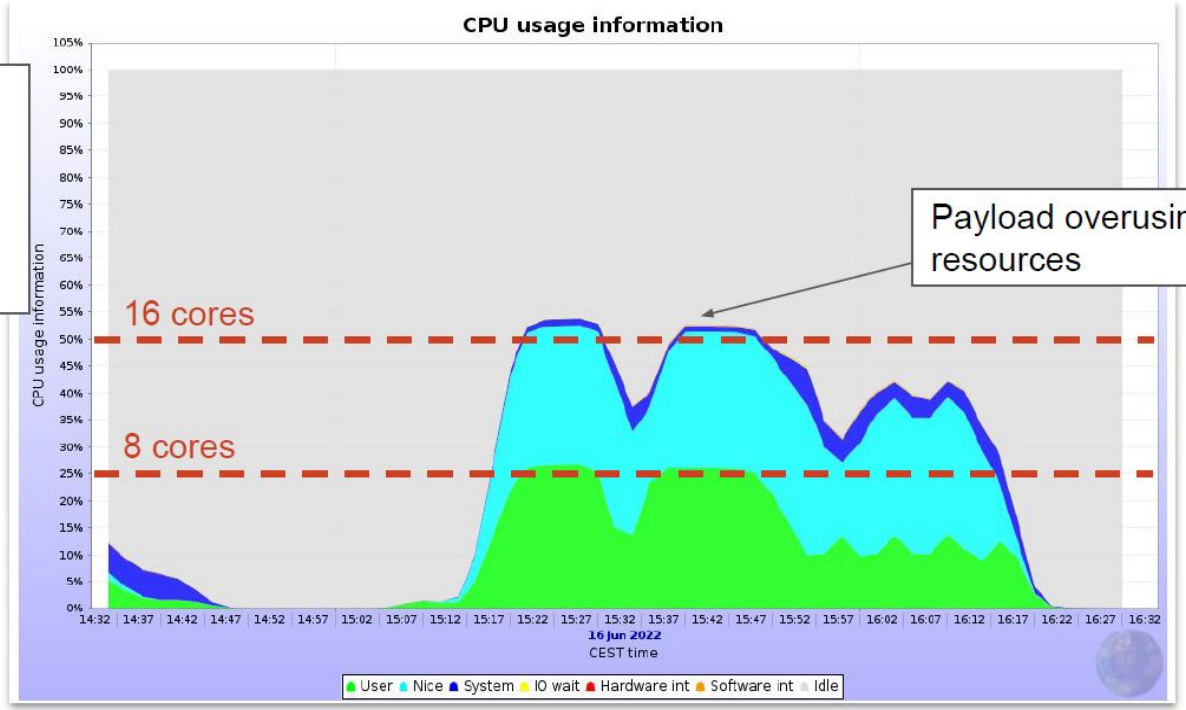
ALICE

Payload containers

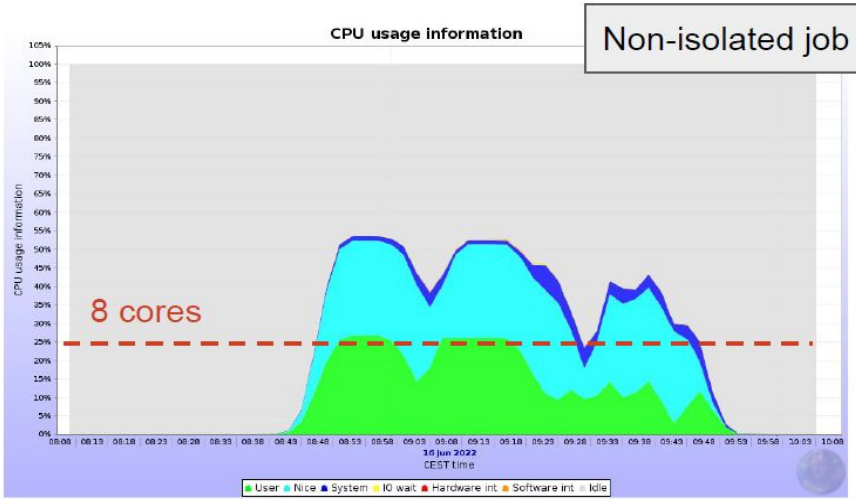
- By default, all jobs are wrapped in a CentOS 7.9 container
- Other images are available
  - Rocky 8.6: For newer payloads and ***GPUs***
  - Rocky 9.0: Testing only
  - Debug containers, for example with vtune, strace
- GPUs are supported when using Apptainer (formerly Singularity), i.e. fully
- All of the above allows for fulfilling various job requirements, independent of the underlying OS
- Use in HPCs or other specialized clusters (for example EPNs)

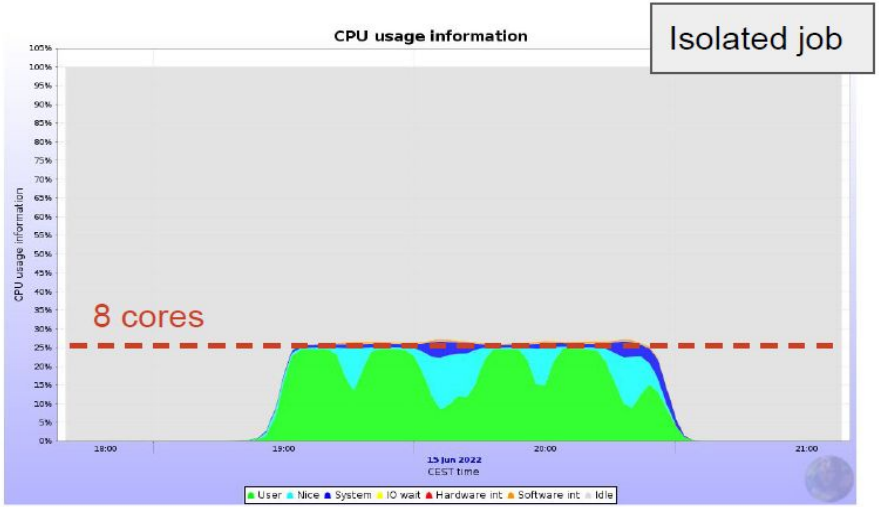
Job isolation and control

CPU usage of a 32-core idle machine running a simulation payload that *requests 8 CPU cores*.



