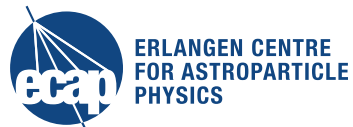
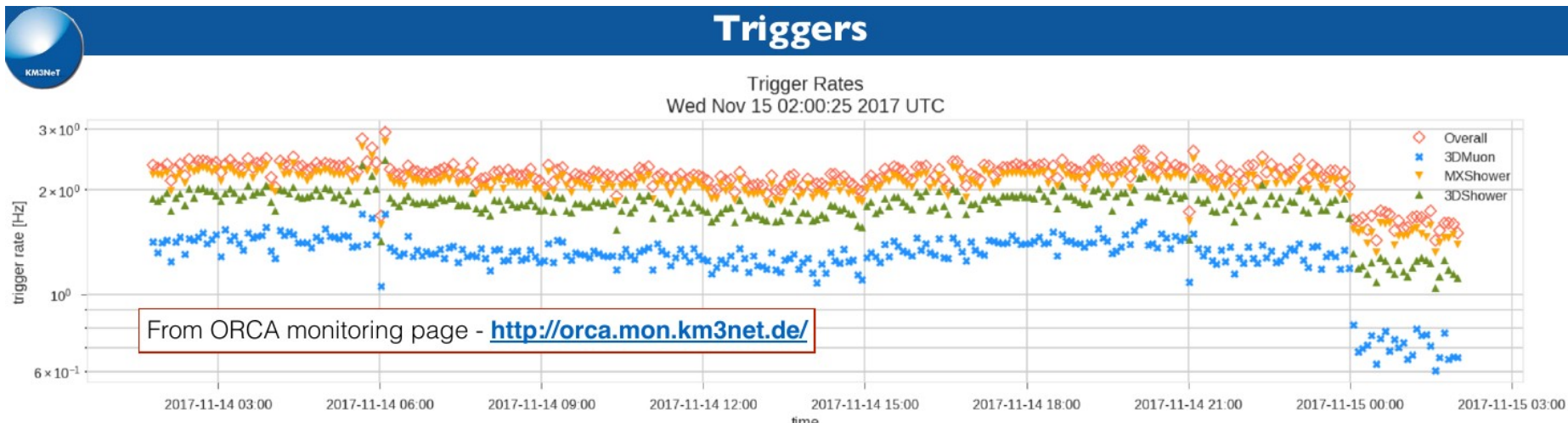


