Analysis in ALICE and User Analysis Support

Andrei Gheata, Mihaela Gheata

Andreas Morsch

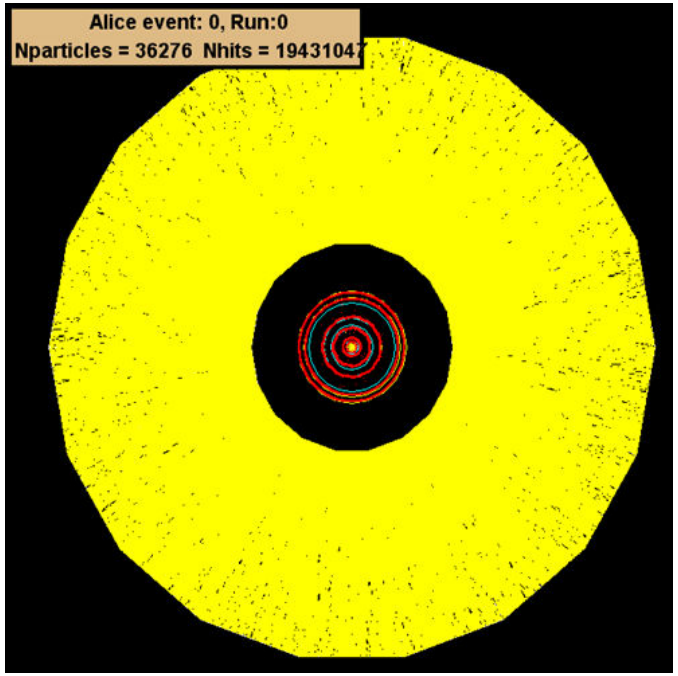
CERN/PH-AIP

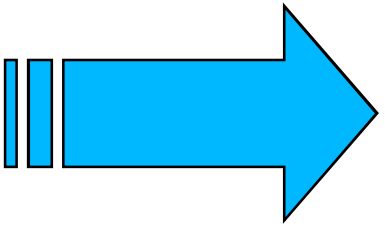
WLCG STEP'09 Post-Mortem Workshop

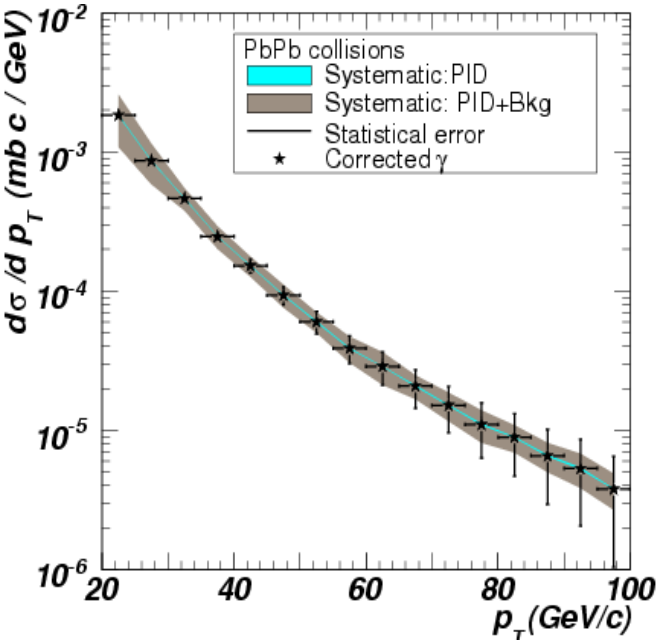
July 10, 2009

Analysis in ALICE

- Three main analysis modes
  - Prompt data processing (calibration, alignment, reconstruction, analysis) @CERN with PROOF
  - User analysis with local PROOF clusters
  - Batch Analysis on the GRID infrastructure