Job isolation and control - applying `taskset`



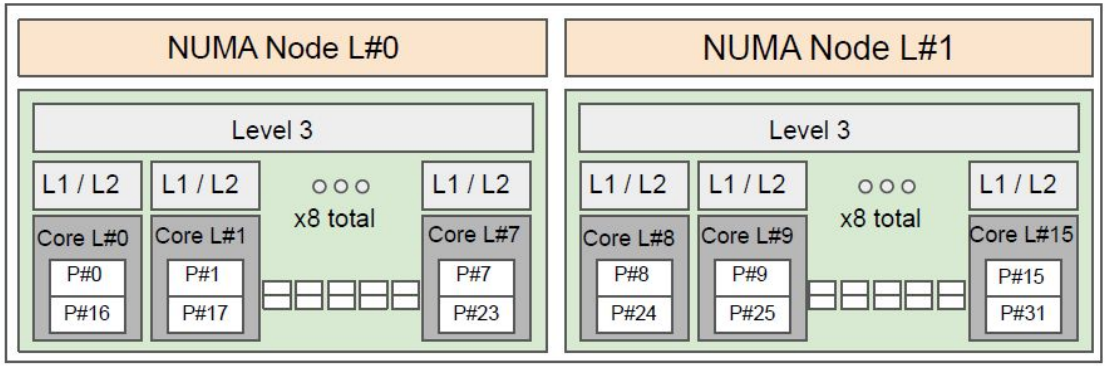


- Total CPU usage goes above the requested 8 cores

- CPU consumption is limited with `taskset`
- Total CPU usage is flat at 8 cores

- Applicable for sites with non-constrained resources and full node submission

Improving job efficiency through CPU pinning

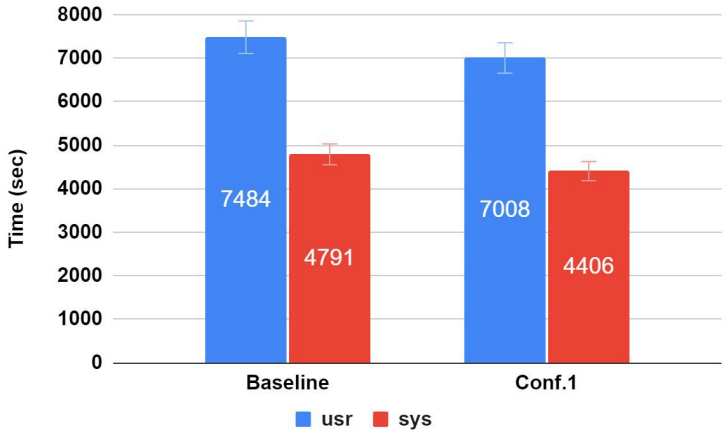


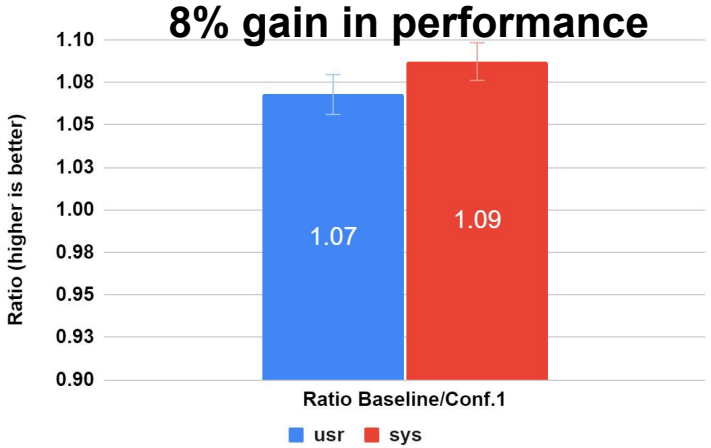
Sample host CPU architecture

- Various core/cache pinning configurations possible
  - Same NUMA Node and independent L1,L2 cache
  - Different NUMA Nodes and independent L1,L2 cache
  - Same NUMA Node and sharing L1,L2 cache
  - Random core assignment
  - No pinning