KM3NeT data processing and use cases

Jutta Schnabel
17th April 2019, Groningen

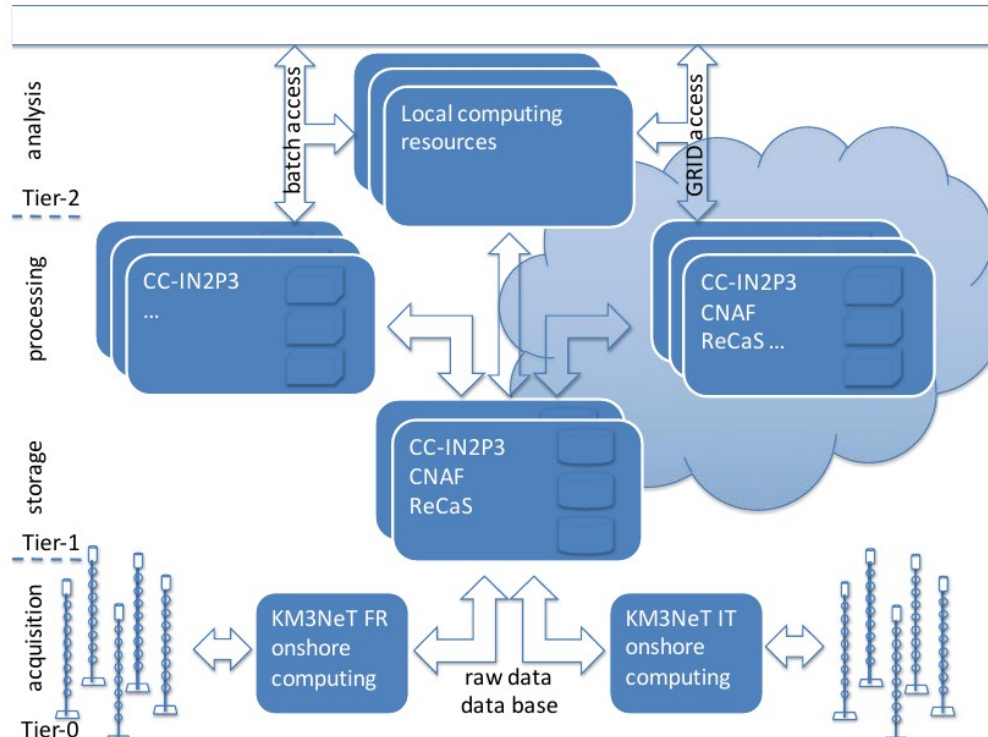


KM3NeT data – Event-based



- Count rate experiment → simulation necessary for analysis and reference
- Slow-control data for deeper understanding of data