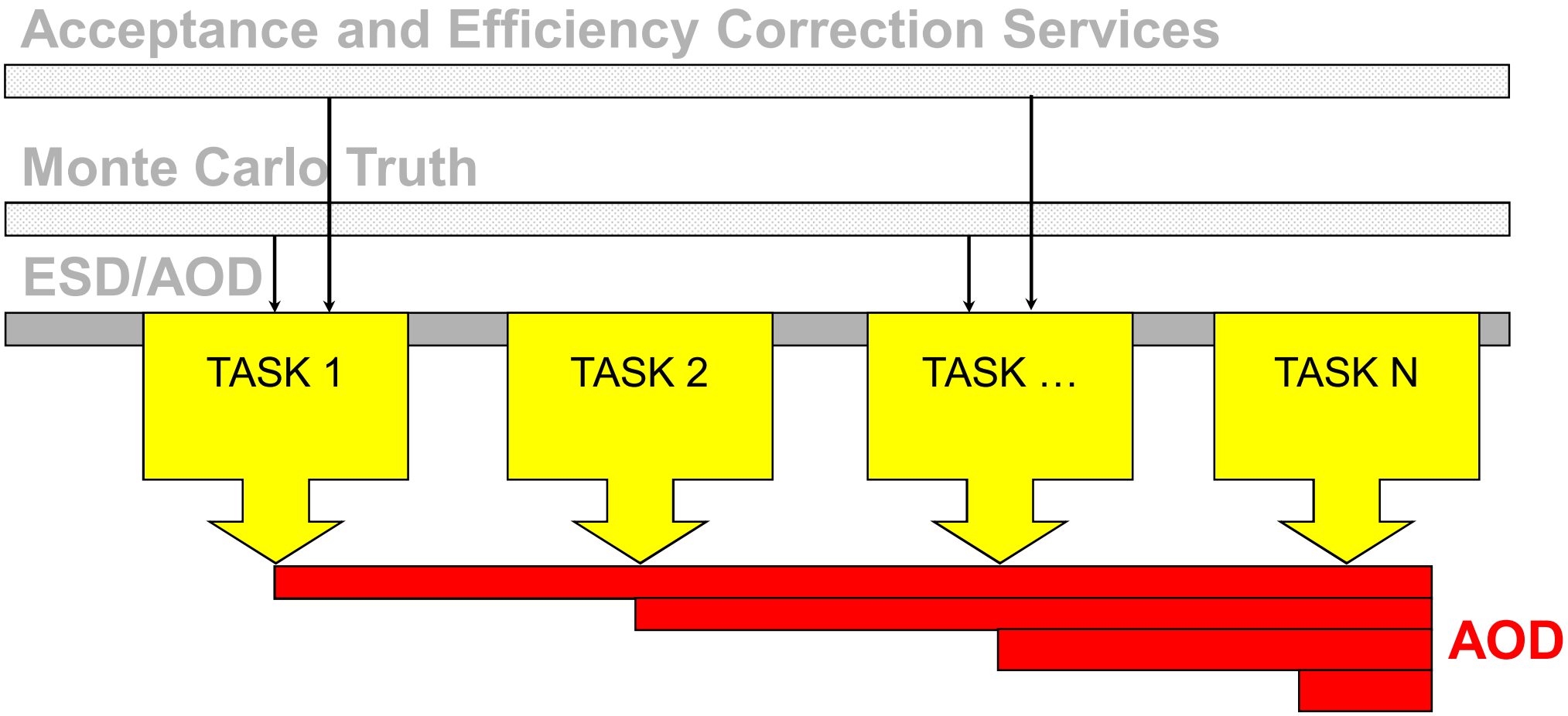




Plans for Scheduled Analysis

- Most efficient way for many analysis tasks to read and process the full data set.
  - In particular if resources are sparse.
  - Optimise CPU/IO ratio
- But also
  - Helps to develop a common well tested framework for analysis.
  - Develops common knowledge base and terminology.
  - Helps documenting the analysis procedure and makes results reproducible.