Improving job efficiency through CPU pinning

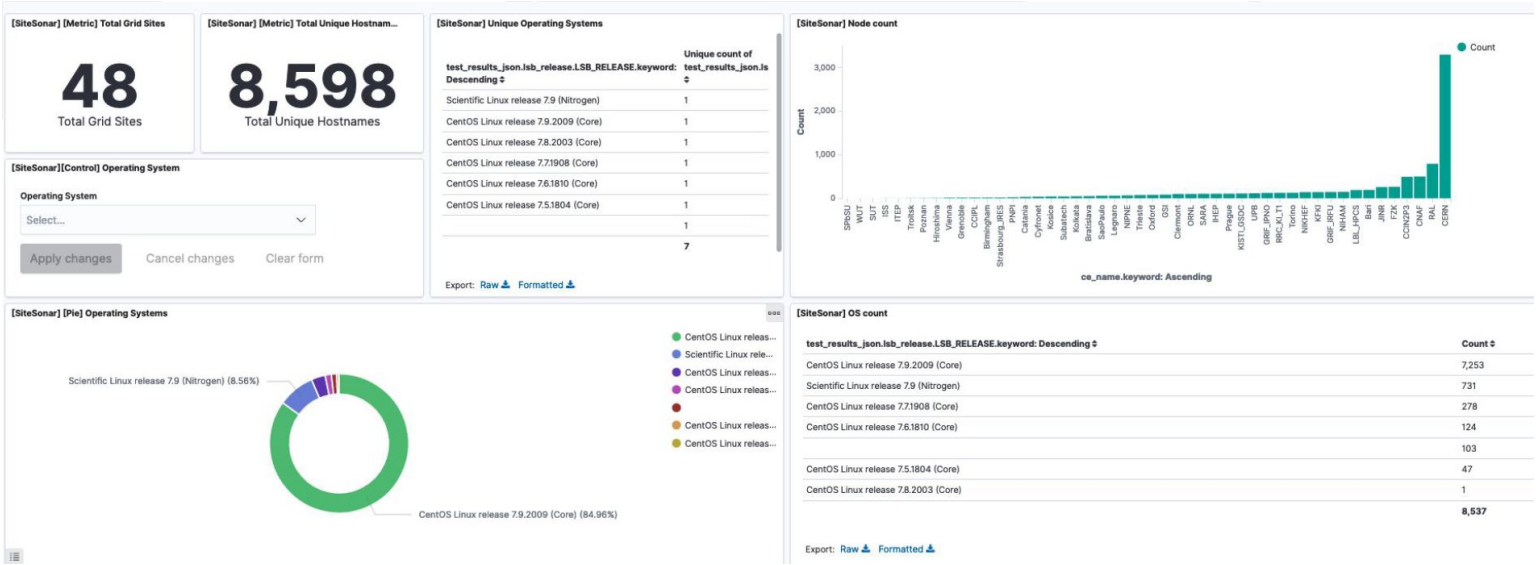
- Most efficient configuration - same NUMA node, independent L1/L2 cache - compared to no pinning
- Only possible if full control of the CPU - whole node
- Already in production at LBNL - Lawrencium HPC





New tools and monitoring

- Kibana dashboard of SiteSonar
  - Tool to evaluate site capabilities and installations - probes invoked at the beginning of execution
  - Collects data from ~10K Grid nodes daily



New tools and monitoring (2)

- File crawler
  - Checks storage integrity on sites by mimicking normal jobs
  - Random files, proportional to the storage size
  - Reporting on file health, throughput and accessibility
  - Early detection of storage issues

Status codes extracted from the crawler

SE Name: ALICE::HIROSHIMA::EOS **Interval:** Last week

Status Type	Status Code	Status Count	Status Code Ratio	Download throughput
FILE_OK	S_FILE_CHECKSUM_MATCH	26972	99.79 %	21.97 Mb/s
	E_CATALOGUE_MD5_IS_BLANK	2	0.01 %	19.04 Mb/s
INTERNAL_ERROR	XRDFS_CANNOT_CONFIRM_UPLOAD	21	0.08 %	
FILE_INACCESSIBLE	XROOTD_EXITED_WITH_CODE	35	0.13 %	
TOTAL		27030	100 %	

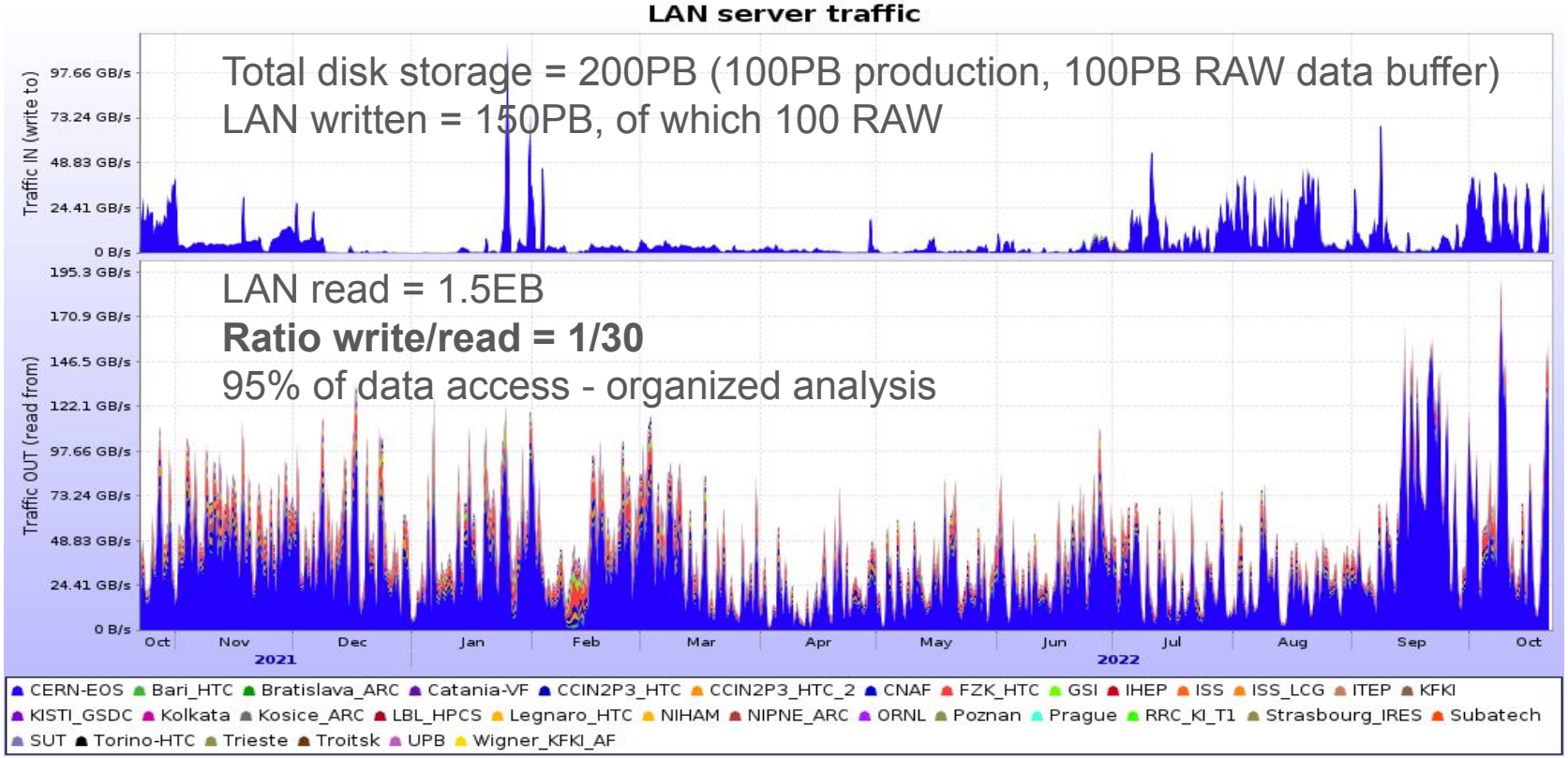
Averaged metrics for the selected interval

