HF jets analysis

30.03.2020 ALICE@IFJ meeting

Sebastian Bysiak

Outline

1. Merging of QA data for ML
2. Progress in HF-jets analysis
3. Questions & issues
4. Plans for next week

Outline

1. **Merging of QA data for ML**
2. Progress in HF-jets analysis
3. Questions & issues
4. Plans for next week

Merging of QA data for ML

- reconstruction (not my merging) of pass3:
  LHC18r completed
  LHC18q: 38/144 runs (Sunday evening)

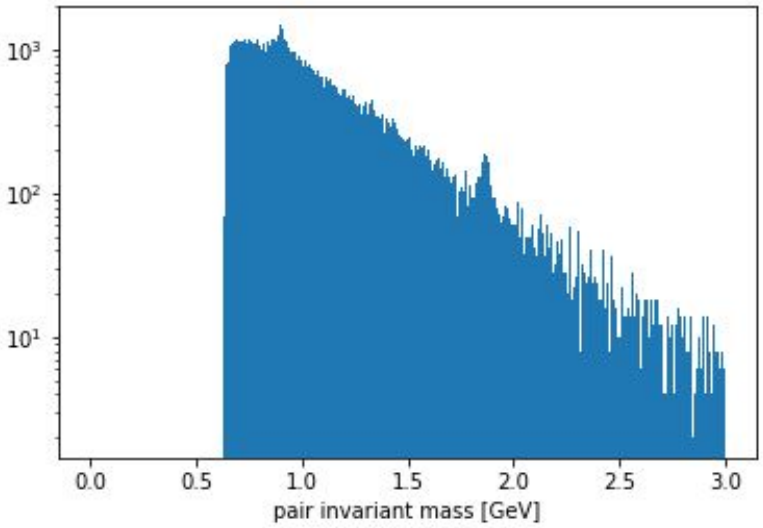
- huge fraction of failing jobs
  -> decision to refactor the code (see backup for story of my life:
  V1 = used in pass1, V2= current)

Outline

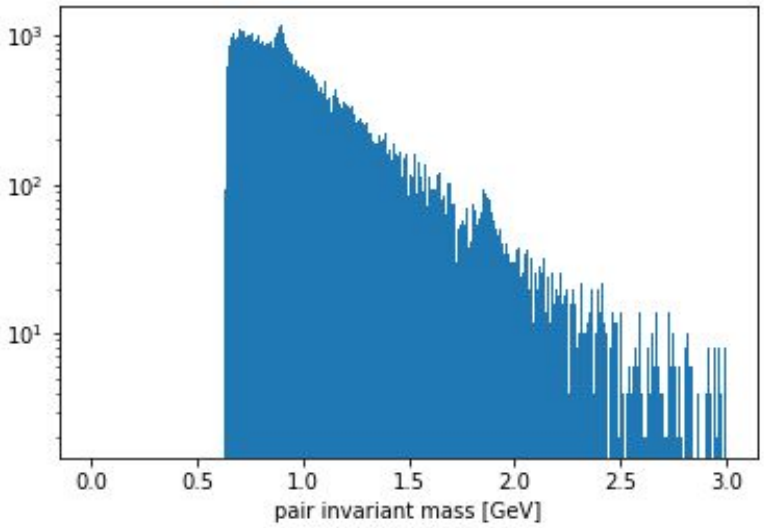
1. Merging of QA data for ML
2. **Progress in HF-jets analysis**
   - D-meson peak
   - chemical composition of jets
3. Questions & issues
4. Plans for next week

PID: D-meson peak

c-jets



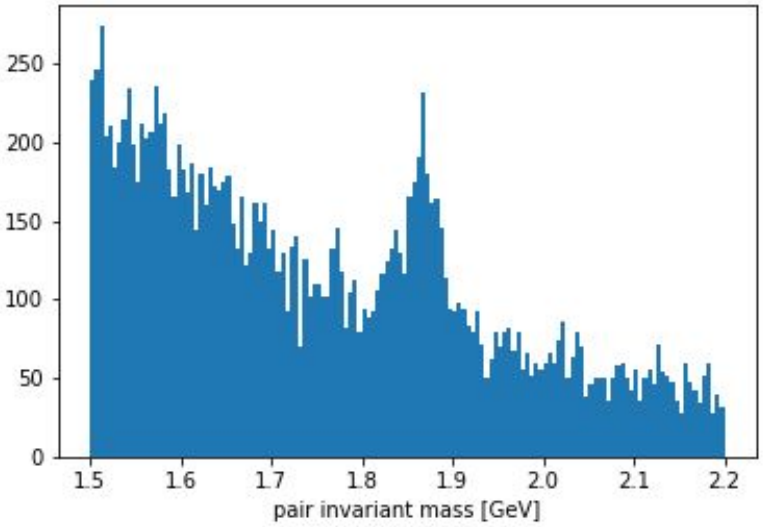
b-jets



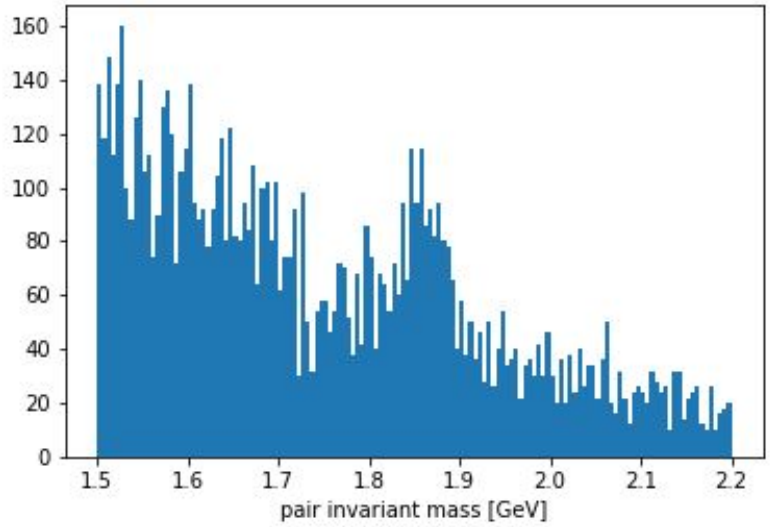
pi-K pairs with opposite signs
clearly visible peak (mass of D+ / D0 / D_s+ is 1.869 / 1.864 / 1.968 GeV)

PID: D-meson peak

c-jets



b-jets



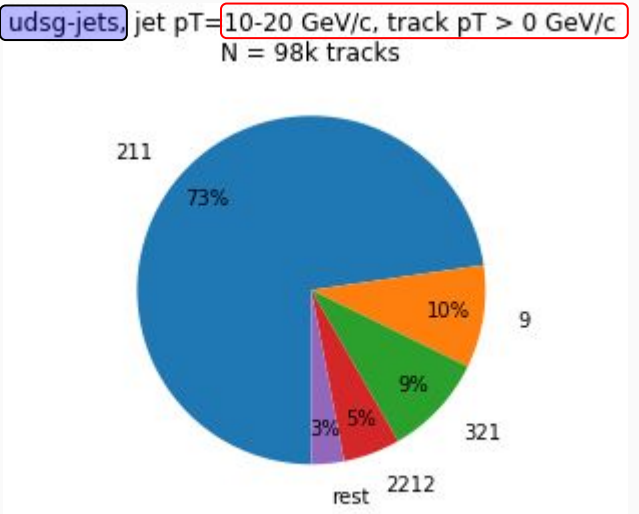
same plot but zoomed and with linear y scale

