

Data Preparation and Tier0

Walter Lampl
University of Arizona

Introduction

- Didn't quite know what to say that isn't covered in other talks
 - Random collection of thoughts
 - Some repetition from last week's Data Preparation Post Mortem meeting

2010 pp prompt processing experience

- Very smooth running
 - I (PROC) was never called a night
 - Few weekend calls
 - We almost never used our emergency handles
 - Switching off failing algorithms, in-situ patching, etc
 - Switched off only monitoring tools (rare)
 - (Heavy Ion processing is a different story)

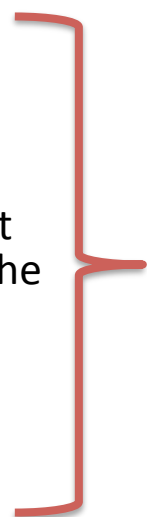
Frozen Tier0 policy

- Successful policy!
 - Allowed ATLAS to show results with maximum luminosity at conferences
 - I am not aware of any results-changing tag that slipped through
- Checks based on TCT: compareTCTs.py
 - Runs automatically every night to compare against rel_(n-1).
 - Can easily re-run by hand to compare any TCT against any other
 - Example: 15.6.X,rel_2 against 15.6.X.Y-VAL,rel_3
 - Is frequently needed because rel_(n-1) is not good for whatever reason
- Experience
 - There were a few occasions of inadvertent ESD/TAG changes that were caught by our checks
 - Increases our reliance on TCT working
 - Danger that we “lose the reference” if too many TCT runs fail or other problems accumulate
 - Have now a script to backup the relevant files of a particular TCT. Requires about 12 Gbytes disk space

Conditions & Calibration Loop

- Conditions are a potential loophole for the frozen-Tier0 policy
 - Ideally calibration updates equalize the detector response over time by compensating hardware drifts
 - We have absolutely no safety net in place to ensure that updates don't do more than that or even completely screw up the reconstruction
 - Rely on subsystem experts being careful
- The calibration loop infrastructure (NEMO, etc) work well
 - Only glitch was once a not-updated website
- The 36 hour delay seem to be suitable
 - Attempts to save time by manual sign-off lead only to modest acceleration

Following up bugs found at the Tier0

- Good news: There were not many bugs, in particular very few fatal ones
 - Most prominent: TRT irreproducibility
 - Bug treatment:
 - Comp@P1 shifter files a bug-report
 - Completely automated, content mostly ok
 - In some cases the automated log-extract does not contain the relevant information. For ERRORS (aka 'needcheck') this is never the case
 - Work by PROC/Reco expert (shifter) to identify the faulty package/subsystem
 - Not always evident
 - Sometimes developers adopt a 'not my problem' attitude and start working only after someone produced court-proof evidence that the bug is in their package
 - Help debugging
 - Downloading the RAW file requires privileged castor access
 - Extracting one problematic event
 - Some of the PROC work here is rather repetitive and could maybe be done by a shifter
- 
- Poorly covered for bugs found during reprocessing

About “Needcheck”

- Jobs that have an ERROR in the log are marked as “needcheck” in the Tier0 monitoring
 - Super-useful feature, I recommend to do something similar for Tier1 processing
- Can work only if there are very few jobs marked
 - Was mostly the case
 - Primary causes:
 - Various errors during HLT deserialization (hints to bugs in HLT)
 - Unconfigured propagator
 - dummy version of propagation method with search of nearest surface called; configure your extrapolator with STEP propagator
 - False alarm during Histogram merging
 - <TH1F::Add>: Attempt to add histograms with different number of bin

Link with Tier0 operations

- Excellent support & response times
- Configuration via AMI tags works nicely
 - Only glitch was when we hit a 4096 bytes limit on one of the fields

