

Grid production with the ATLAS Event Service

CHEP2018 – Computing in High Energy and Nuclear Physics 2018, 9-13 July 2018, Sofia, Bulgaria

Funded by MICINN under contracts FPA2016-75141-C2-1-2-R and FPA2016-80994-C2-2-R

D. Benjamin¹, P. Calafiura², T. Childers³, K. De⁴, E. Fullana Torregrosa^{5*},
 P. Van Gemmeren³, A. Di Girolamo⁶, W. Guan⁷, M. Lassnig⁶, T. Maeno⁸,
 P. Nilsson⁸, V. Tsulaia², T. Wenaus⁸, W. Yang⁹ on behalf of the ATLAS Collaboration
¹Duke University (US); ²Lawrence Berkeley National Lab. (US); ³Argonne National Lab. (US);
⁴University of Texas at Arlington (US); ⁵Instituto de Fisica Corpuscular (IFIC), UV-CSIC, (Spain);
⁶CERN (Switzerland); ⁷University of Wisconsin (US); ⁸Brookhaven National Lab. (US); ⁹SLAC National Accelerator Lab.
 *Corresponding author

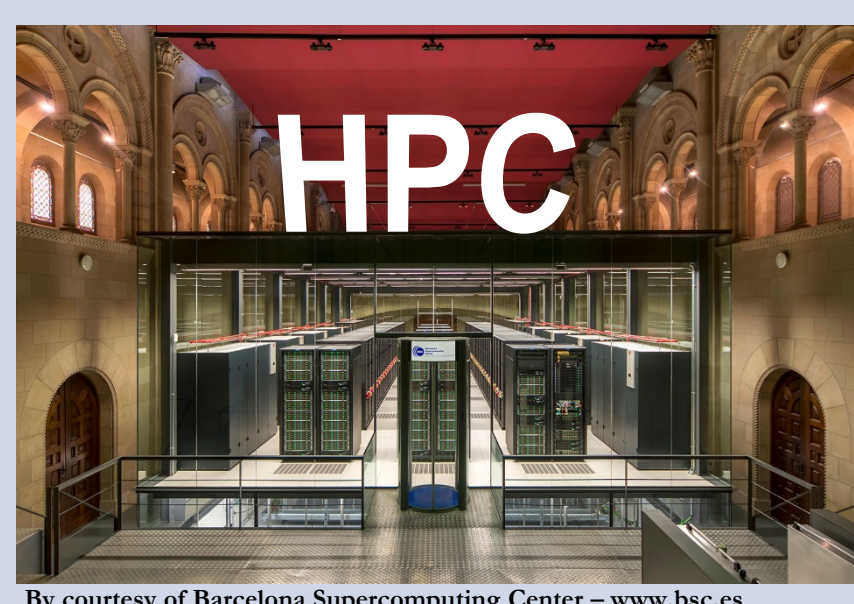


Why ATLAS Event Service ?

Efficiently and flexibly exploit any CPUs available:

Opportunistic computing

CLOUD COMPUTING



VOLUNTEER COMPUTING

...but also better use of all the resources :