Plans for scheduled analysis: Analysis train



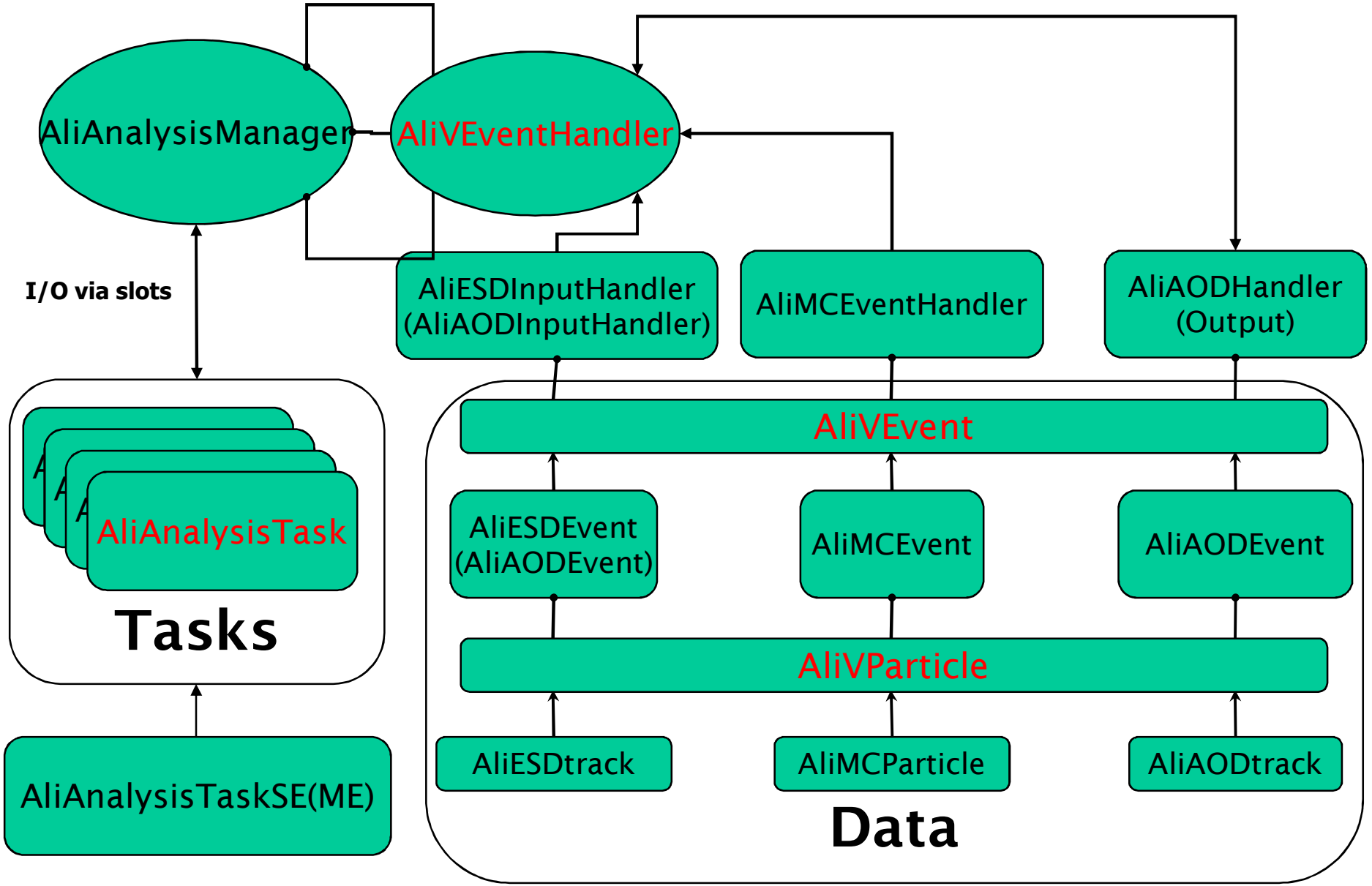
What the Analysis Framework does in ALICE

- Transparent access to all resources with the same code
  - Usage: Local, AliEn grid, CAF/PROOF
- Transparent access to different inputs
  - ESD, AOD, Kinematics tree (MC truth)
- Allows for „scheduled“ analysis
  - Common and well tested environment to run several tasks

How ?