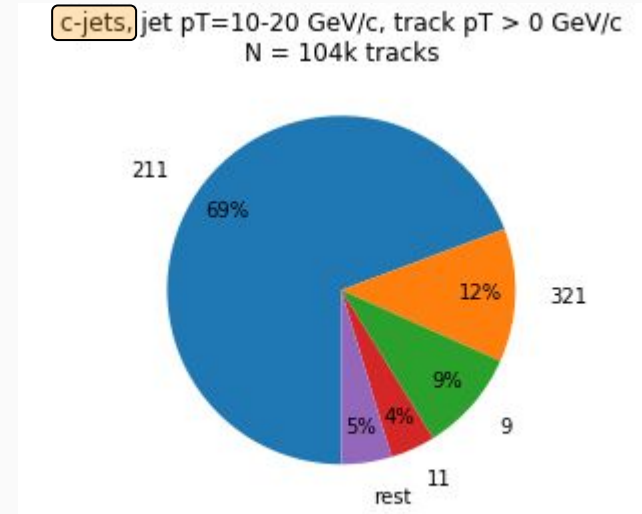
PID: chemical composition of the jet

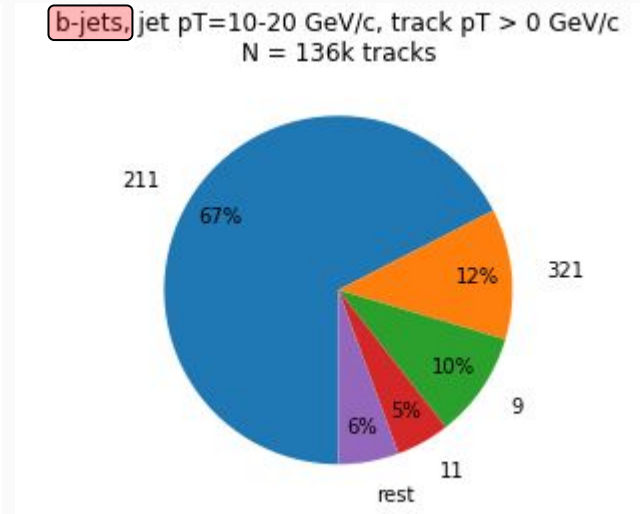
- PID observables inspired by: https://arxiv.org/pdf/2003.09517.pdf
- IDEA: usage of PID for jet tagging is not well recognized in literature due to lack of good PID capabilities (beyond leptons/muons?) in other LHC experiments

- our jet multiplicities: 10-20 GeV/c: 6+/- 2 tracks
  50-100 GeV/c: 10+/- 4 tracks

PID: low pT jets, all tracks



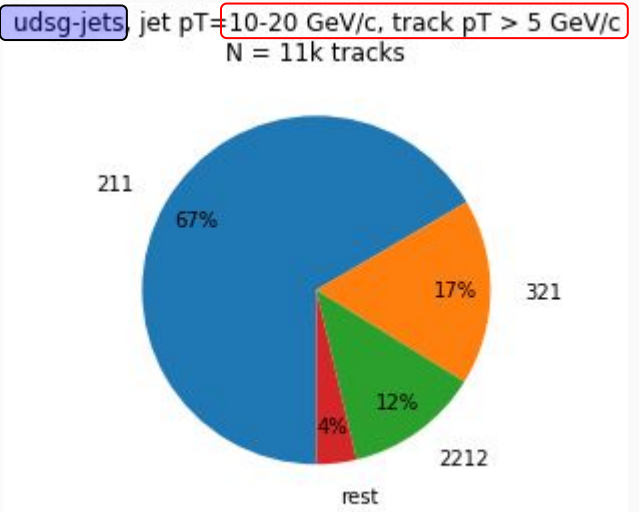


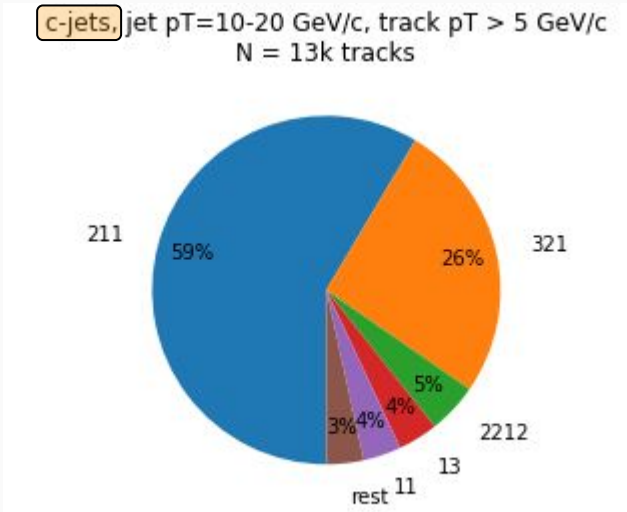


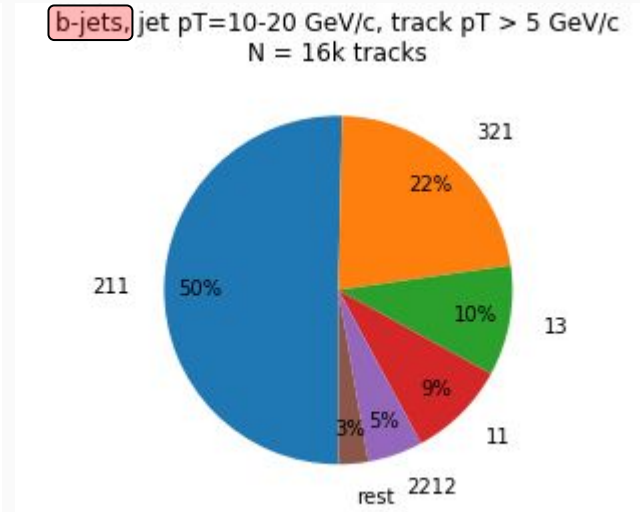
- similar composition, with larger contribution of electrons in HF
- “9” -> 10%

codes:
211=pi, 321=K, 2212=p, 11=e, 13=mu, 9=**
warning: colors does not match PID!

