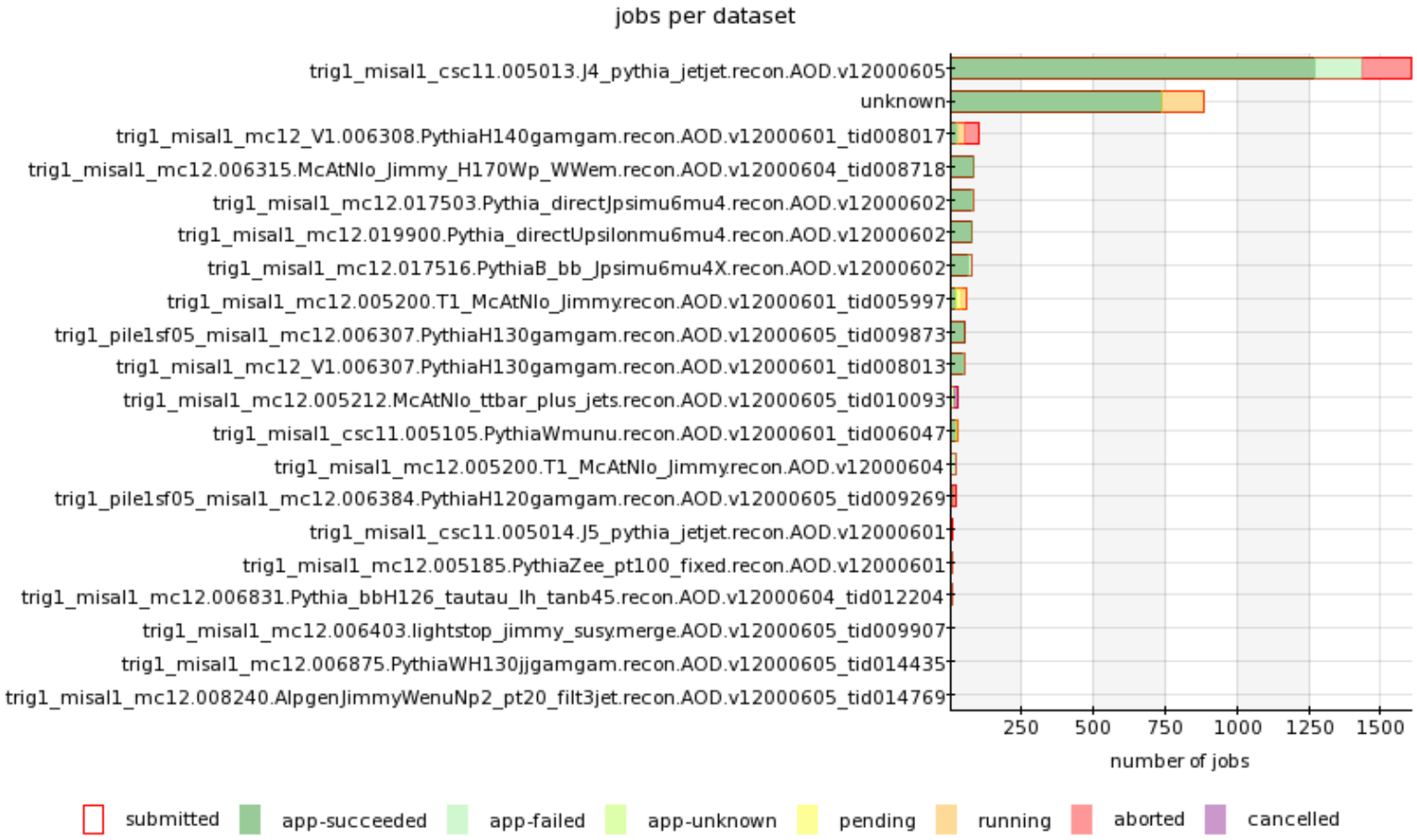
Access to Dataset usage information

Dietrich Liko

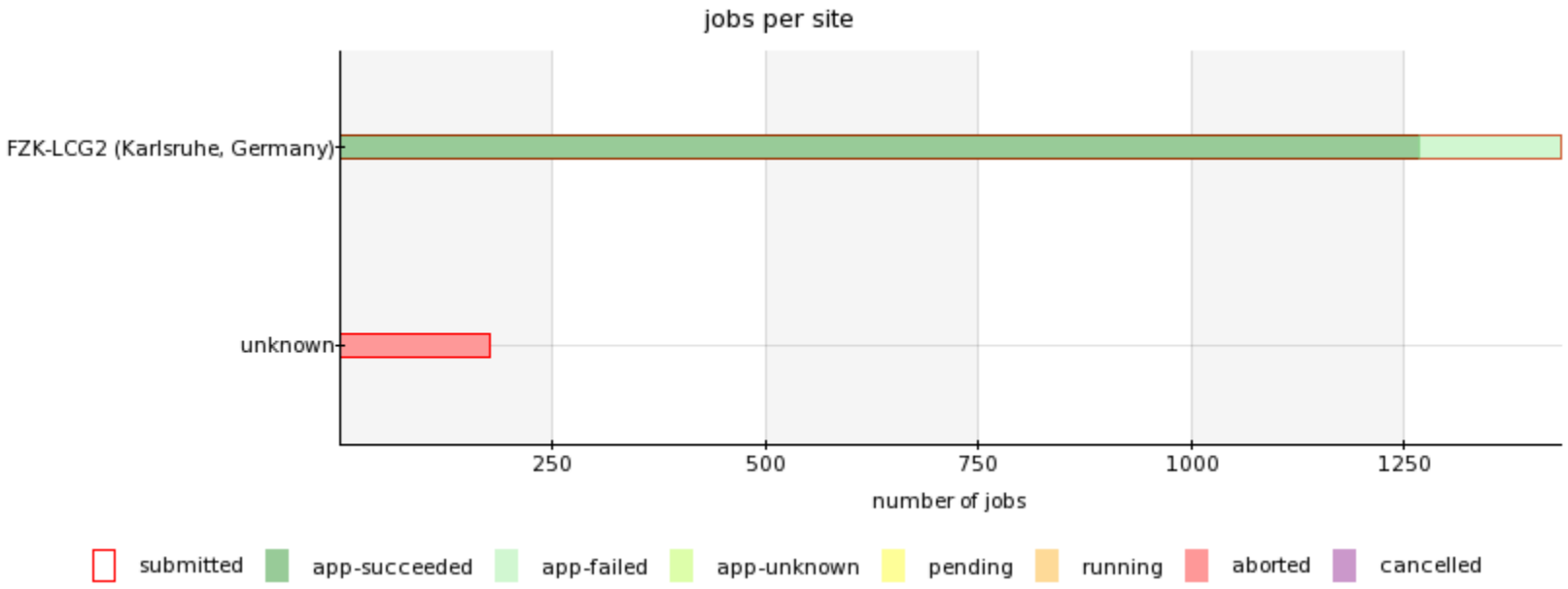
Motivation

- ATLAS is replicating dataset to many Tier-2
  - Some datasets are seem never to be used
  - We need feedback to optimize our dataset distribution strategy
- Information already available via Dashbord
  - A summary page/table is needed
  - Access by some tool to be used in a script

Just the last hours



Site info


jobs per site
FZK-LCG2 (Karlsruhe, Germany)
unknown
250
500
750
1000
1250
number of jobs
submitted
app-succeeded
app-failed
app-unknown
pending
running
aborted
cancelled

Other important elements

- Access to PANDA database for info on PANDA based jobs
  - Central taskqueue, is already a database
  - Seems a better way then instrumentation
- Access to jobs submitted to Nordugrid
  - Instrumentation of the job a la EGEE