- Transparent access to computing resources
  - Hide computing scheme dependent code in one Manager Class
- Transparent access to data
  - Make intensive use of interfaces
    - *VEventHandler*
    - *VEvent*
    - *Vtrack*
    - *Vvertex*

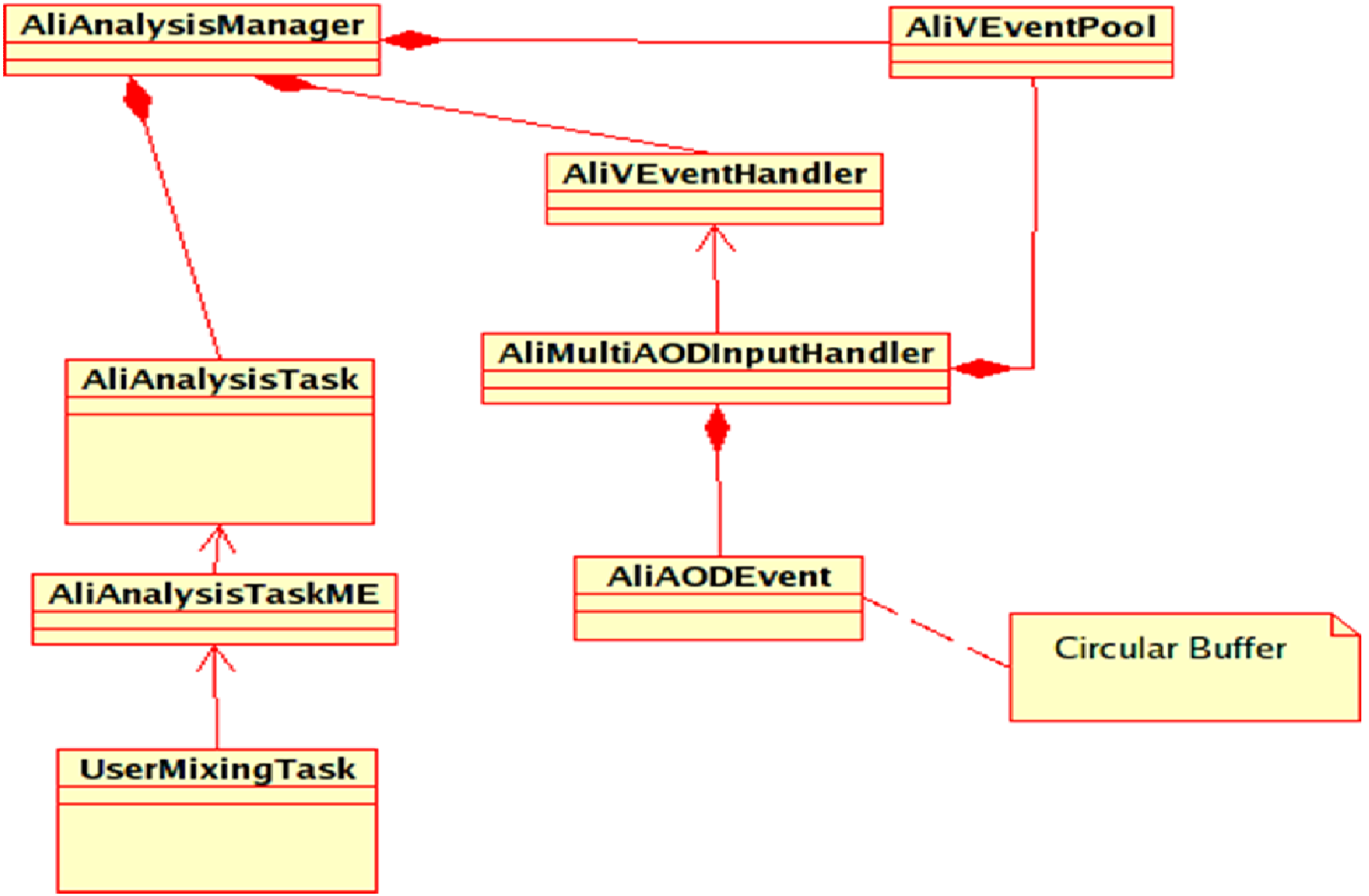
Tasks and Common I/O



User tasks derive from …

- *AliAnalysisTaskSE*
  - Single Event Analysis
  - Access to Input(ESD/AOD), output AOD and MC truth
- *AliAnalysisTaskME*
  - Multi Event Analysis (Event Mixing)
  - Access to Buffer of events which are similar according to user defined metric