PID: low pT jets, high pT tracks



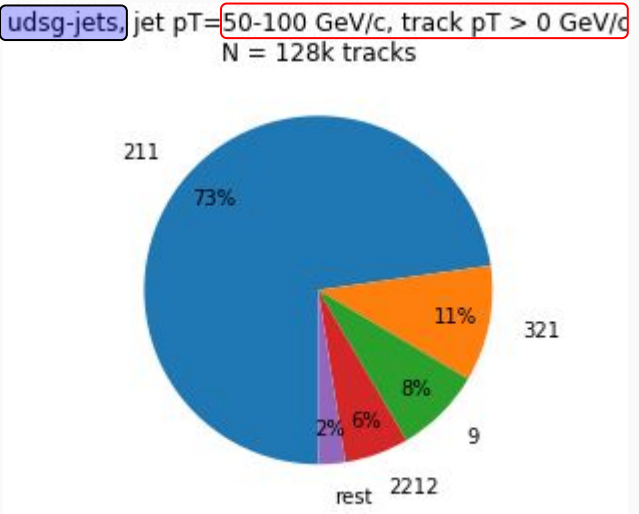


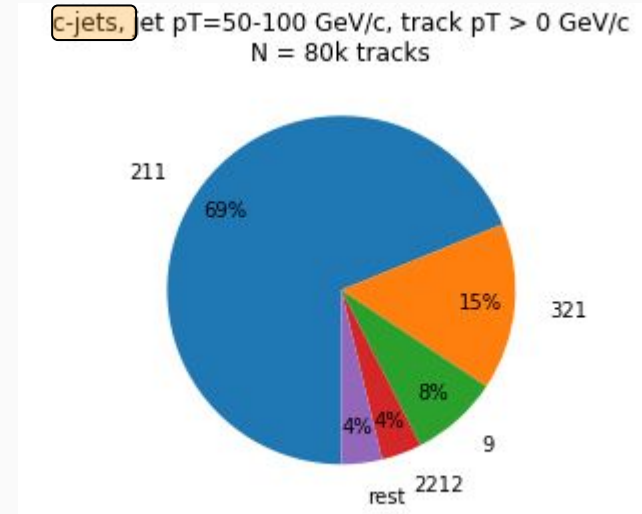


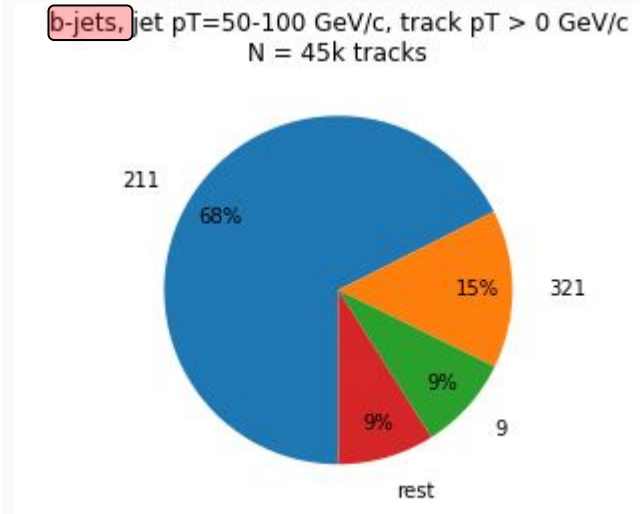
- ~2x larger contribution of kaons, largest for *c-jets*
- “9” greatly reduced
- leptons: 8% and 19% for *c-jets* and *b-jets*
- protons: 12% / 5% / 5% for *udsg-* / *c-* / *b-jets*

codes:
211=pi, 321=K, 2212=p, 11=e, 13=mu, 9=**
warning: colors does not match PID!

PID: high pT jets, all tracks



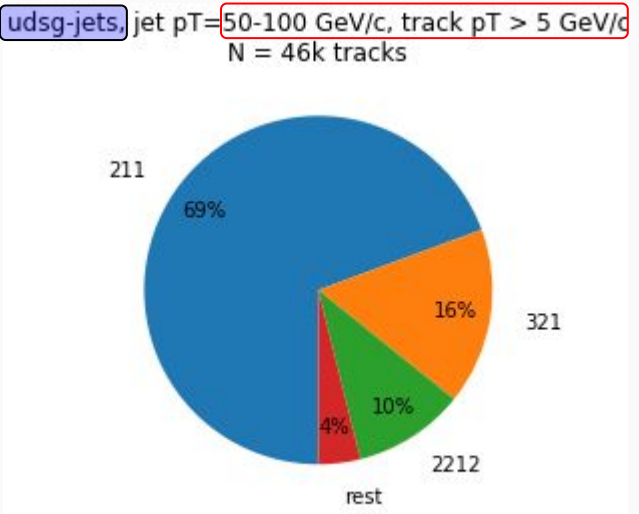


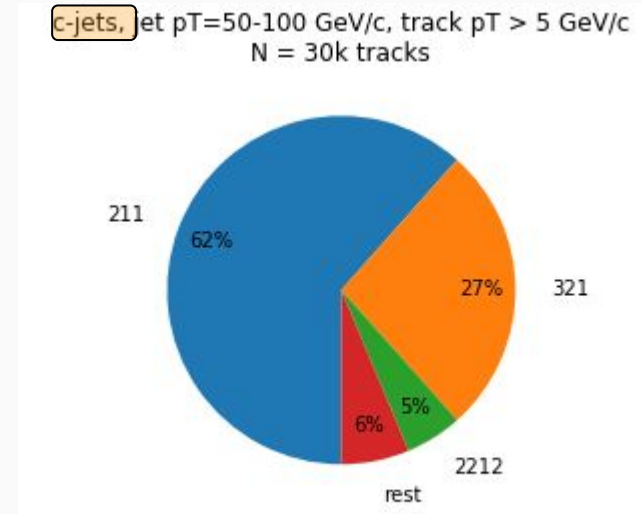


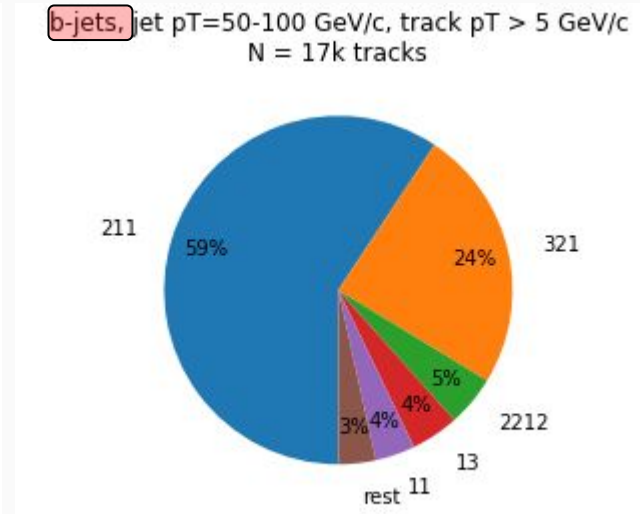
- similar to low pT jets, kaons +1-3 %

codes:
211=pi, 321=K, 2212=p, 11=e, 13=mu, 9=**
warning: colors does not match PID!

PID: high pT jets, high pT tracks







- larger kaons contribution compared to all tracks
- less leptons in HF compared to low pT jets: only 8% for *b-jets*!
- larger proton frac. in *udsg* than HF

codes:
211=pi, 321=K, 2212=p, 11=e, 13=mu, 9=**
warning: colors does not match PID!

