Karlsruher Institut für Technologie

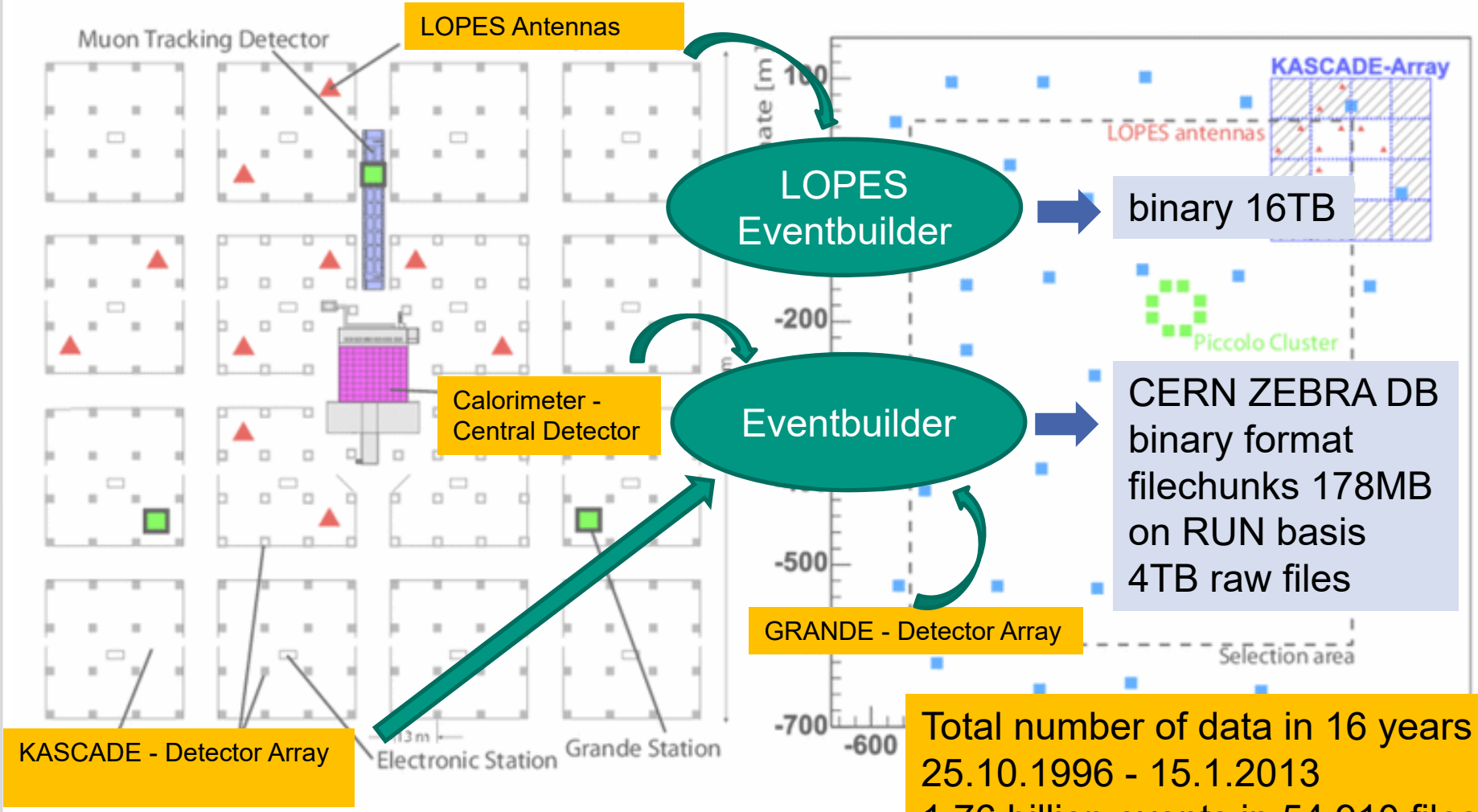
KASCADE Storage

Doris Wochele, Jürgen Wochele
03. April 2019

Institut für Kernphysik (IKP)