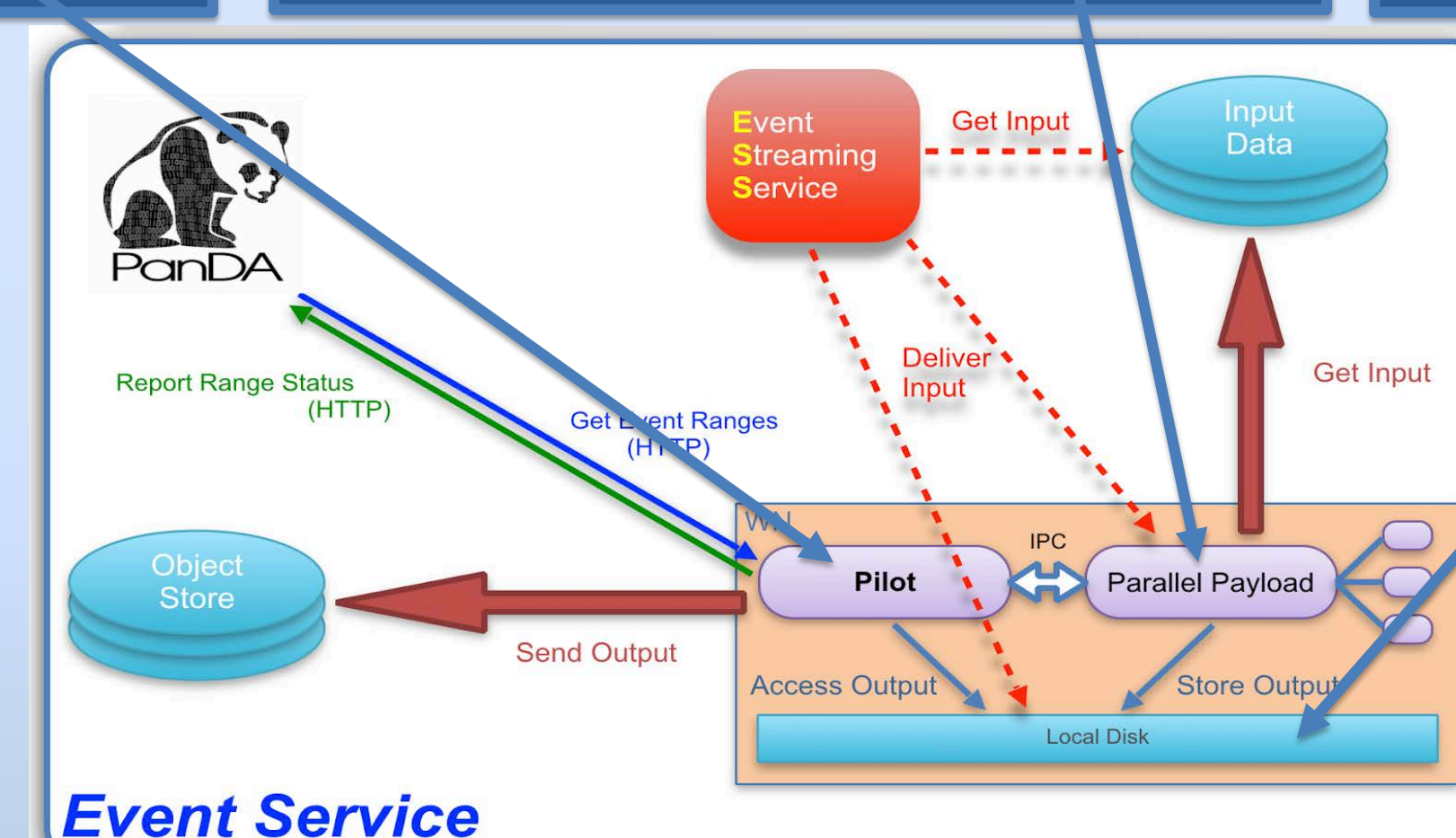
- Full utilization of cycles
- A job can be split into event granular subjobs to be processed in multiple resources
- Multiple jobs can be created and scheduled to different resources for processing the same event. The first available processes the event.

What is ATLAS Event Service ?

→ Pilot delivers fine-grained workload to running payload application in real time
 • Workload: Event Ranges
 • Event Range: identifier string with Input File ID, event positions within the file, etc.