PID: possible observables

- fraction of pT/multiplicity carried by certain types of particles
  + robust to misidentification as some errors cancel out
  + utilize a lot of information without large number of new columns
  + possible fully probabilistic approach, without hard PID (even more robust):
  e.g. track1 = 80/20/0% and track2=40/50/10% pion/kaon/proton
  -> pion/kaon/proton frac. in jet = 60/35/5%
- focus on single particles: e.g. leptons or tracks with high-pT or high IP
  + again, instead of binary flag “there is lepton” one can use
  (1 - (1-0.1)*(1 - 0.8)) = 0.82 for two tracks with 10% and 80% proba. of being lepton
- overall:
  + well physically motivated and experimentally verified signals
  + low correlation (?) with so far used observables
  - but high inside correlation e.g. between pi and K fraction
  + visible differences especially for low pT jets, where classification is the most difficult
  - reproduction in MC?
  - PID at higher momenta

anther pp 5.02 TeV: LHC17p & LHC17q

- DATA:
  LHC17p: 302k chunks, 400-780M events
  LHC17q: 58k chunks, 60-83M events
  LHC17p+LHC17q: 360k chunks, 450-860M events
  compared to LHC15n: 73k chunks, 180M events
- MC:
  - MC for HF jets in pp anchored to LHC17p/q:
    LHC18k6a2 (bb all, modified generator), LHC18k6a (bb all), LHC18k5a (bb), LHC18k6b (cc all), LHC18k5b(cc)
  - others: general-purpose, HF prod., pythia8+jetjet, incjected J/Psi and Upsilon

Plans for next week (discussion)

- several ideas from JacekB - WIP

BACKUP

particle with code 9

Two quotations from http://pdg.lbl.gov/2007/reviews/montecarlorpp.pdf :

" The gluon, when considered as a gauge boson, has official number 21. In codes for glueballs, however, ***9*** *is used to allow a notation in close analogy with that of hadrons."*

" As usual, ***9*** *rather than 21 is used to denote a gluon/gluino in composite states"*

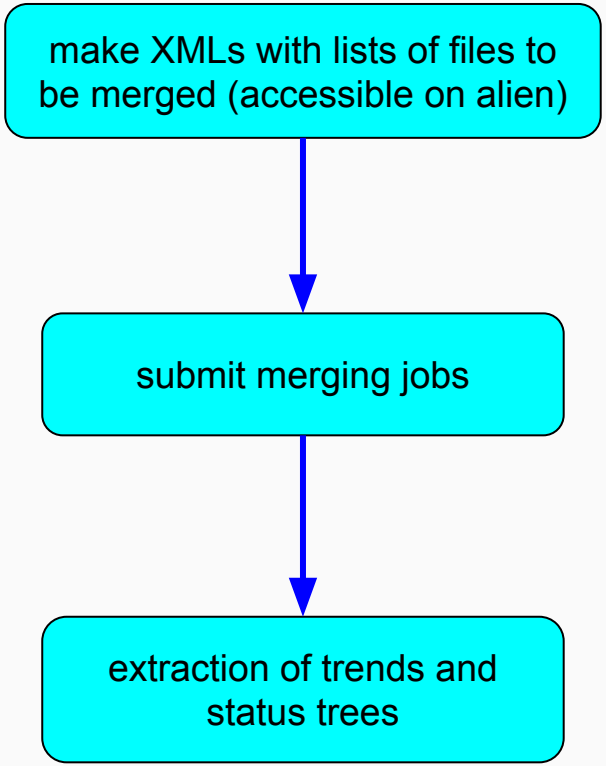
LHC17p & LHC17q VS LHC15n

						Raw data			Reconstructed					
LHC17p											ESDs			Output
Production	Description		Col.	Status	Run Range	Runs	Chunks	Size	Chunks	%	Size	%	Events	Size
LHC17p_VdM	LHC period LHC17p - Full production, VdM scan runs, ALIROOT-7634		pp	Completed	282026 - 282027	2	684	1.006 TB	684	100%	344.9 GB	33%	4,384,193	426.5 GB
LHC17p_pass1_CENT_woSDD	LHC period LHC17p - Full production pass 1, CENT trigger selection, without SDD, ALIROOT-7582		pp	Completed	282008 - 282343	42	301,167	503.8 TB	300,463	100%	51.18 TB	10%	396,530,021	66.84 TB
LHC17p_pass1_CENT_wSDD	LHC period LHC17p - Full production pass 1, CENT trigger selection, with SDD, ALIROOT-7582		pp	Completed	282008 - 282343	42	301,167	503.8 TB	300,783	100%	51.73 TB	10%	396,923,297	67.86 TB
LHC17p_pass1_FAST	LHC period LHC17p - Full production pass 1, FAST trigger selection, without SDD, ALIROOT-7582		pp	Completed	282008 - 282343	42	301,167	503.8 TB	301,063	100%	92.71 TB	18%	781,811,750	115.4 TB
LHC17p_muon_calo_pass1	LHC period LHC17p - Muon+Calorimeters reconstruction pass 1, ALIROOT-7583		pp	Completed	282008 - 282343	42	301,167	503.8 TB	300,845	100%	29.73 TB	5%	1,178,239,939	40.91 TB
5 productions							**1,205,352**	**1.969 PB**	**1,203,838**		**225.7 TB**		**2,757,889,200**	