III International Workshop "Data life cycle in physics", DLC-2019
2-7 April 2019
Irkutsk State University

Main KASCADE detector components



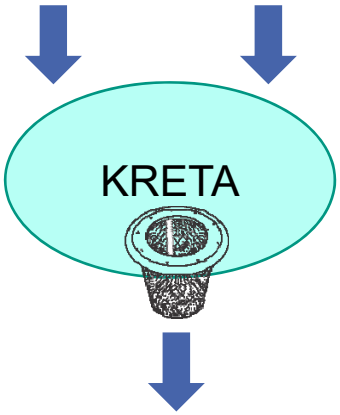
Total number of data in 16 years
25.10.1996 - 15.1.2013
1.76 billion events in 54.910 files
20 TB stored on RAID6 and Tape

All subcomponents used for data refinement and quality in reconstruction can be found in KCDC Manual Chap. 2

KASCADE analysis data

KASCADE

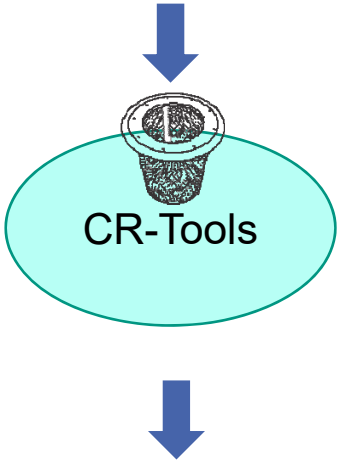
CORSIKA + CRES



4+2TB, ZEBRA + HBOOK → ROOT

430 Mio events, ~ 200 observables
metadata / calibrated / reconstructed

LOPES



4000 events, ROOT

LOPES has to be offline correlated with GRANDE events !

KRETA - KASCADE Reconstruction of Extensive Air showers

135k lines of FORTRAN in 860 CERN CMZ decks
analyse of all main and subcomponents in 3 levels
recently compiled on Ubuntu 18 with gcc with static cernlibs

KASCADE data published via KCDC

1. release **Wolf359** 18.11.2013 / 158 Mio events KASCADE
7 meta + 8 reconstructed scalar quantities KASCADE in **mysql**

2. release **Vulcan** 12.11.2014
7 meta + 11 rec. scalar quantities (KASCADE + Calorimeter) in **mysql**

3. release **Meridian** 22.3.2016
7 meta + 11 rec. scalar quantities in **mongodb**
+ 3 data arrays (each 252 calibrated observables per event)

4. current release **Naboo** 3.2.2017 / 430 Mio events KASCADE + GRANDE
7 meta + 11 rec. scalar quantities + 3 data arrays in **mongodb**
+ GRANDE Component with 7 rec. scalar quant. and 2 data arrays (37 calibrated observables per event)

5. next release **Oceanus** June/2019
+ UUID + LOPES Component in **mongodb** (→ talk Frank Polgart 4.4.)

KCDC data selection

430 Mio KASCADE
4000 LOPES

430 Mio KASCADE
3000 LOPES

(42) Document (3)	58889c4868b7e2343a714754...	Document
_id	58889c4868b7e2343a714754	ObjectId
array (8)		Document
E	14.201095581054688	Double
Xc	51.5677490234375	Double
Yc	-65.46118164062 5	Double
Ne	3.8891110420227051	Double
Nmu	3.3134799003601074	Double
Age	0.95509117841720581	Double
Ze	6.5479892322830962	Double
Az	232.0608953157091	Double
general (8)		Document
T	4.57958984375	Double
P	1014.0003662109375	Double
Gt	1078759321	Int32
M	669606400	Int32
R	4775	Int32
Ev	71658	Int32
Datetime	2004-03-08T15:22:01Z	DateTime
KDetectorsId	58889c4868b7e2343a714755	String

quality criteria for published data
→ applied when filling mongodb
→ separately for each detector component

mongodb ~2 TB on single server RAID6
2 collections for scalars and arrays
Indices on all scalars for user cuts

What physics is possible?
What cuts make sense?
What documentation is needed?
What simulations are needed?

Thank You!