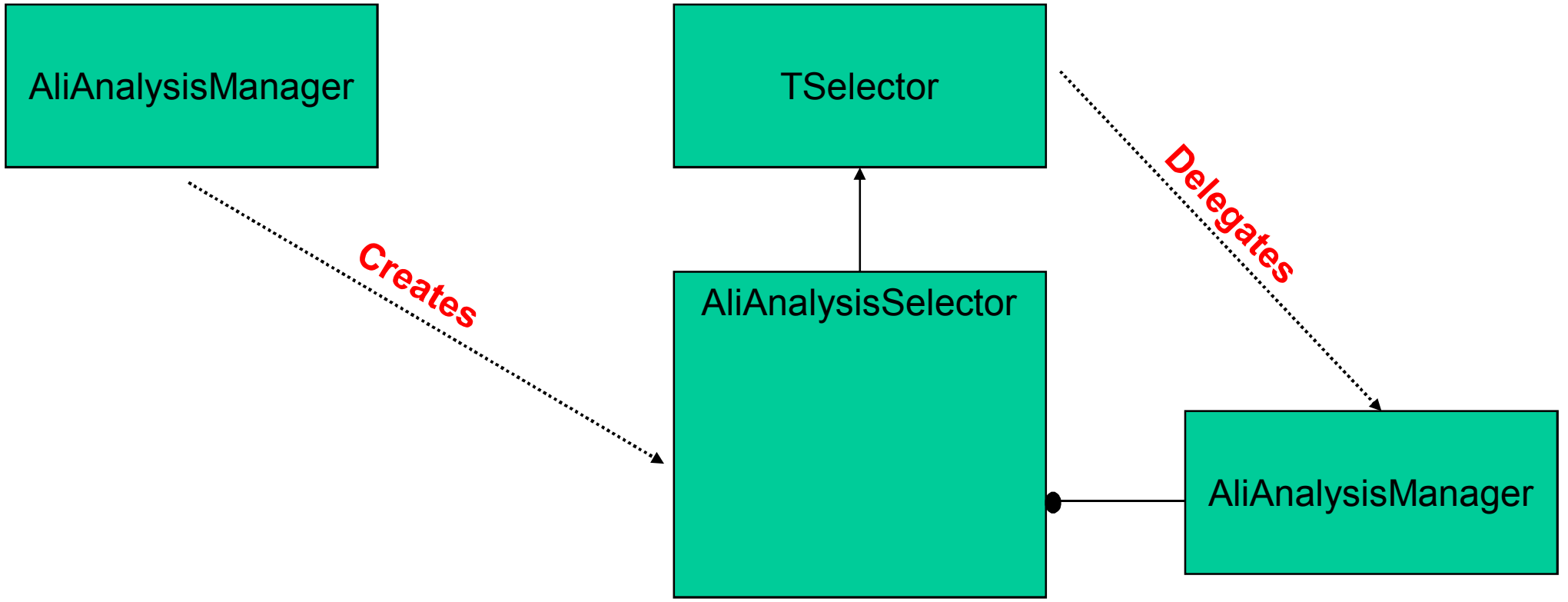
Analysis With Event Mixing



Transparent Access to Computing Resources

- Aim reduce computing scheme dependence in user code to
  - mgr->StartAnalysis(“local”);
  - mgr->StartAnalysis(“proof”);
  - mgr->StartAnalysis(“grid”);
- Some extra code for configuration is unavoidable but minimal
  - PROOF
    - TProof::Open(….)
    - gProof->UploadPackage(…)
  - GRID Configuration
    - Software versions
    - User space directory names

Example: *AliAnalysisManager* and PROOF:



GRID: Alien-Plugin

- Run local user code on the GRID without quitting the ROOT session
  - Main features
    - Automatic JDL creation
    - Generation of XML collection on demand
    - Detection of usage of tags
    - Analysis macro + validation script generation
    - Job submission
    - Automatic merging of outputs.