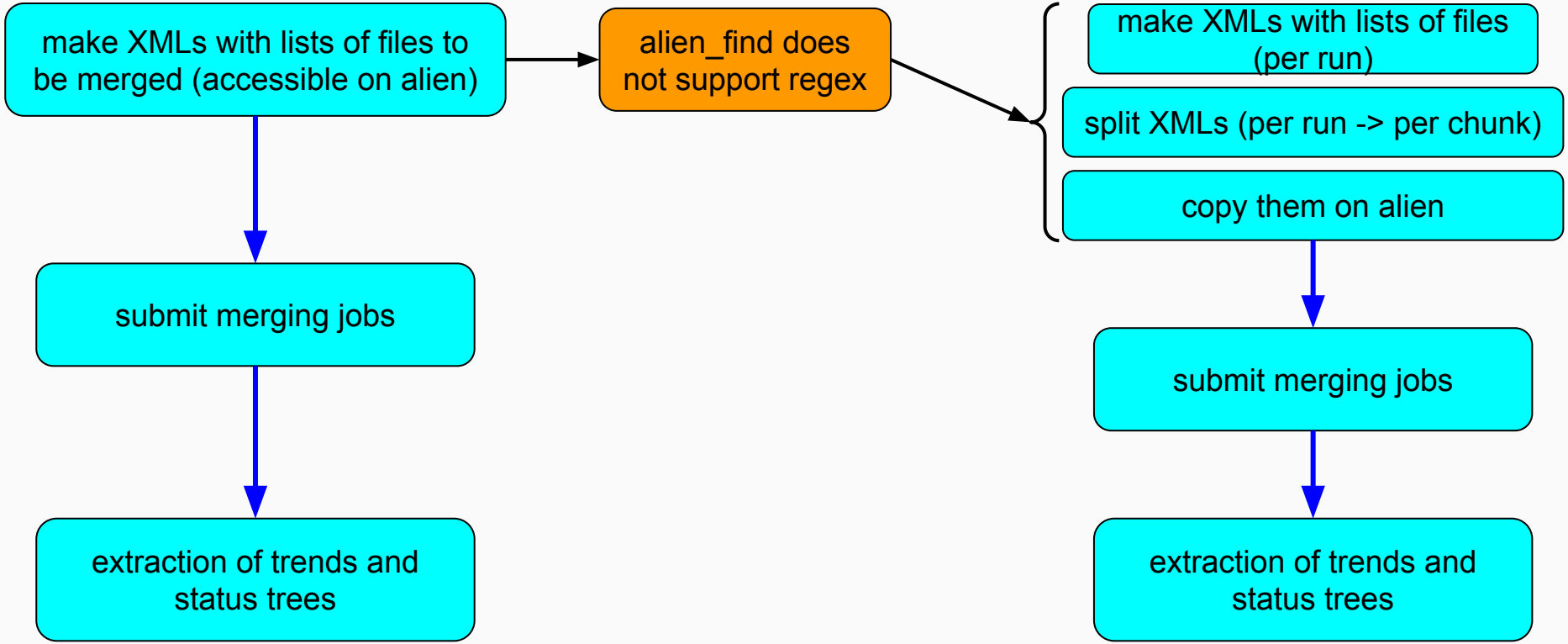
					Raw data			Reconstructed					
LHC17q										ESDs			Output
Production	Description	Col.	Status	Run Range	Runs	Chunks	Size	Chunks	%	Size	%	Events	Size
LHC17q_pass1_CENT_woSDD	LHC period LHC17q - Full production pass 1, CENT trigger selection, without SDD, ALIROOT-7599	pp	Completed	282365 - 282441	15	58,234	86.47 TB	58,214	100%	14.86 TB	17%	60,063,309	18.62 TB
LHC17q_pass1_CENT_wSDD	LHC period LHC17q - Full production pass 1, CENT trigger selection, with SDD, ALIROOT-7599	pp	Completed	282365 - 282441	15	58,234	86.47 TB	58,213	100%	15.04 TB	17%	60,062,531	18.87 TB
LHC17q_pass1_FAST	LHC period LHC17q - Full production pass 1, FAST trigger selection, without SDD, ALIROOT-7599	pp	Completed	282365 - 282441	15	58,234	86.47 TB	58,231	100%	17.49 TB	20%	83,972,738	21.63 TB
LHC17q_muon_calo_pass1	LHC period LHC17q - Muon+Calorimeters reconstruction pass 1, ALIROOT-7600	pp	Completed	282365 - 282441	15	58,234	86.47 TB	58,233	100%	3.657 TB	4%	144,036,841	5.301 TB
4 productions						**232,936**	**345.9 TB**	**232,891**		**51.05 TB**		**348,135,419**	

						Raw data			Reconstructed					
LHC15n_pass4											ESDs			Output
Production	Description		Col.	Status	Run Range	Runs	Chunks	Size	Chunks	%	Size	%	Events	Size
LHC15n_pass4	LHC period LHC15n - Full production pass 4, ALIROOT-7217		pp	Completed	244340 - 244628	27	75,647	121.2 TB	73,324	97%	20.33 TB	17%	180,620,663	25.48 TB
1 productions							**75,647**	**121.2 TB**	**73,324**		**20.33 TB**		**180,620,663**	

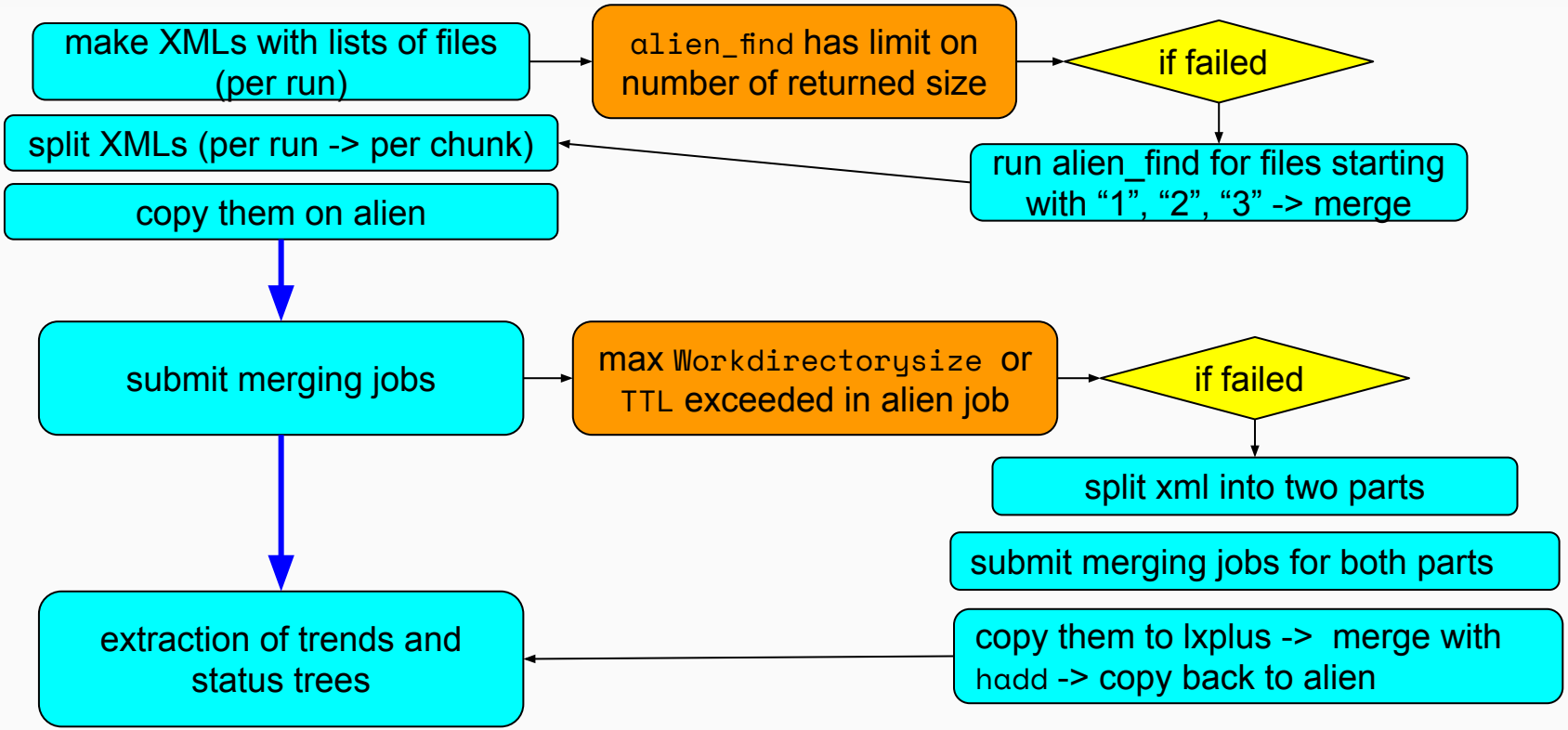
Merging procedure (concept)