→ Payload application: AthenaMP(*) with a special configuration for the Event Service. Payload directly reads input files for the event data. Either local or remote inputs. The payload makes a sequence of fine-grained output files. **One output per input event range.**

→ Event Service outputs are first stored into **Object Stores** (or to normal storages)
 Final outputs are produced later by dedicated merge jobs

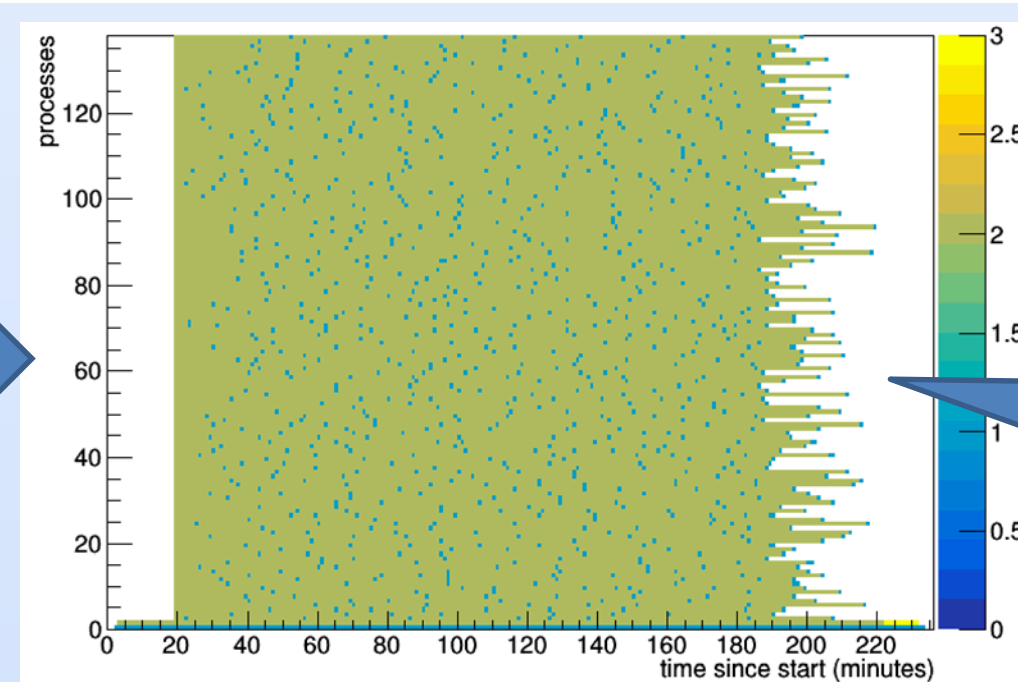


(*) Athena is the name of the ATLAS software framework that manages almost all ATLAS production workflows. AthenaMP adds multiprocessing functionality to Athena.

Use case 1 : Opportunistic resources : HPCs

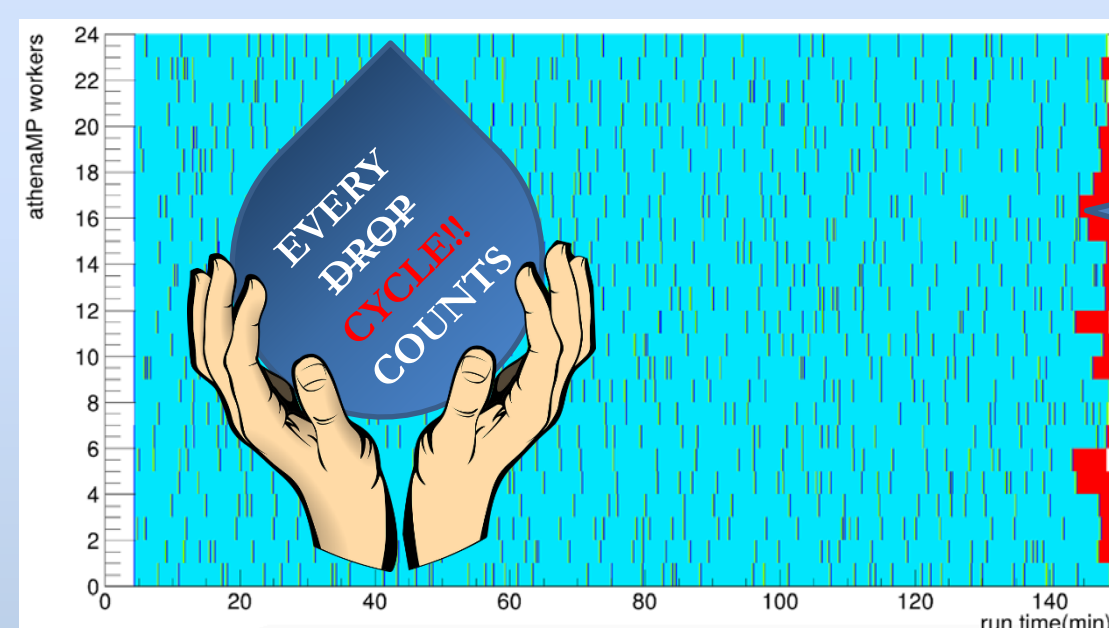
NERSC utilization per core (1 node)

