

# Tier-0 Feedback

---

Armin NAIRZ

(CERN, PH-ADP-CO Group)

# Contents

---

- Tier-0 feedback to Autumn-2010 reprocessing campaign
  - TAGs and TAG uploading
- Tier-0 resources and capacity
  - CPU (LSF farm)
  - Disk buffer, bandwidth (CASTOR)
- Bulk processing at Tier-0 and Tier-1s

# Feedback to Autumn-2010 Reprocessing

---

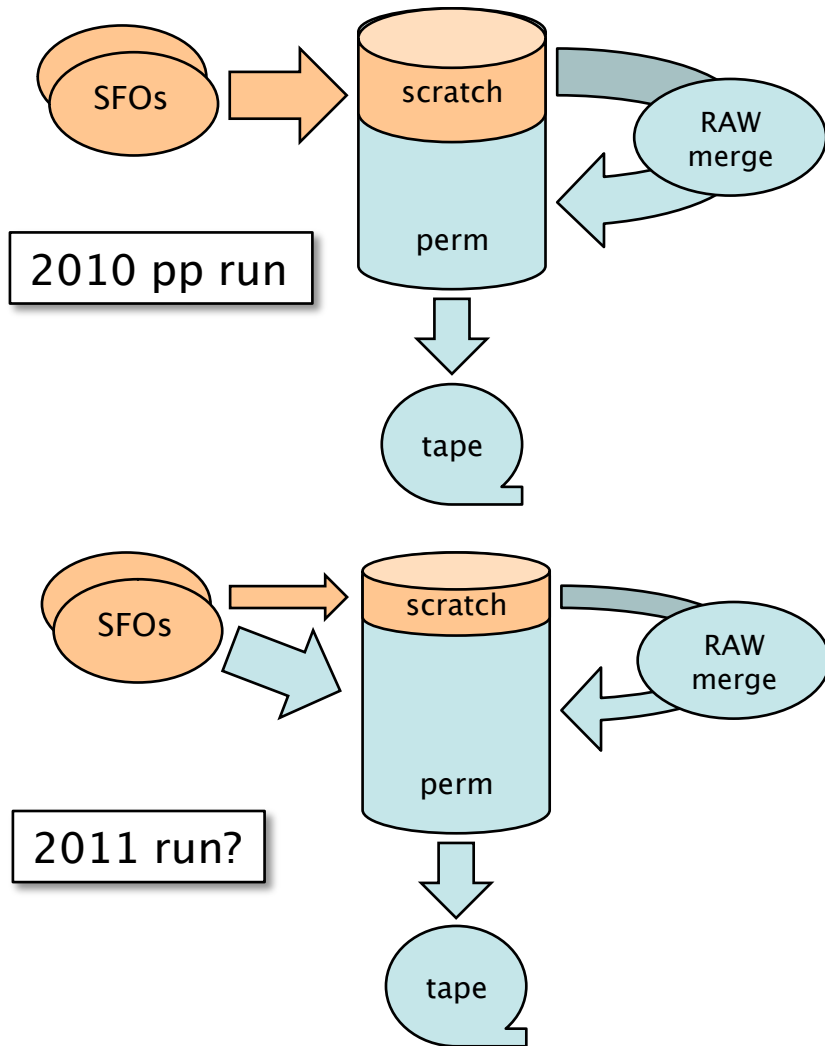
- Tier-0 involvement in reprocessing: TAGs
  - Shipped to CERN/DATADISK
  - Picked up in cron jobs (E. Vinek), using AMI and DDM info, inserted into Tier-0 databases
  - Complicated chain of pre-, upload- and post-processing steps to several TAG databases, run by Tier-0
- TAGs had arrived at CERN, we started uploading and processing, then learned “by chance” that TAGs were broken
  - After one week of processing
- “Awareness” for TAGs and their processing at the Tier-0 has to be raised
  - Includes also TAG validation (physics groups?) before a reprocessing campaign is launched

# Tier-0 Resources and Capacity (I)

---

- Scale of Tier-0 resources according to Computing Model + margins
  - 200 Hz EF output rate, 60 HS06-s (= 7-8 s wall-time) reco time/event
  - Extra resources for Tier-0/CAF activities
- What we got was much more
  - Up to 600 Hz EF output rate, 15-20 s reco time/event (wall-time!)
    - RAW→ESD, ESD→AOD, ESD→DESD, DQ monitoring in one job
- CPU resources (LSF batch farm)
  - ~3400 cores (= max number of parallel jobs)
    - ~200 dedicated to CAF TMS, max 3000 reco jobs, rest needed for merging
    - Max achievable reco rate (@ 15-20 s per event): 150-200 Hz
  - For likely data-taking scenarios of next year:
    - 400 Hz @ 50% LHC duty cycle: reco time < 15 s
    - 400 Hz @ 70% LHC duty cycle: reco time < 10 s

# Tier-0 Resources and Capacity (II)



- CASTOR disk buffer (T0ATLAS pool only)
  - ~630 TB, 65 servers
  - CM requirement: ~10 days' buffer for RAW data
  - For likely data-taking scenarios of next year:
    - Assumptions:
      - 1.5 MB/event RAW, 1 MB/event ESD
      - 90% T0ATLAS occupancy
      - 80% of RAW don't have to be merged
    - 400 Hz @ 50% LHC duty cycle:
      - ~8 days' buffer (full RAW merging)
      - ~12 days' buffer (partial RAW merging)
    - 400 Hz @ 70% LHC duty cycle:
      - ~6 days' buffer (full RAW merging)
      - ~8 days' buffer (partial RAW merging)

# Tier-0 Resources and Capacity (III)

---

- CASTOR bandwidth (T0ATLAS pool only)
  - Network allows max 4 GB/s concurrent reading/writing
    - Operationally smooth up to ~3 GB/s
  - For likely data-taking scenarios of next year:
    - Assumptions:
      - 1.5 MB/event RAW, 1 MB/event ESD
      - 80% of RAW don't have to be merged
      - Only Tier-0 activity and tape migration, export not taken into account
    - 400 Hz @ 50% LHC duty cycle:
      - >800 MB/s writing, >1100 MB/s reading (full RAW merging)
      - Save ~240 MB/s writing/reading each for partial RAW merging
    - 400 Hz @ 70% LHC duty cycle:
      - >1100 MB/s writing, >1500 MB/s reading (full RAW merging)
      - Save ~340 MB/s writing/reading each for partial RAW merging

# Bulk Processing at Tier-0 and Tier-1s

---

- Bulk processing at Tier-0 is fully automated
  - From task definition, via processing, to registration, TAG upload, etc.
- S/w configuration is defined by PROC via AMI tags and read from AMI
  - Sufficient to configure Tier-0 reconstruction
- Data are released for processing automatically, after calibration loop
  - Regular look-up in COOL
- HI bulk processing on the Tier-1s needs manual task definitions, run by run, by PROC
  - Automation in case of 2011 pp bulk processing (?)
- TAGs from potential Tier-1 bulk processing would have to be treated like the ones from reprocessing
  - Shipped to CERN, loaded to TAG databases
- Caveat: mixture of f/m and r/p/t AMI configuration tags
  - Issue for applications (e.g. ELSSI), confusing for the users?