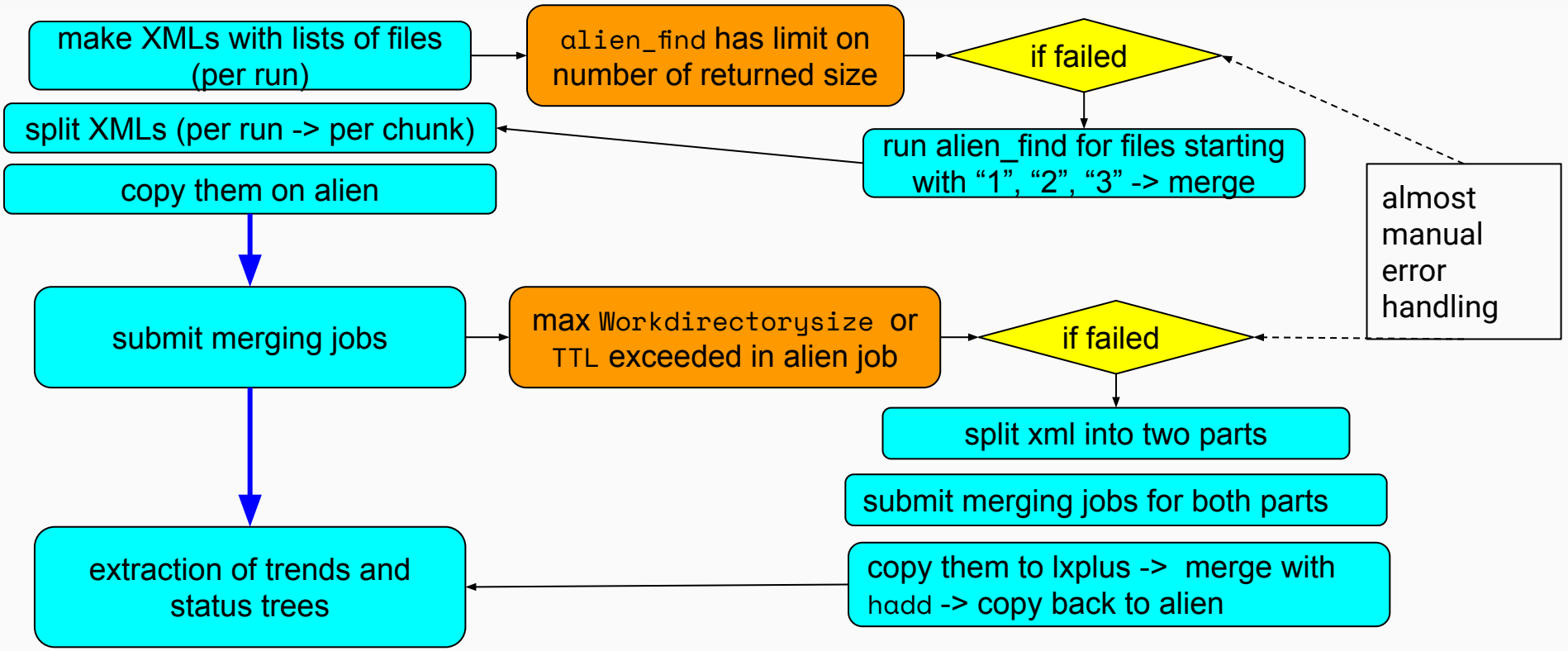
Merging procedure (concept -> V1)



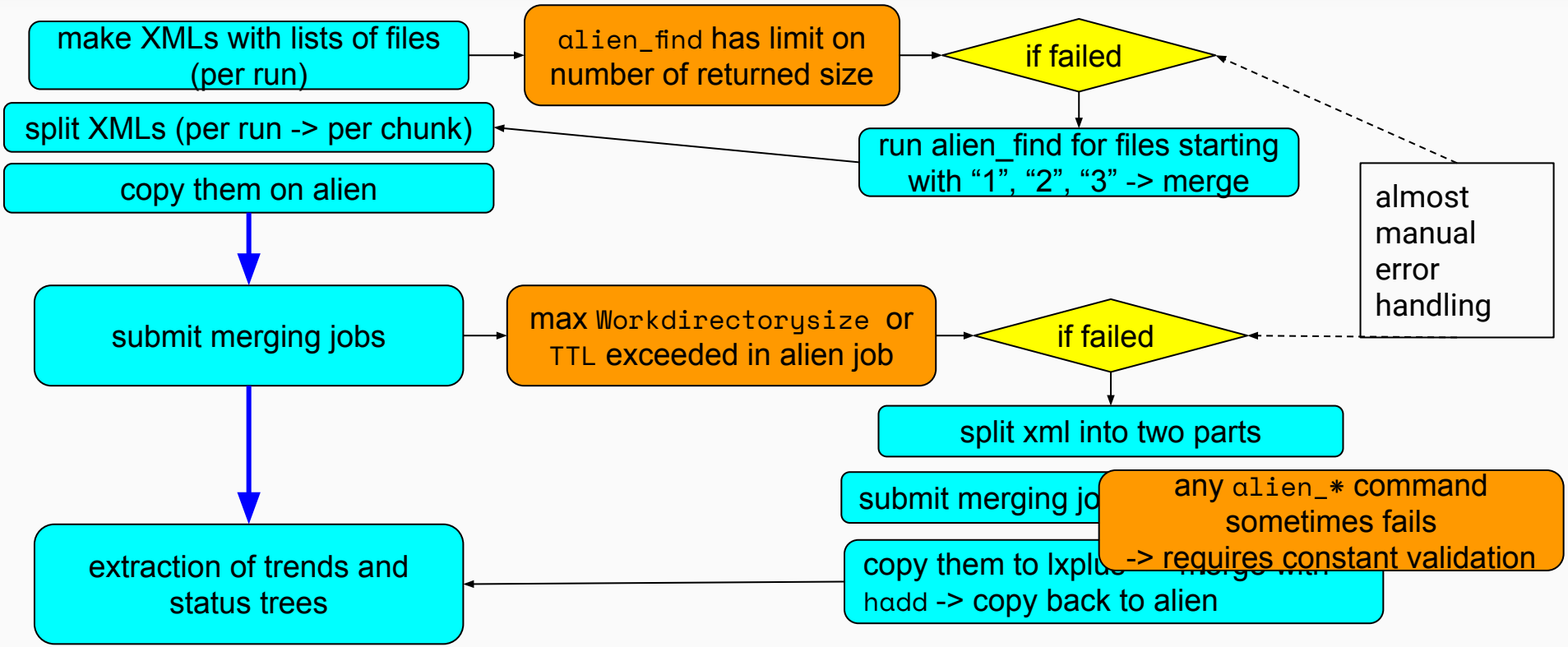
Merging procedure (V1 -> V1+patches)