SE Name	Start	End	Success ratio	Corrupt ratio	Inaccessible ratio	Internal error ratio
SARA::DCACHE	18 Oct 2022 06:08	17 Nov 2022 10:32	99.87 %	0.09 %	0.05 %	0.00 %
Hiroshima::EOS	18 Oct 2022 06:08	17 Nov 2022 10:33	99.73 %	0.00 %	0.18 %	0.09 %
SNIC::DCACHE	18 Oct 2022 06:12	17 Nov 2022 10:28	99.68 %	0.02 %	0.27 %	0.03 %
Vienna::EOS	18 Oct 2022 06:07	17 Nov 2022 10:38	99.60 %	0.24 %	0.16 %	0.00 %
NIPNE::EOS	18 Oct 2022 06:09	17 Nov 2022 13:03	99.58 %	0.03 %	0.37 %	0.03 %
Trieste::SE	18 Oct 2022 06:11	17 Nov 2022 12:11	99.54 %	0.11 %	0.35 %	0.00 %
Bari::SE	18 Oct 2022 06:04	17 Nov 2022 12:22	99.50 %	0.08 %	0.42 %	0.00 %
IHEP::SE	18 Oct 2022 06:07	17 Nov 2022 10:20	99.35 %	0.11 %	0.53 %	0.01 %
Torino::SE2	18 Oct 2022 06:09	17 Nov 2022 11:07	99.34 %	0.13 %	0.53 %	0.00 %
Troitsk::SE	18 Oct 2022 06:04	17 Nov 2022 10:43	99.26 %	0.54 %	0.19 %	0.01 %
CERN::EOS	18 Oct 2022 06:12	17 Nov 2022 10:47	99.19 %	0.08 %	0.65 %	0.07 %
CNAF::SE	18 Oct 2022 06:10	17 Nov 2022 10:35	99.06 %	0.02 %	0.92 %	0.00 %
FZK::SE	18 Oct 2022 06:11	17 Nov 2022 10:33	98.86 %	0.06 %	1.07 %	0.01 %
Legnaro::SE	18 Oct 2022 06:04	17 Nov 2022 10:26	98.54 %	0.03 %	1.34 %	0.09 %
UPB::EOS	18 Oct 2022 06:08	17 Nov 2022 10:32	98.49 %	0.07 %	1.44 %	0.00 %
ORNL::EOS	18 Oct 2022 06:06	17 Nov 2022 10:31	98.18 %	0.46 %	1.36 %	0.00 %
NDGF::DCACHE	18 Oct 2022 06:04	17 Nov 2022 10:30	97.89 %	0.23 %	1.87 %	0.00 %
NIHAM::EOS	18 Oct 2022 06:08	17 Nov 2022 10:49	97.75 %	0.12 %	2.13 %	0.00 %
GRIF::EOS	18 Oct 2022 06:05	17 Nov 2022 10:31	97.75 %	0.05 %	2.20 %	0.00 %
Subatech::EOS	17 Oct 2022 17:38	16 Nov 2022 16:28	97.46 %	0.06 %	0.91 %	1.57 %
JINR::EOS	18 Oct 2022 06:11	17 Nov 2022 12:13	95.93 %	0.13 %	3.92 %	0.03 %
RRC_KI_T1::EOS	18 Oct 2022 06:06	17 Nov 2022 10:28	95.86 %	0.09 %	1.47 %	2.57 %
KISTI_GSDC::EOS	18 Oct 2022 06:07	17 Nov 2022 10:57	95.04 %	3.49 %	1.47 %	0.01 %
CCIN2P3::SE	18 Oct 2022 06:11	17 Nov 2022 10:37	94.27 %	0.02 %	5.69 %	0.02 %
Kosice::EOS	18 Oct 2022 06:07	17 Nov 2022 11:40	93.05 %	0.11 %	6.84 %	0.00 %
Prague::SE	18 Oct 2022 06:06	17 Nov 2022 10:44	90.18 %	0.02 %	9.79 %	0.01 %
Birmingham::EOS	18 Oct 2022 06:05	17 Nov 2022 10:26	87.70 %	0.06 %	12.23 %	0.01 %
Strasbourg_IRES::SE2	18 Oct 2022 06:04	17 Nov 2022 12:46	87.68 %	0.03 %	12.26 %	0.03 %
Catania::SE	18 Oct 2022 06:07	17 Nov 2022 10:23	86.12 %	0.03 %	13.84 %	0.00 %
KISTI_GSDC::SE2	18 Oct 2022 06:07	17 Nov 2022 10:41	86.03 %	0.17 %	13.80 %	0.00 %
LBL_HPCS::EOS	18 Oct 2022 06:04	17 Nov 2022 10:23	85.88 %	1.21 %	12.90 %	0.00 %
Poznan::SE	17 Oct 2022 23:50	17 Nov 2022 10:42	79.63 %	0.33 %	20.04 %	0.00 %
ISS::FILE	18 Oct 2022 06:07	17 Nov 2022 05:19	78.76 %	0.07 %	21.12 %	0.04 %
Kolkata::EOS2	18 Oct 2022 06:09	17 Nov 2022 12:09	68.71 %	0.61 %	30.57 %	0.12 %