without Event Service



Parallel processing leaves wasted cycles due to uneven event processing times.

with Event Service

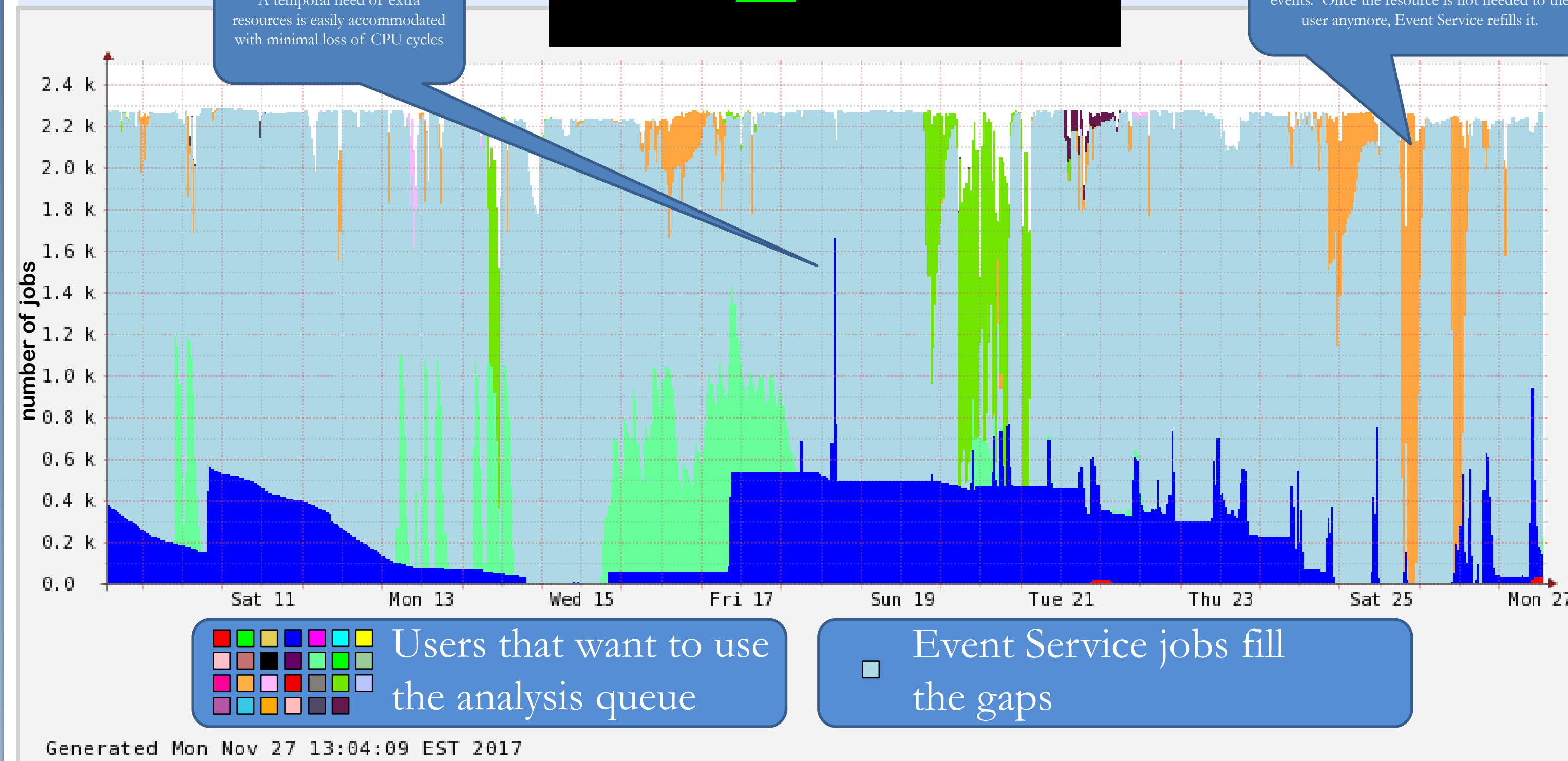


Event Service fills the holes with events
 Only the last interrupted event is lost (in red)

Use case 2 : Opportunistic resources : Tier 3



BNL_LOCAL



A temporal need of extra resources is easily accommodated with minimal loss of CPU cycles