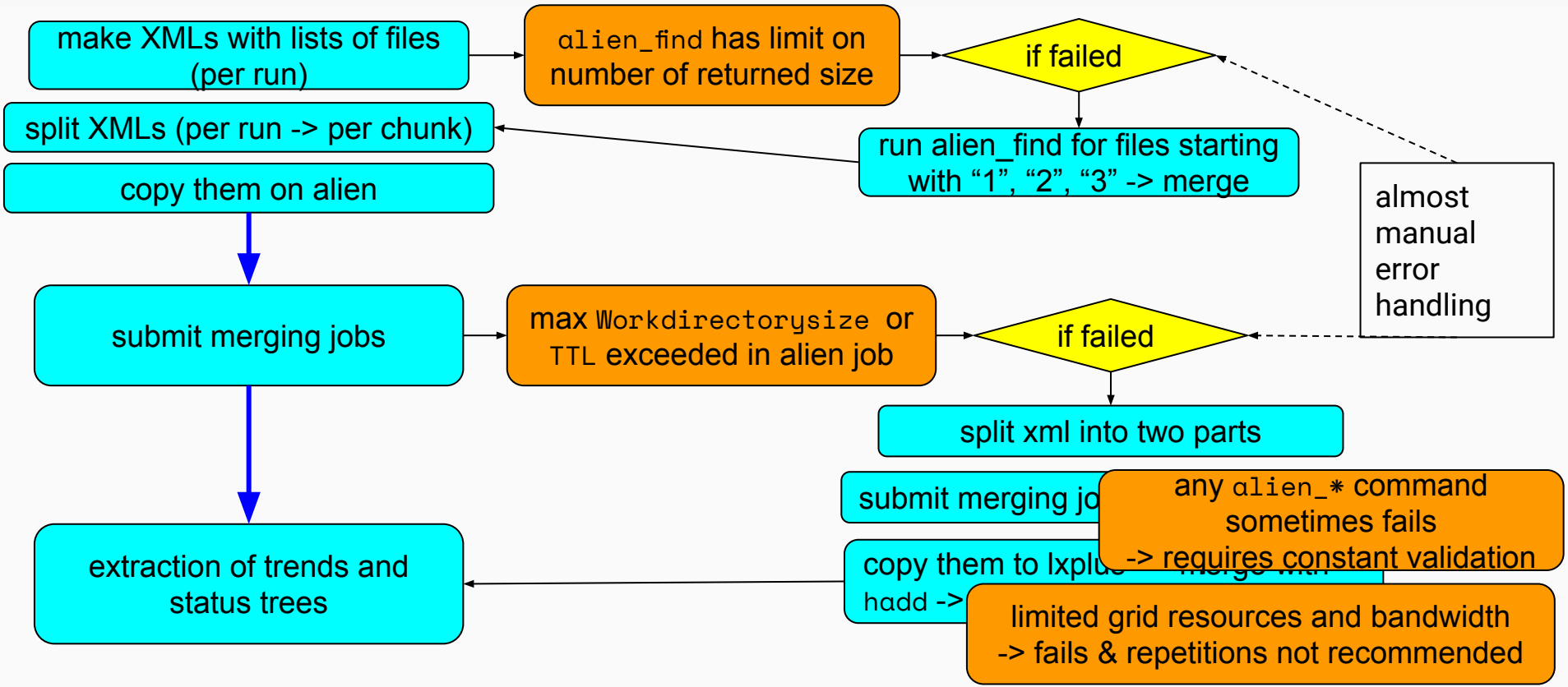
Merging procedure (V1 -> V1+patches)



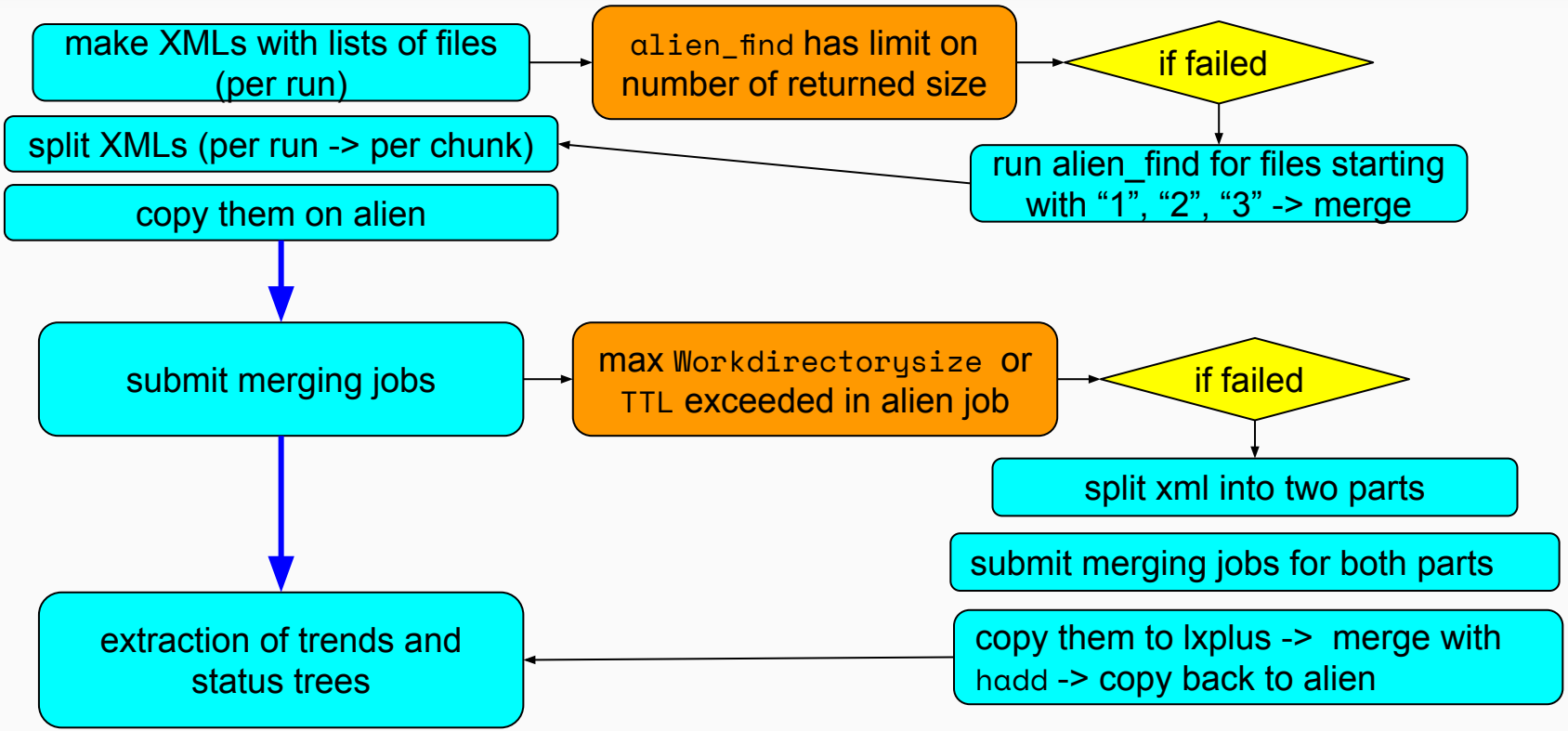
Merging procedure (V1 -> V1+patches)