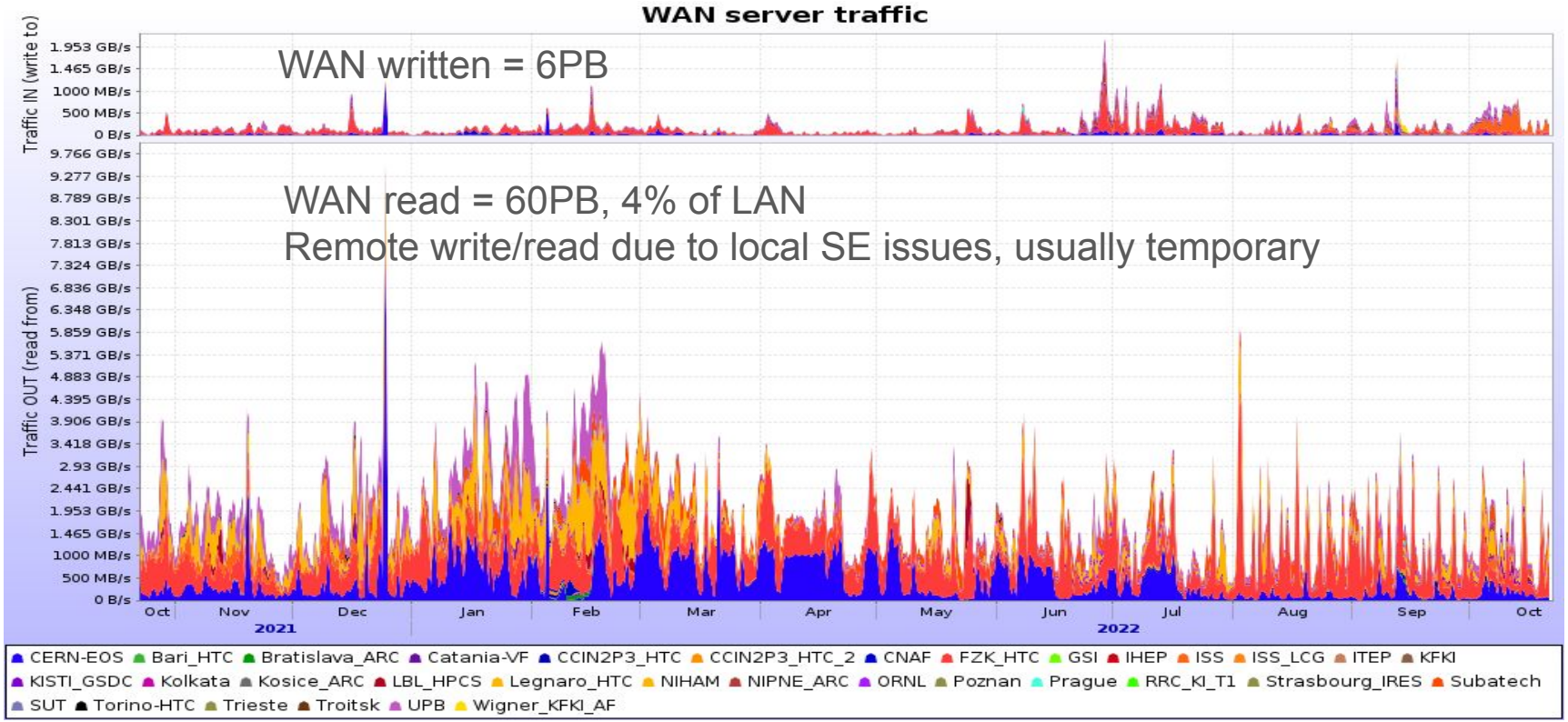
Network and data processing

- Jobs are dispatched to the Grid sites that already have the data
  - Minimizes WAN traffic and RTT efficiency penalty
- Grid site local file access (95%), remote (5%)
  - Remote access due to local SE issues, usually temporary
- Multiple replicas sorted topologically: apps first access local replica, then the next closest
  - Sorting by network topology, availability, network quality, geo-location and other metrics
- Storing multiple replicas
  - One replica is written to the local storage element
  - The other replicas are written to the remote (but close) storage elements
  - Remote writes might go through LHCOPN / LHCONE

Data access - LAN

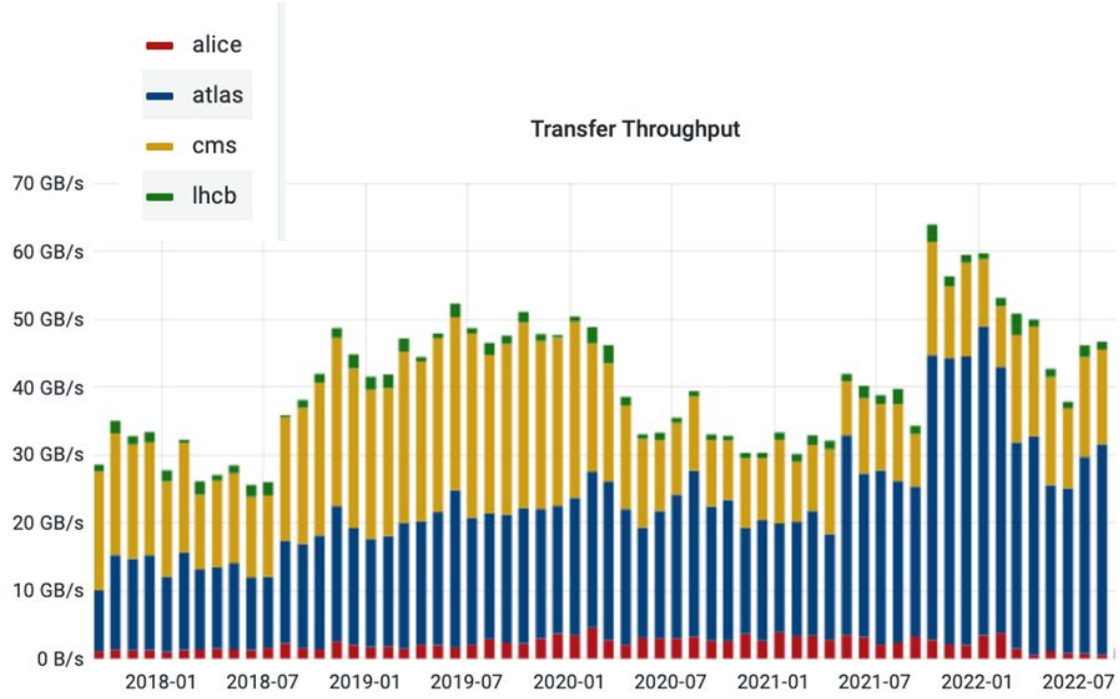


Data access - WAN (LHCONE/LHCOPN)