Outsourcing DESD production Tier1s

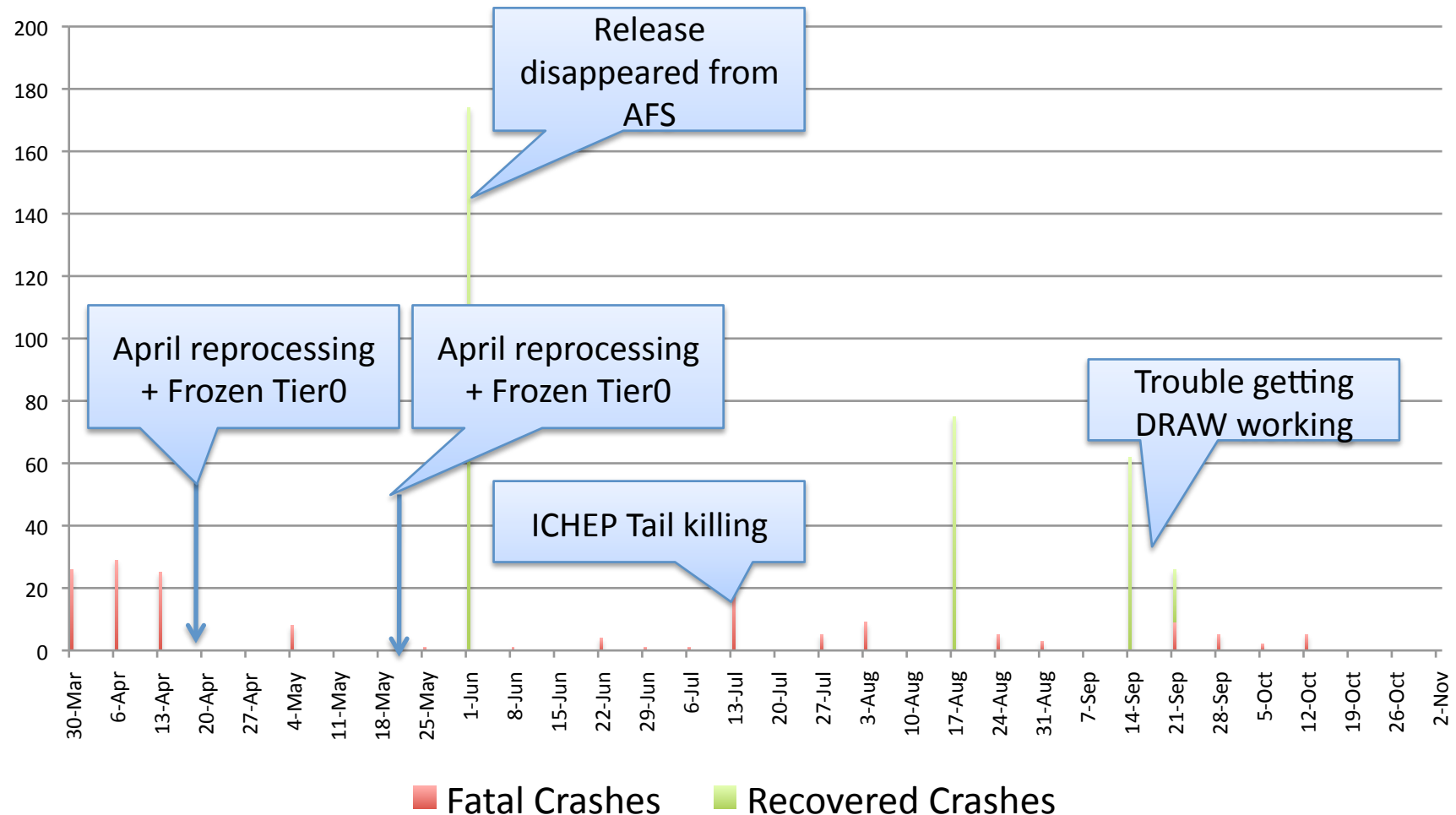
- Discussed a number of times
- IMHO a good idea, would relieve the Tier0
- Open questions
 - How to define the release/configuration?
 - For example, the mapping of trigger stream to DESD type changed a few times
 - At the Tier0 everything is done through AMI tags, can we do the same for Tier1?
 - Database access
 - In theory a ESDtoDESD jobs does not need DB access, in practice I am not so sure...
- Monitoring the DESD skimming rate is an issue
 - Changes because trigger changes
 - In 2010, reaction was often very slow, leading to problems and waste of disk-space

Failed jobs at the Tier0

- Various cases
 - Crashes on individual events, because of real SW bug, out-of-mem, etc.
 - Are not recovered at the Tier0 (unlike reprocessing)
 - Lead to fewer ESD/AOD than RAW files
 - Massive failures after re-configuration or infrastructure problems
 - Typically recovered by reverting/fixing configuration
 - Manual ‘tail killing’
 - We manually killed a few jobs to get the bulk of the data faster, eg before ICHEP
- In addition: Sometimes trouble with file merging
 - Mostly because files grew beyond 10 Gbytes/LB
 - A very few cases of corrupt temporary files

Number of failed jobs per week

(numbers from Armin)



Total number of failures: $159/1313944=0.012\%$
(roughly 1 failure per 10^7 events)

Christmas Wishlist

- Move StoreGateSvc WARNING about in-existent object to DEBUG.
 - Most (all?) clients check the StatusCode and react appropriately
- Fix coral bug #58522
 - Database connection failure causes infinite loop of Tier0 jobs
- Get rid of AlgErrorAuditor
 - See also 'needcheck' slide
- Better support for Reco_trf + PyJobTransformsCore
 - Alvin left Atlas
- **Speed-up**, in particular the initialize/finalize steps
 - I suspect we are wasting time and memory by loading things that are not used later on (in particular during ESDtoAOD/DPD)
 - I suspect there are also unnecessary overheads in the transform

Discussion on optimizing Tier0 running

- Proposal from DavidR yesterday evening
- Should involve operations experts and software experts (Reco, persistency, etc)
- Incomplete list of items to discuss
 - Experiment with athenaMP
 - Output file validation
 - Output file optimization
 - Athena steps per Reco_trf job

Conclusions

- Prompt processing of 2010 pp data at the Tier0 was a very successful operation
- The frozen-Tier0 policy and the 36h calibration loop fulfilled their purpose flawlessly
 - To adopt this strategy is a policy decision by ATLAS
 - I am confident that it was the right strategy for 2010
- The overall load on the Tier0 was $\ll 100\%$ in 2010
 - We had always resources to catch up in case of intermittent problems
 - Not necessarily the case in the future