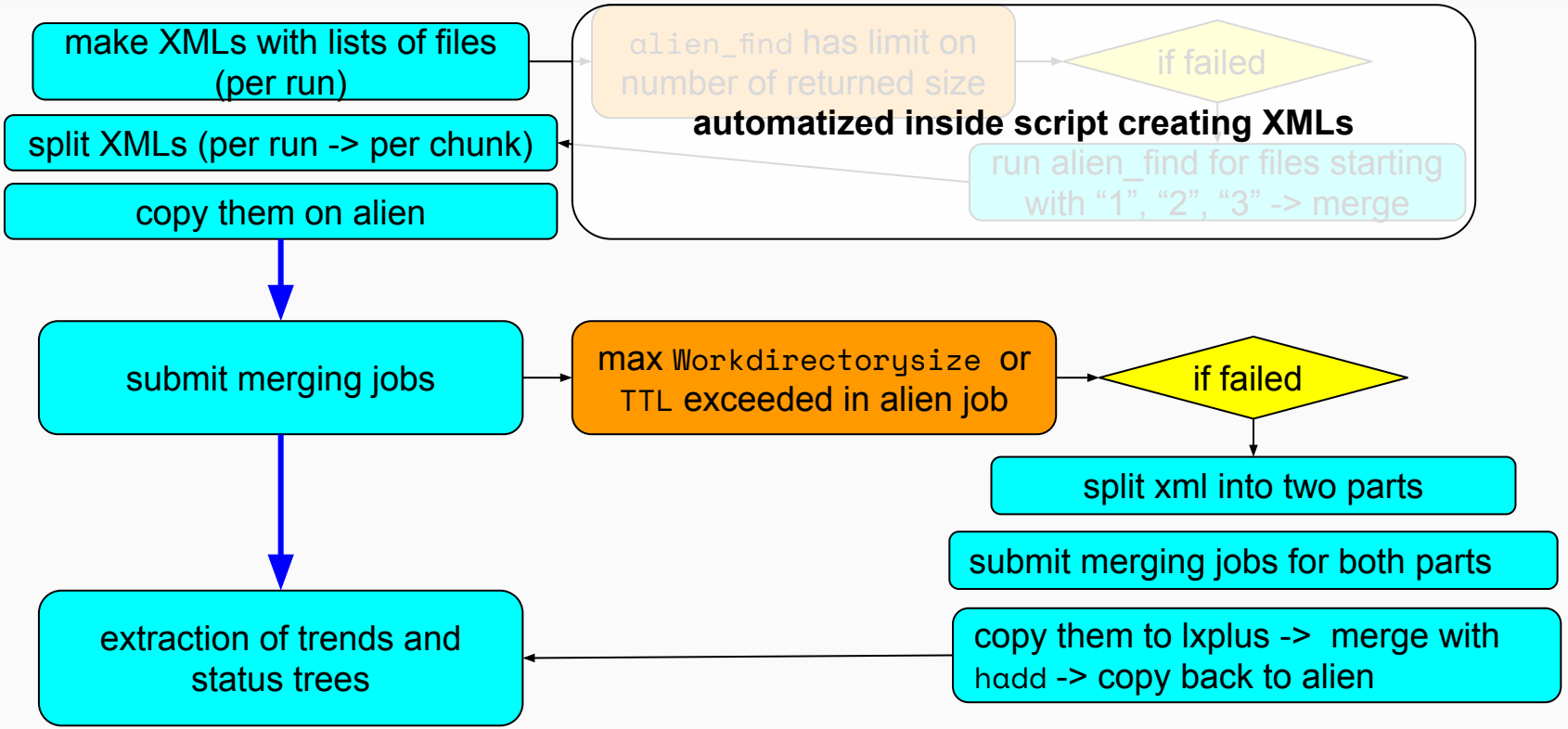
Merging procedure (V1 -> V1+patches)



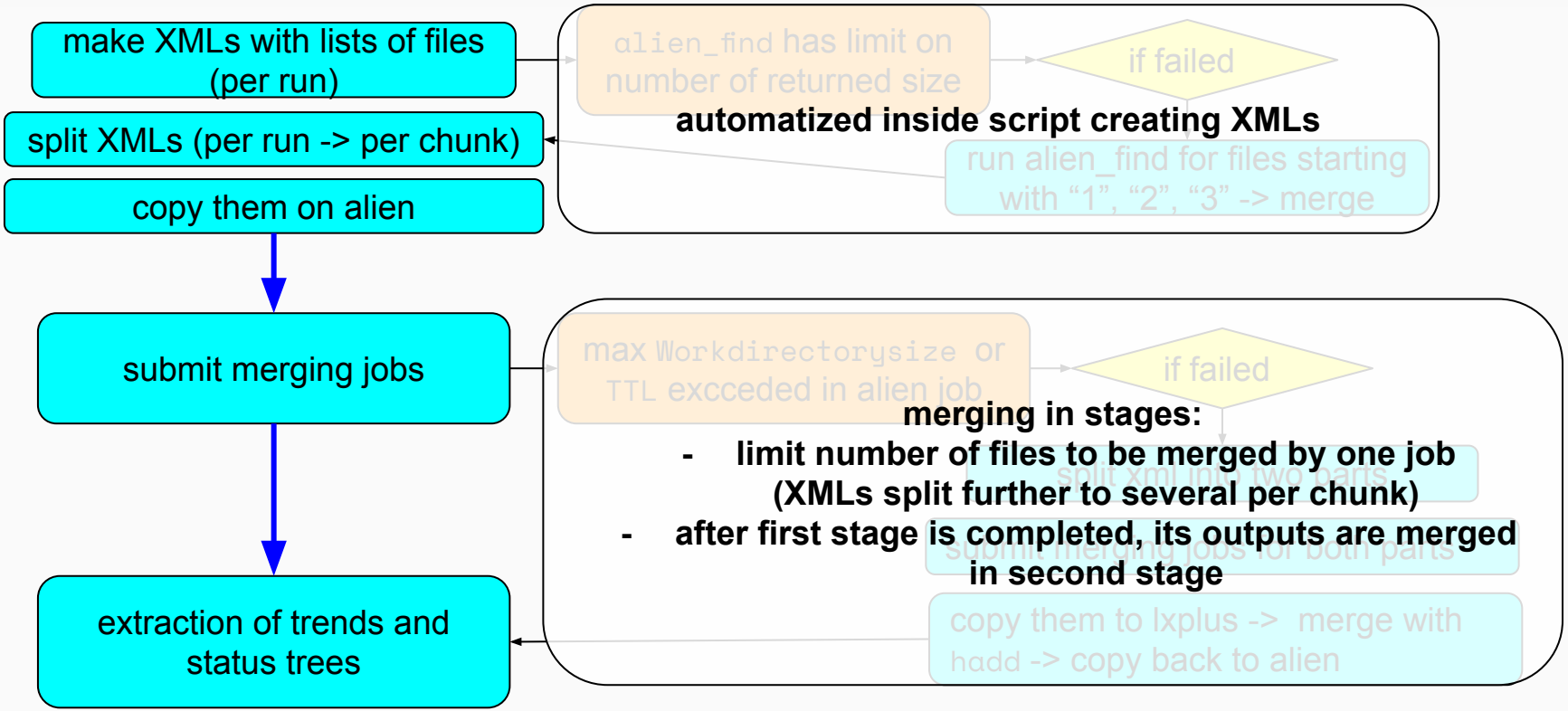
Merging procedure (V1+patches -> V2)



Merging procedure (V1+patches -> V2)



Merging procedure (V1+patches -> V2)



Merging procedure (V1+patches -> V2)