ALICE in the big picture - WLCG data transfers

- Includes RAW data distribution and other LHCONE/LHCOPN transfers
- ALICE computing model and network use is beneficial to remote sites
- Network requirements are mild and well within the capabilities of regional T2s

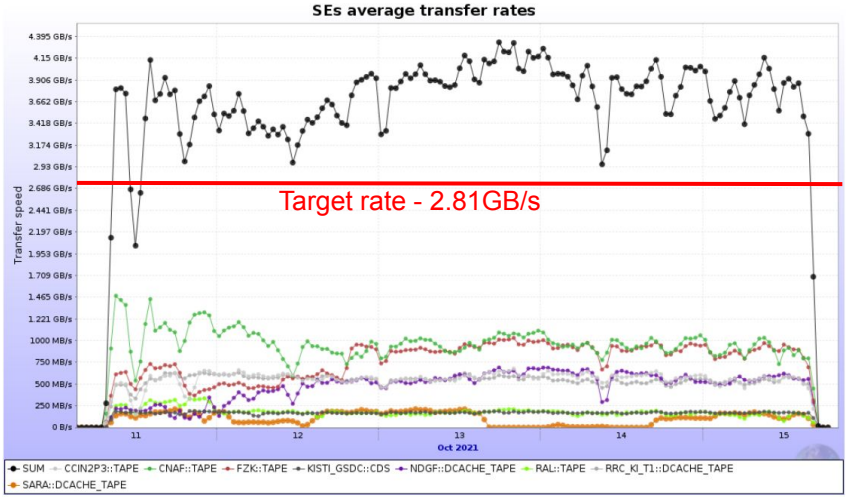


Expected data rates in Run3 - replication of RAW

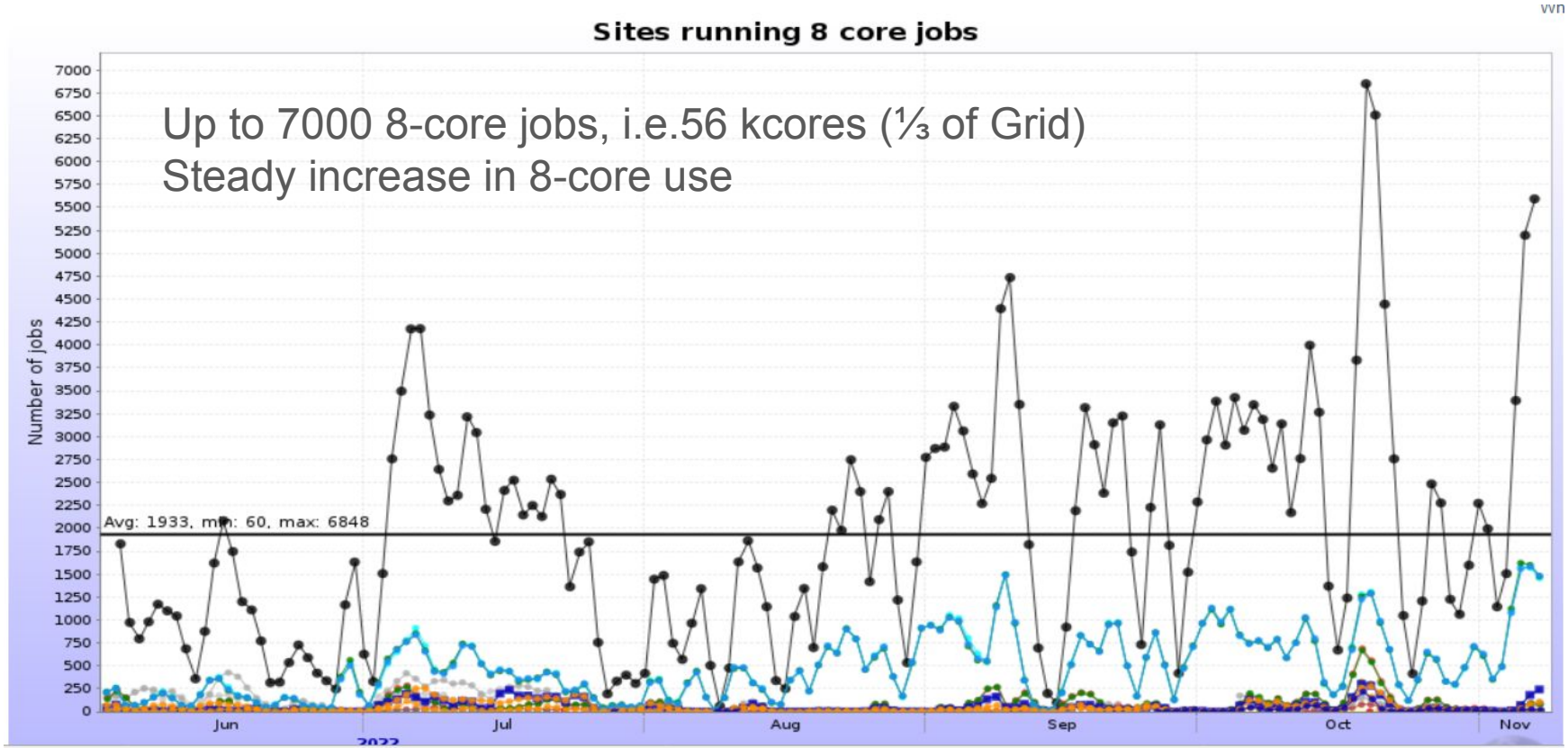
T1 Centre	Target rate GB/s	Achieved rate GB/s
CNAF	0.8	0.94 (116%)
IN2P3	0.4	0.54 (130%)
KISTI	0.15	0.16 (106%)
GridKA	0.6	0.76 (123%)
NDGF	0.3	0.47 (144%)
NL-T1	0.08	0.1 (122%)
RRC-KI	0.4	0.53 (128%)
RAL	0.08	0.17 (172%)