KM3NeT Computing Modell/Data Management Plan

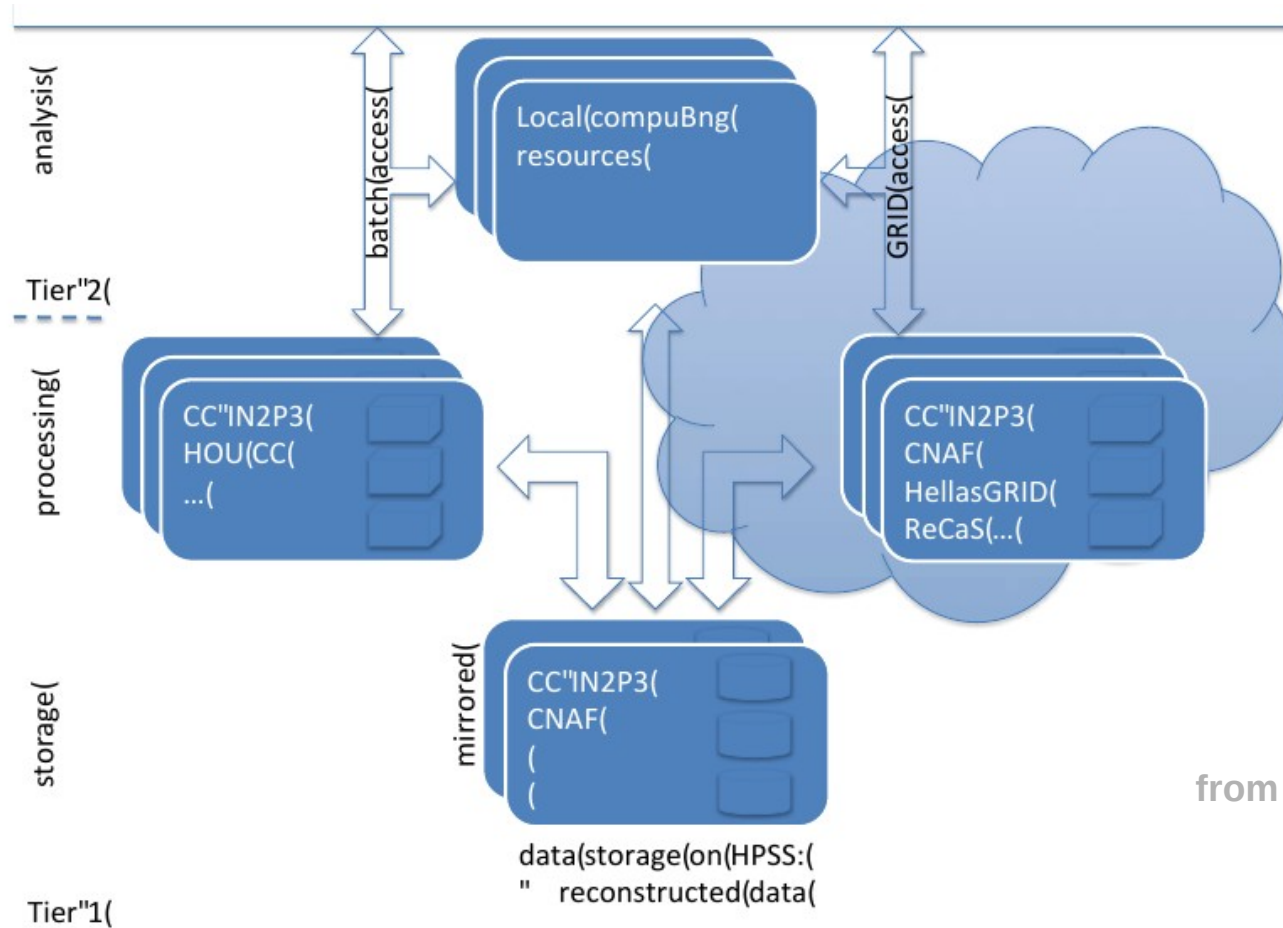


Data Management Plan
is first
written deliverable
for KM3NeT2.0

from Kay Graf

Tier	Computing Facility	Processing steps	Access
Tier-0	at detector site	triggering, online-calibration, quasi-online reconstruction	direct access, direct processing
Tier-1	computing centres	calibration and reconstruction, simulation	direct access, batch/grid processing
Tier-2	local computing clusters	simulation and analysis	varying

Tier-1: Include or not?



from Kay Graf

Currently: User as developer

Software Strategy

- implemented data-format-driven strategy for the KM3NeT software (both simulation and experiment data processing):

KM3NeT adopts a data processing chain with well-defined interfaces (mainly i/o and configuration data formats) at each processing step. One official software package per step will be maintained (with alternatives for crosschecks), which can read/write the official data formats directly.

- Data format bases:
 - ROOT
 - ASCII (plain or XML)
 - ⇒ added HDF5

from Kay Graf

Current frameworks

aanet

- low level analysis framework with high level capabilities
- main dev: Aart Heijboer - Nikhef / Amsterdam

OrcaNet

- deep learning training organiser
- main devs: Michael Moser, Stefan Reck - ECAP / Erlangen

KM3Pipe

- multi-purpose framework
- main devs: Tamás Gál, Johannes Schumann - ECAP / Erlangen
- credits: Moritz Lotze

Typical usage examples

Long-term work: Development of reconstruction strategy

- Integration of own software into existing chain
- Standardized interfaces and easy understanding of requirements

Short term work: Access existing software for target-specific analysis

- Customizable software chain (e.g. integrating deep learning tools)

External access: Easy accessibility of high-level data

- Limited access to fully processed data sets and simulation

Comp. Needs for 1 Building Block (per year) - estimates

processing stage	size per proc. (TB)	time per proc. (HS06.h)	size per year (TB)	time per year (HS06.h)	periodicity (per year)
Raw Data					
Raw Filtered Data	300	-	300	-	1
Monitoring and Minimum Bias Data	150	-	150	-	1
Experimental Data Processing					
Calibrated Data (only temporarily stored)	750	24 M	1500	48 M	2
Reconstructed Data	150	119 M	300	238 M	2
DST	75	30 M	150	60 M	2
Simulation Data Processing					
Air showers	100	14 M	50	7 M	0.5
atm. Muons	50	1 M	25	638 k	0.5
neutrinos	2	22 k	20	220 k	10
total:	827	188 M	995	353 M	

Vielen Dank für Ihre Aufmerksamkeit

ecap



Bundesministerium
für Bildung
und Forschung



FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG