

Random thoughts on Accounting

...and the long-standing issue with the average HS value

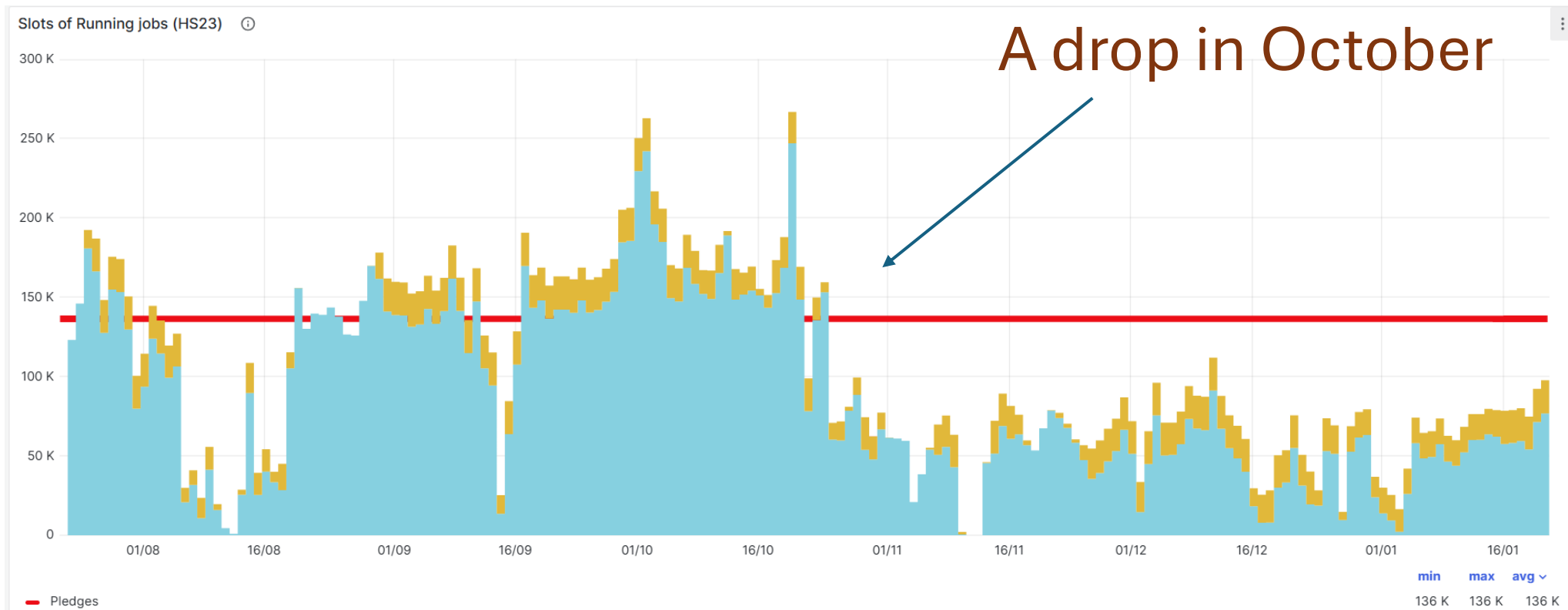
ATLAS Distributed Computing Technical Interchange Meeting
January 21, 2025

David Rebatto <David.Rebatto@mi.infn.it>

Disclaimer

- The subject came up only a few days ago, so this presentation lacks a few important things on my side:
 - A thorough investigation into the current accounting machinery
 - A survey among sites about the most common problems
 - An investigation of existing work or plans on accounting
- As a result, this presentation is primarily a collection of personal reflections on a recurring issue reported during meetings. It may contain gaps or inaccuracies due to my limited understanding of the subject.

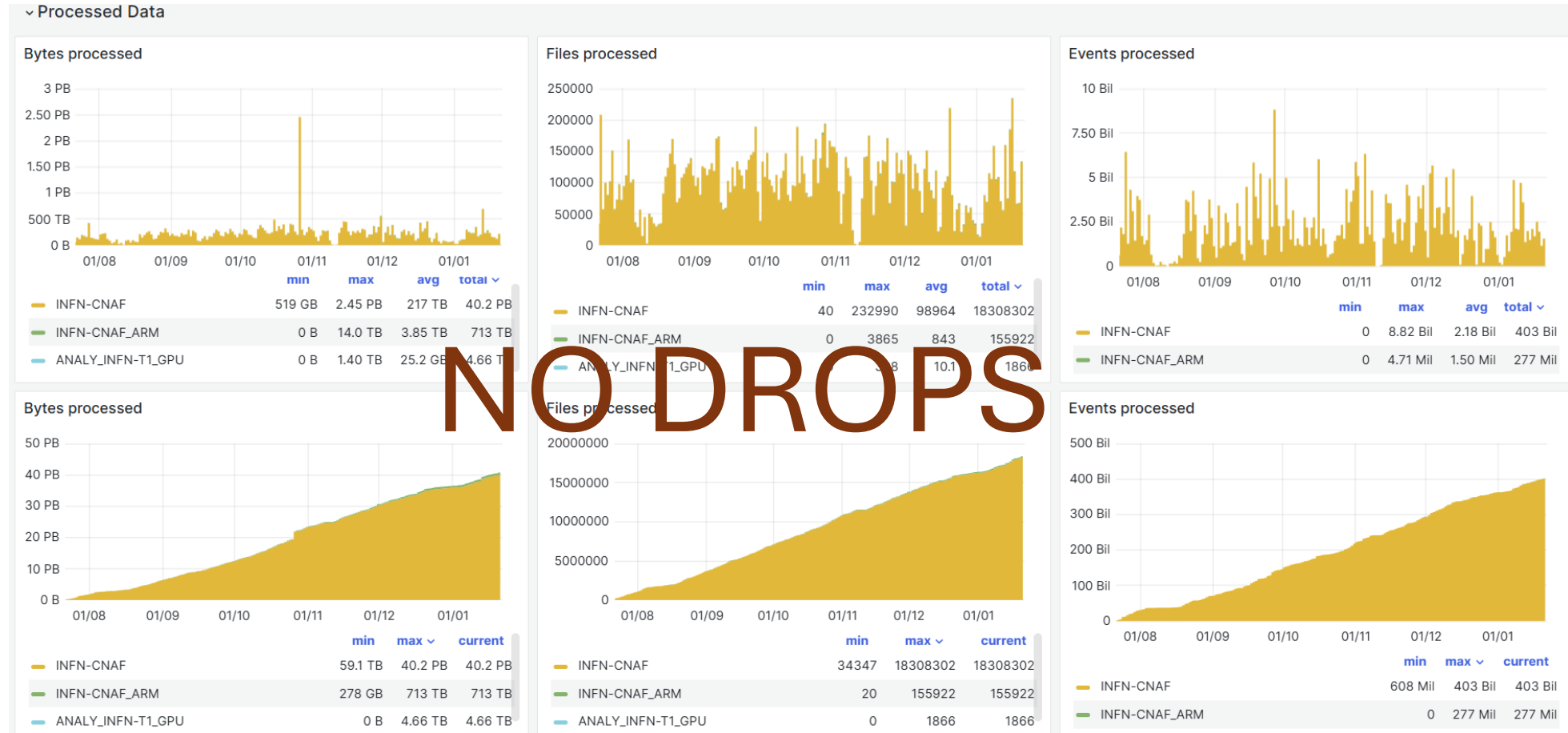
Slots of running jobs at T1 (last 6 months)



From ATLAS Job Accounting monitoring

David Rebatto <David.Rebatto@mi.infn.it>

Work done (last 6 months)