Sum 2.81GB/s

- Full traffic simulated during data challenge
- Channels tuned to slightly above the target rate, within reasonable limit
- The bulk of the bandwidth will be used after the Pb-Pb data taking period, for ~3 months
  - Since there is no Pb-Pb this year, we remain at the level of data challenges



Multicore and full node use



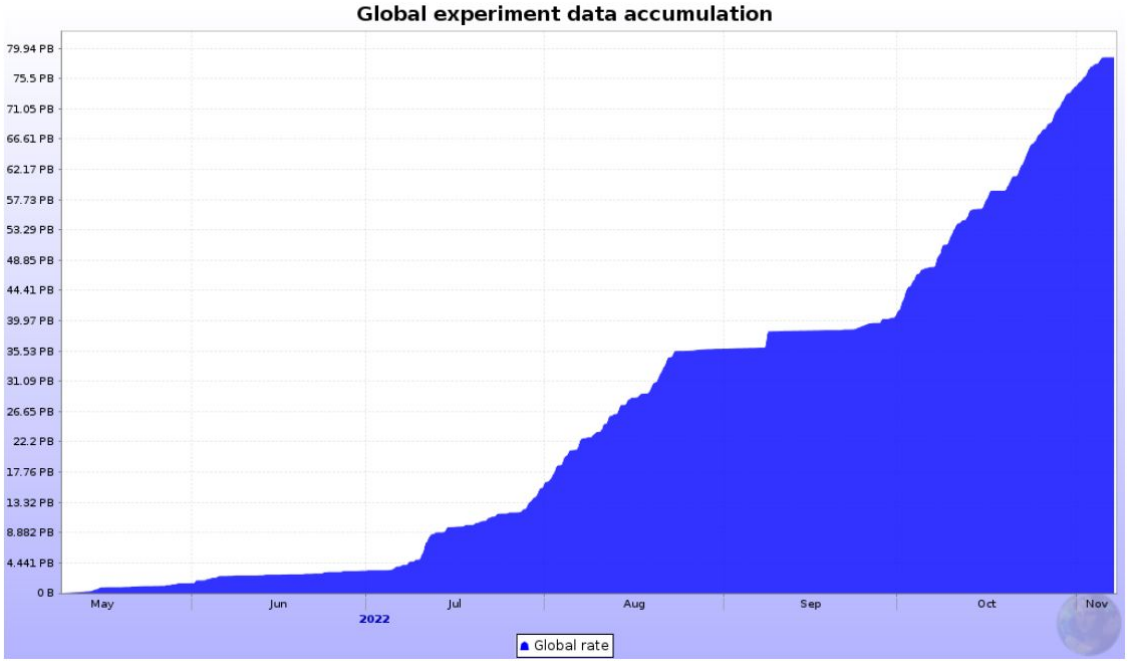
Site upgrade to 8-core

- Site conversion to 8 cores ongoing
  - ~60% of capacity already there
  - By January 2023 - 95% should be ready
- Good experience with whole node submission
  - Steady running (LBNL Lawrencium) + ORNL
  - All HPC resources are whole-node, use will expand
  - Possibility to improve job performance
    - ~8% reduction in execution time through optimal NUMA assignment
    - More flexibility with CPU vs. I/O intensive tasks

	AliEn proxy		LDAP		AliEn Tests JAliEn cert	
Service	Status	Time left	Status	Cores	Status	Time left
1. Altaria		1d 23:04		8		59d 22:58
2. Bari_HTC		1d 23:52		1		29d 22:55
3. Birmingham		1d 23:42		1		29d 23:01
4. Bratislava_ARC		1d 23:50		1		29d 22:59
5. Capella		1d 23:25		1		59d 23:05
6. Catania-VF		1d 23:59		1		29d 23:05
7. CCIN2P3_HTC		1d 23:57		8		29d 23:03
8. CCIN2P3_HTC_2		1d 23:58		8		29d 23:00
9. CERN-AURORA		1d 23:48		1		59d 23:01
10. CERN-CORONA		1d 23:43		8		59d 22:57
11. CERN-MIRAGE		1d 23:46		8		59d 22:56
12. CERN-SIRIUS		1d 23:39		8		59d 22:52
13. CERN-TRITON		1d 23:56		8		59d 23:05
14. CERN-ZENITH		1d 23:46		8		59d 22:59
15. Clermont_ARC		1d 23:20		1		29d 22:55
16. CNAF		1d 23:51		8		29d 23:02
17. CNAF-DUE		1d 23:17		8		29d 22:52
18. CORI	-	-		-	-	-
19. Cyfronet_ARC		1d 23:12		1		29d 22:56
20. DCSC_KU		29d 3:34		-	-	-
21. EPN	-	-		16		314d 23:53
22. FZK		1d 23:30		8		29d 22:56
23. FZK_HTC		1d 23:21		8		29d 23:05
24. Grenoble_ARC		1d 23:50		1		29d 23:06
25. GRIF_IPNO_IJCLAB		1d 23:20		1		17d 6:38
26. GRIF_IRFU		1d 23:15		1		29d 22:54
27. GSI_8core	-	-		8		68d 22:12
28. HIP	-	-		1		162d 2:57
29. Hiroshima		1d 23:25		8		29d 23:05
30. HPCS_Lr	-	-		0		222d 3:35
31. IHEP		1d 23:07		1		29d 22:59
32. ISS		1d 23:30		1		335d 15:39
33. ITEP		1d 23:27		1		29d 23:05
34. JINR_ARC		1d 23:30		1		29d 22:57
35. KFKI		1d 23:08		1		29d 23:03
36. KISTI_GSDC		1d 22:59		1		29d 22:55
37. Kolkata		1d 23:10		1		29d 23:02
38. Kosice_ARC		1d 23:59		1		29d 23:05
39. LBL_HPCS	-	-		1		112d 7:50
40. Legnaro_HTC		1d 23:07		1		29d 23:02
41. Minerva		1d 23:39		1		59d 22:54
42. Nemesis		1d 23:46		8		93d 5:19
43. NIHAM	-	-		1		95d 20:22
44. NIKHEF		1d 23:20		1		29d 23:07
45. NIPNE_ARC		1d 23:26		8		29d 22:54
46. ORNL	-	-		0		112d 7:37
47. Oxford		1d 23:07		1		29d 23:02
48. Phoenix		1d 23:39		1		93d 5:24