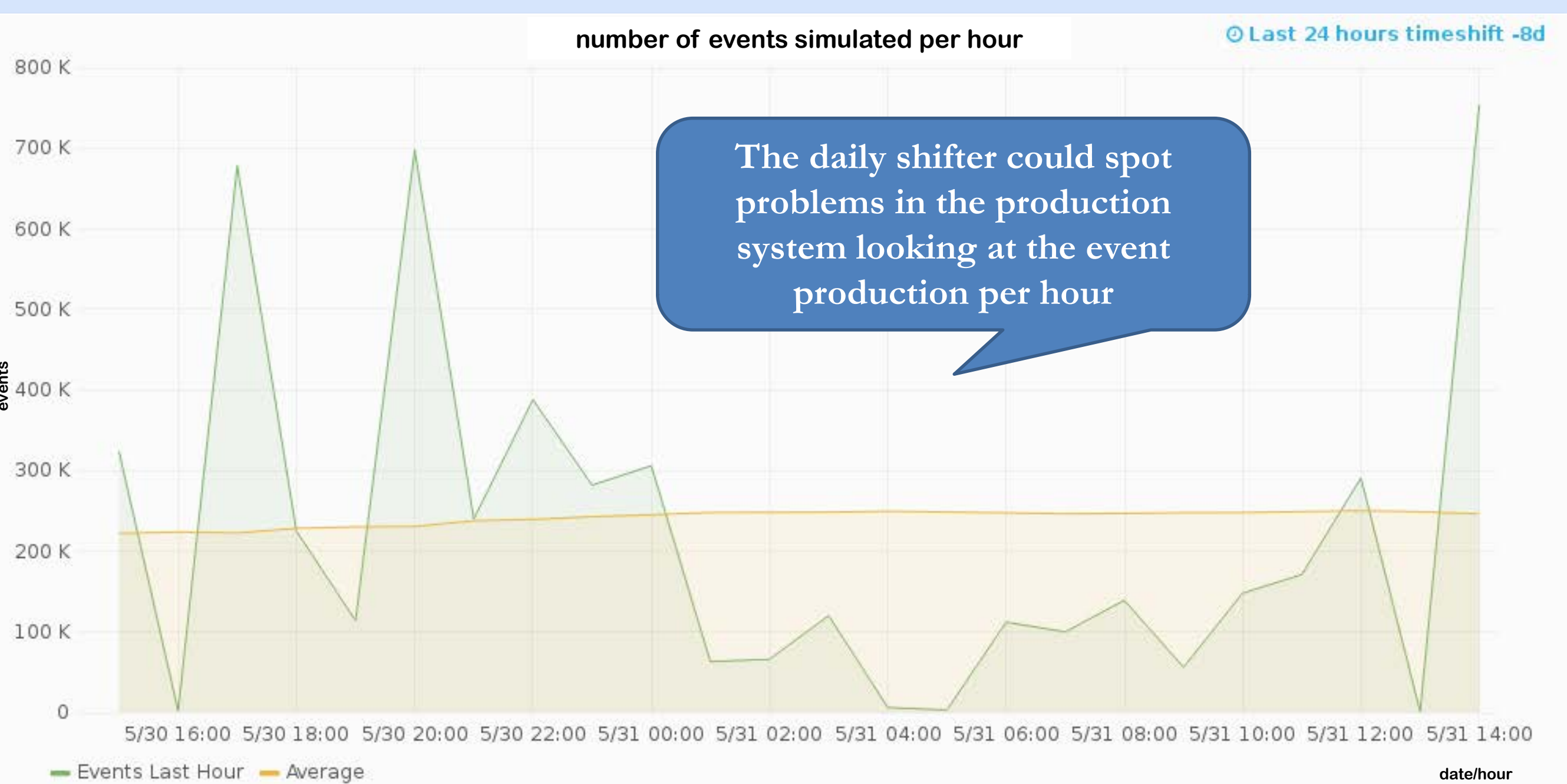
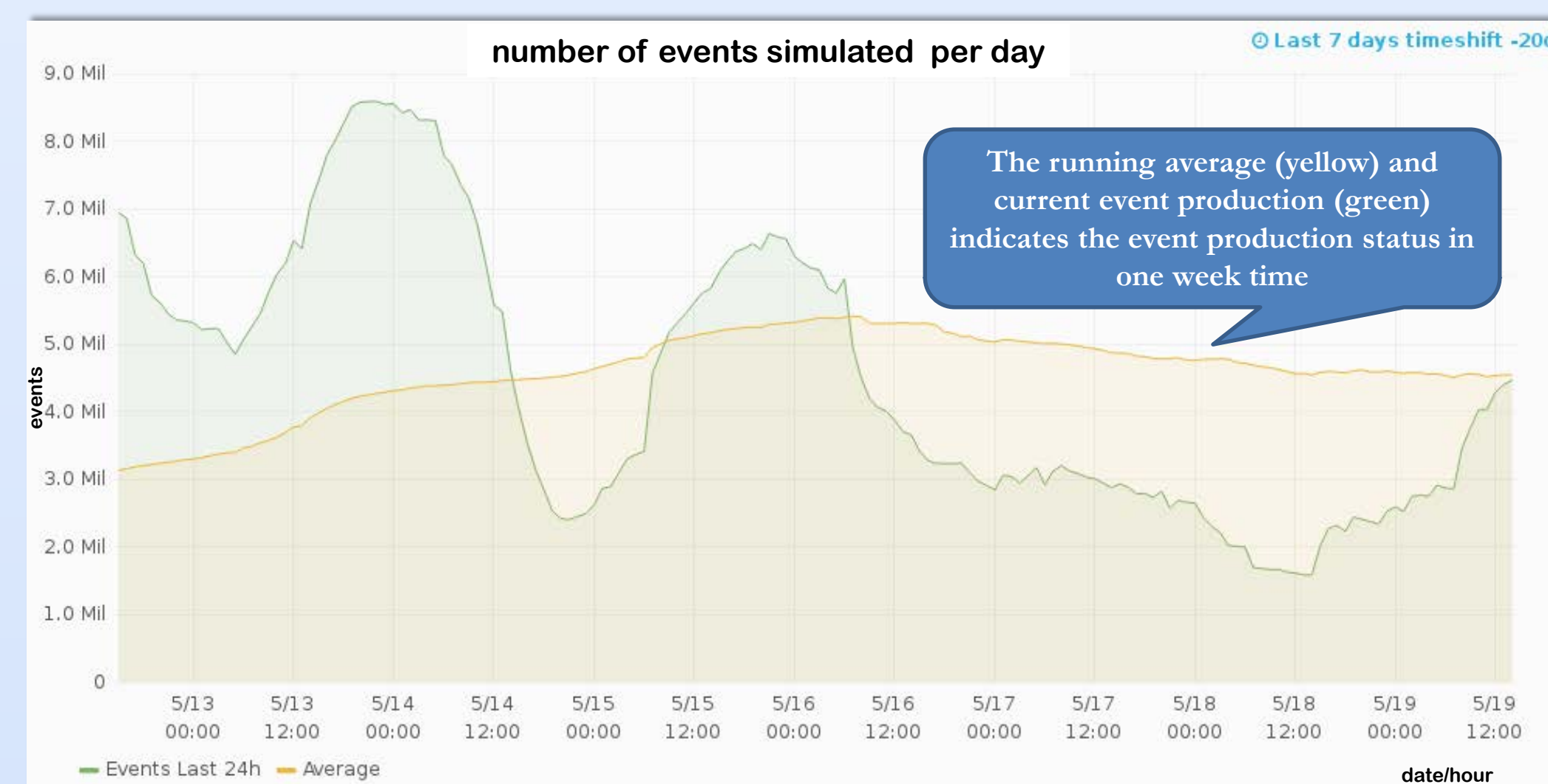
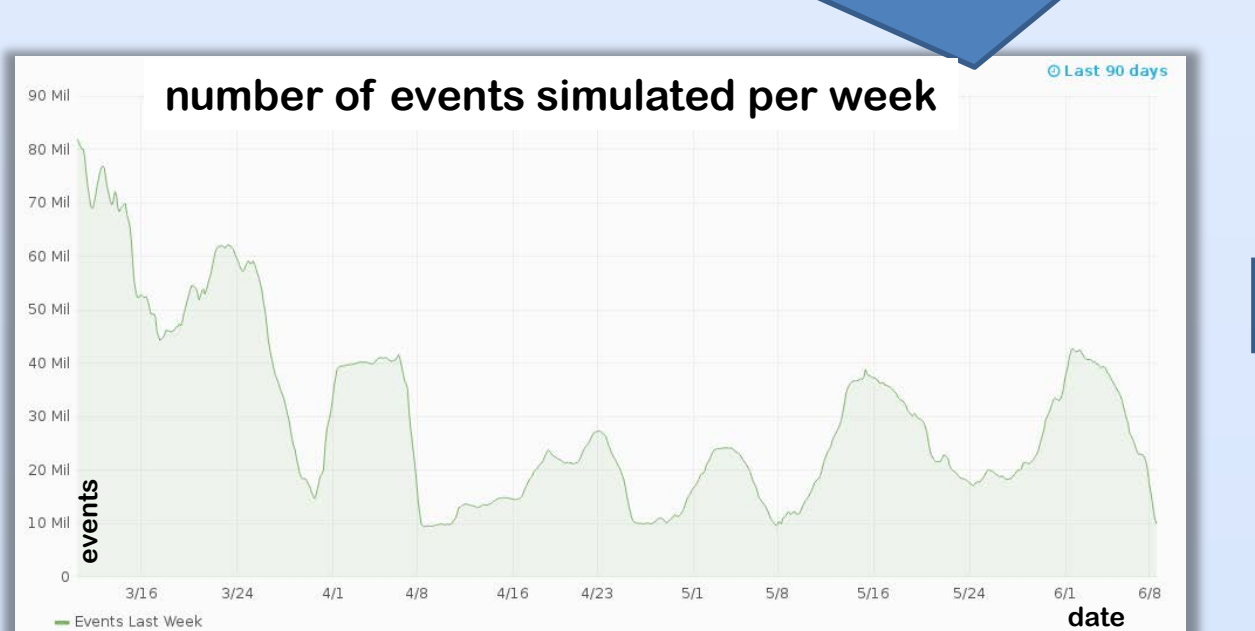
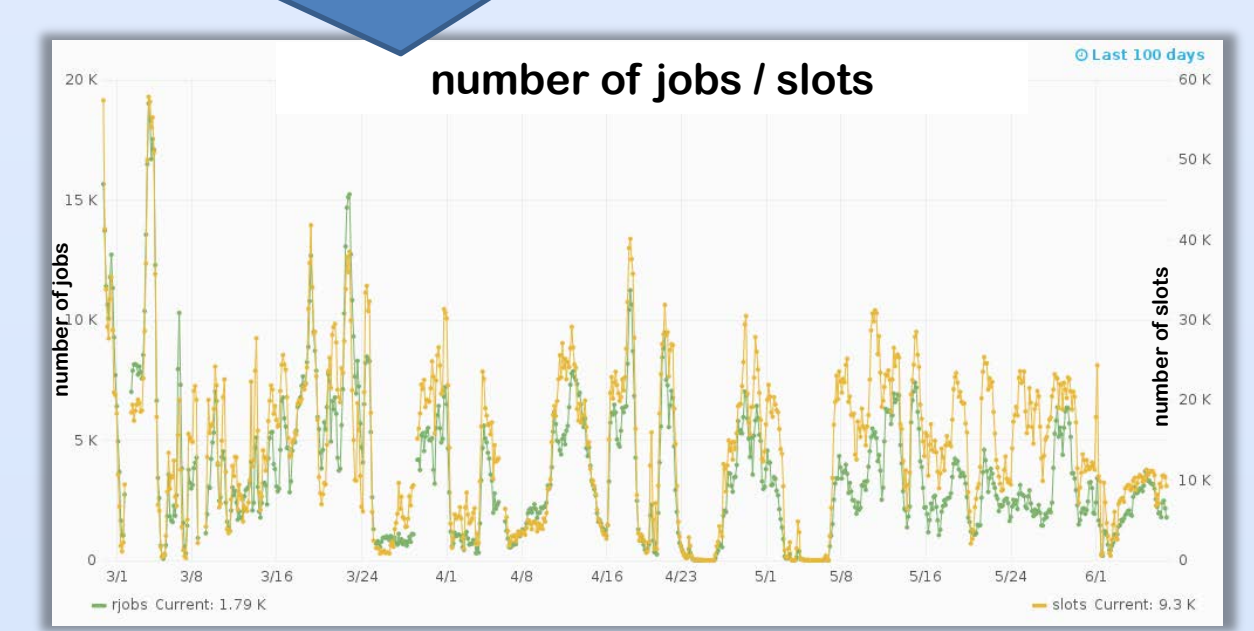
When the resource is fully needed by a user, Event Service allows for a complete termination of the jobs with no loss of the processed events. Once the resource is not needed to the user anymore, Event Service refills it.

Change of paradigm → New monitoring tools

Starting last year ATLAS simulation jobs were automatically assigned to be run under the new Event Service framework. Since then many brand new monitoring tools have been developed to understand the performance and status of the simulation

As a single job can produce an unknown number of events the numbers of jobs (or slots = jobs * cores) is less significant

The monitoring of the rate of number of events produced gives a more realistic picture of the status of the production system



Conclusions

- The new Event Service framework has been commissioned in the ATLAS grid production system
 - After more than 3 years of R&D
- Using opportunistic platforms have proven to be very efficient
- Sometimes the combined deployment of a brand new computing framework (Event Service) with a brand new storage architecture (Object Store) introduced challenging issues...
 - ...but resulted in improvements in both
- The change in paradigm from files to individual events made us rethink the monitoring quantities and to develop a brand new monitoring system.