User Response/Support

- All user analysis code has been written using the analysis framework
  - Reduced complexity of the user code
  - Support is much easier
- Support issues reduce to
  - Training
    - Monthly tutorials
    - Explaining frame work features (mailing list)
  - Debugging on Grid and CAF (PROOF)
    - PROOF problems mainly related to
      - streaming of objects
        - Generally leads to code improvement
      - Compilation of “private” code
    - GRID/Alien
      - Data access/availability
      - Output SE
      - Software version (private code)

Train Status

PWG2 SPECTRA	AliAnalysisTaskProtons	p/pbar analysis	ESD/AOD +MC	Histograms, CF containers	OK	OK	OK
PWG2 SPECTRA	AliAnalysisTaskCheckCascade	QA for cascades	ESD/AOD	Histograms	OK	OK	OK
PWG2 SPECTRA	AliAnalysisTaskCheckPerformanceCascade	Performance study for cascade identification	ESD/AOD +MC	Histograms	OK	OK	OK
PWG2 SPECTRA	AliAnalysisTaskFemto	Femtoscopy	ESD/AOD +MC	Histograms	OK	OK	OK
PWG2 SPECTRA	AliAnalysisTaskCheckV0	V0 check	ESD/AOD	Histograms	OK	OK	OK
PWG2 SPECTRA	AliAnalysisTaskStrange	Strangeness	ESD/AOD	Histograms	OK	OK	OK
PWG2 FLOW	AliAnalysisTaskFlowEvent	Fill flow events from AOD/ESD/MC for flow analysis	ESD/AOD +MC	transient AliFlowEventSimple QA hists	OK	OK	OK
PWG2 FLOW	AliAnalysisTaskScalarProduct	Flow analysis using scalar product method	FlowEvent	Histograms	OK	OK	OK
PWG2 FLOW	AliAnalysisTaskLeeYangZeros (SUM & PROD)	Flow analysis using LeeYang zeros method	FlowEvent	Histograms	OK	OK	OK
PWG2 FLOW	AliAnalysisTaskCumulants	Flow analysis with cumulants method	FlowEvent	Histograms	OK	OK	OK
PWG2 FLOW	AliAnalysisTaskQCumulants	Flow analysis with Qcumulants method	FlowEvent	Histograms	OK	OK	OK
PWG2 RESONANCES	AliRsnAnalysisSE	Resonances analysis	ESD +MC	Histograms	OK	OK	OK

Train Status

PWG2 KINK	AliAnalysisTaskKinkESDMC	Kink topology study	ESD +MC	Histograms	OK	OK	OK
PWG2 KINK	AliAnalysisTaskKinkResonance	Reco. for resonances with a kaon kink	ESD +MC	Histograms	OK	OK	OK
PWG2 KINK	AliResonanceKinkLikeSign	Background computation for resonances with kaon kinks	ESD	Histograms	OK	OK	OK
PWG2 EVCHAR	AliAnalysisTaskSPDdNdEta	dN/dEta reco. with SPD	ESD +MC +TrackRefs	Histograms	OK	?	?
PWG2 UNICOR	AliAnalysisTaskUnicor	Unicor analysis modules	?	?	OK	?	?
PWG3 VERTEXING	AliAnalysisTaskSEVertexingHF	HF vertexing	ESD/AOD	delta AOD	OK	OK	OK
PWG3 MUON	AliAnalysisTaskESDMuonFilter	MUON filtering	ESD	AOD+ (adding muons to tracks)	OK	OK	OK
PWG3 MUON	AliAnalysisTaskFromStandardToMuonAOD	MUON AOD	AOD	new AOD with muons only	OK	?	?
PWG4 PARTCORR	AliAnalysisTaskParticleCorrelation (EMCAL, PHOS) + different plugged-in analysis	Particle correlations	ESD/AOD/MC	AOD+ (adding pi0, gamma) Histograms	OK	OK	OK
PWG4 JETAN	AliAnalysisTaskJets	JETAN	ESD/AOD/MC	AOD+ (reconstructed jets) Histograms	OK	OK	OK