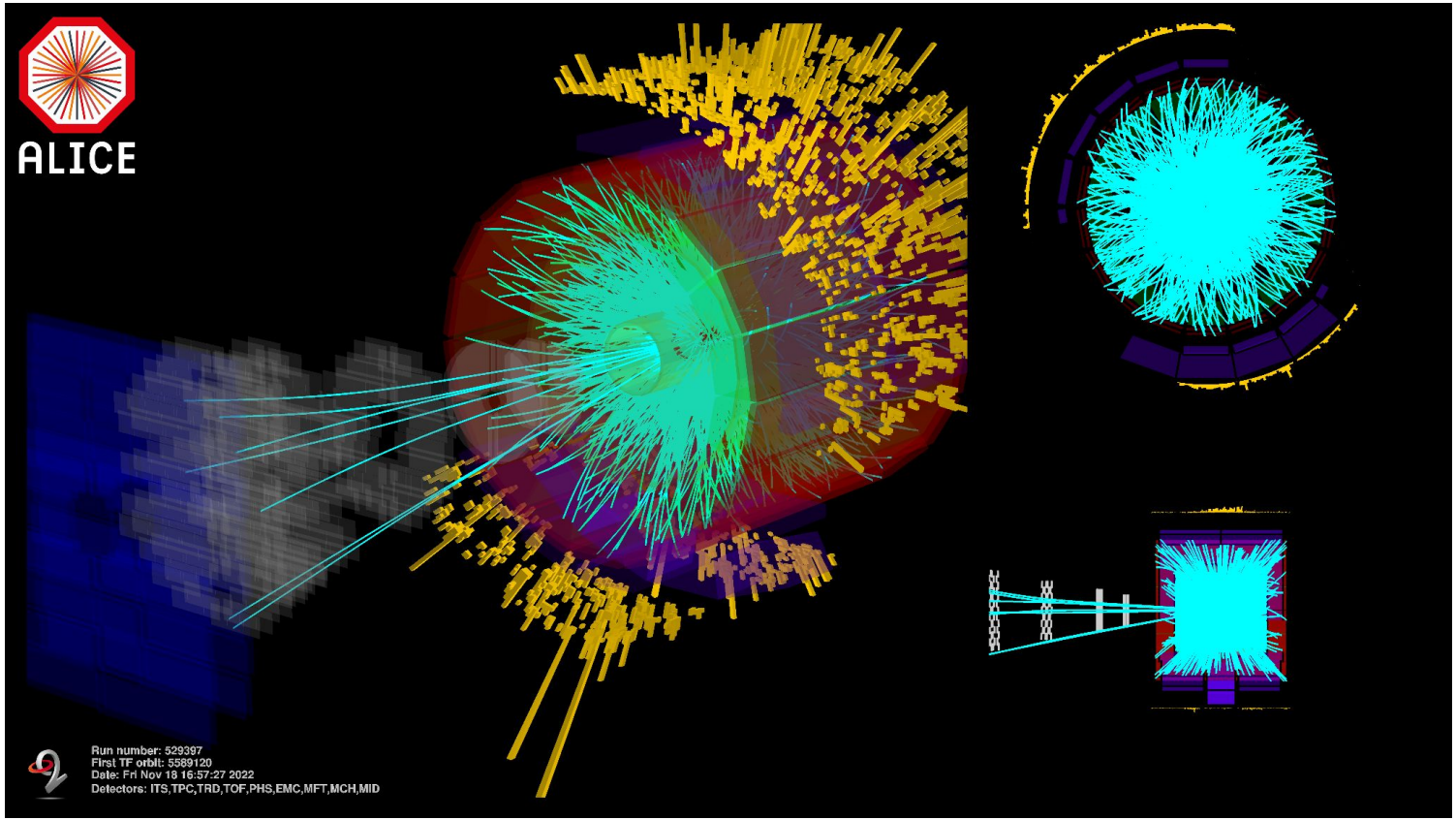
2022 data collection and processing

- ALICE is taking p-p data since July
- Record-breaking data volume - so far 80PB, all stored on O2 EOS buffer
- January - massive processing, followed by skimming
- Exciting times for everyone involved
- 2023 - first year with Pb-Pb beam (looking forward to it)



ALICE

Event from 18 Nov 2022, low IR Pb-Pb@5.36TeV



Summary

- After a long (COVID-interlaced) pause, the LHC commenced its Run3 in the spring of 2022
  - ALICE collected record amount of p-p data with upgraded detector, new online, offline and Grid software JAliEn
- The Grid sites are updated and continue to be the backbone of the ALICE data storage and processing
  - Still to do 8-core/whole node conversion
  - Number of ongoing projects to make it better and include new resources - ***participation welcome!***
- The processing strategy continues to depend on good network connectivity for data exchange
  - The AFs behaviour is still to be tested with real analysis scenarios
- ALICE computing requirements will increase and we count on the Asian sites to continue their growth and involvement