From ATLAS Job Accounting monitoring

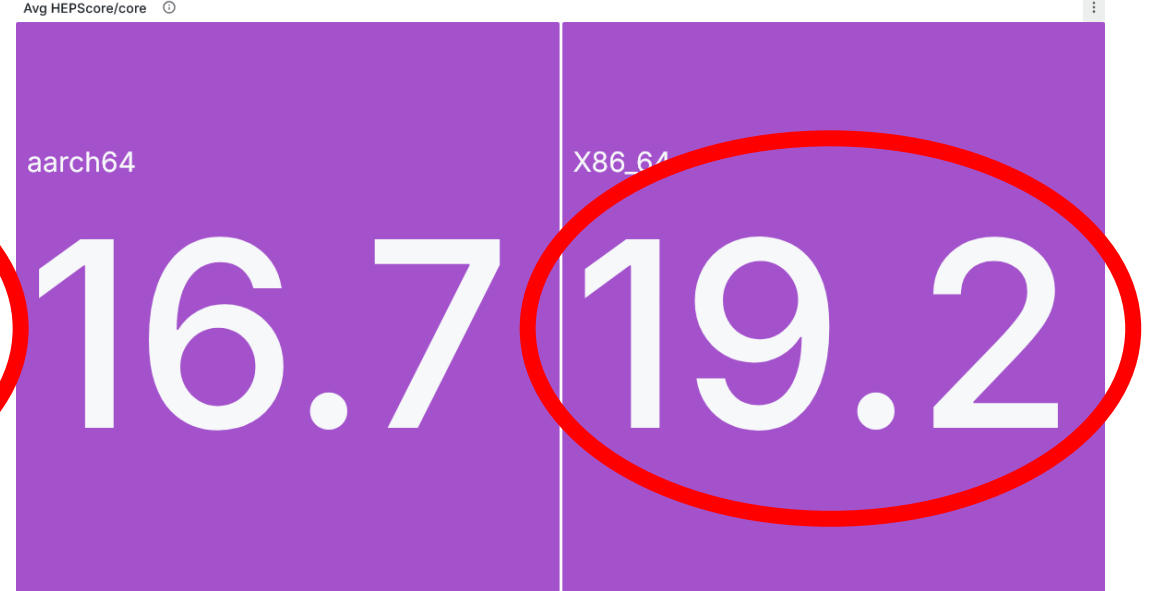
David Rebatto <David.Rebatto@mi.infn.it>

Possible explanation

HS23 on October 1st



HS23 on November 1st



From internal CNAF monitoring

Issues with scaling by average HS23

- **Error-prone:** Administrators sometimes forget to update the status when setting nodes offline, bringing them back online, installing new nodes, etc.
- **Challenging to update in CRIC:** Special permissions are required, making updates somewhat difficult.
- **Limited administrative control:** At some sites (e.g., CNAF), outsourced computing clusters (e.g., Cineca, Leonardo) might not be under direct administrative control, causing downtimes to be reflected inconsistently in the average value.
- **Timing mismatch:** The information is decoupled from when the accounted job was actually executed.
- **Uneven job distribution:** On sites with diverse node types and multiple VOs, jobs may not be evenly distributed across all types of nodes.

Issues with scaling by average HS23

- **Error-prone:** Administrators sometimes forget to update the status when setting nodes offline, bringing them back online, installing new nodes, etc.
- **Challenging to update in CRIC:** Special permissions are required, making updates somewhat difficult.
- **Limited administrative control:** At some sites (e.g., CNAF), outsourced computing clusters (e.g., Cineca, Leonardo) might not be under direct administrative control, causing downtimes to be reflected inconsistently in the average value.
- **Timing mismatch:** The information is decoupled from when the accounted job was actually executed.
- **Uneven job distribution:** On sites with diverse node types and multiple VOs, jobs may not be evenly distributed across all types of nodes.

CRIC API?



An idea towards job-grained accounting

- Store in Panda DB, along with CPU time and execution WN, the HS23 value of the WN itself
- Sites must make the information available
 - With HTCondor, in the 'startd' classad
 - In a well-know file (e.g. in /etc)?
 - Using a lookup table?
- The job pilot collects the value and passes it up to Panda

An idea towards job-grained accounting

- Store in Panda DB, along with CPU time and execution WN, the HS23 value of the WN itself
- Sites must make the information available
 - With HTCondor, in the 'startd' classad
 - In a well-know file (e.g. in /etc)?
 - Using a lookup table?
- The job pilot collects the value and passes it up to Panda
- **The solution seems too simple. What are the challenges that make it difficult to implement?**

Thank you.