KM3NeT e-Needs

Kay Graf
e-IRG open Workshop Luxembourg
November 2015

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

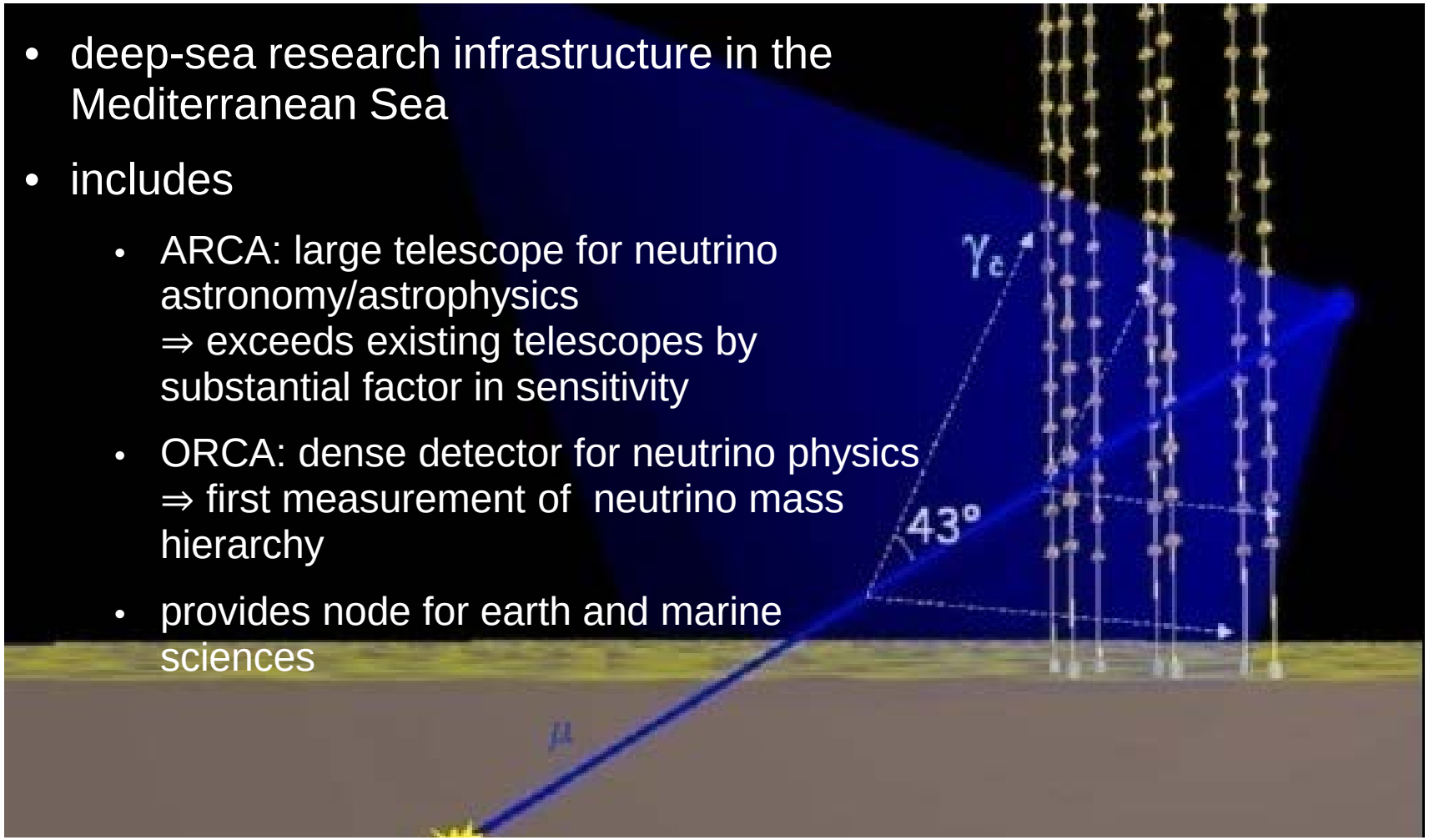
FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Astroparticle Physics e-Needs

- Astroparticle physics uses
  - detection techniques from particle physics to study astrophysical phenomena
  - different messenger particles: gamma, Cosmic Rays, **neutrinos**,…
- especially neutrino astronomy:
  - 1990/2000s: “branched” from particle physics
  - following: technology and analysis development phase (first generation projects)
    ⇒ “low” e-Needs, branch out at the “pre-grid” time
  - in between: intermediate generation experiments
  - now: the second/third generation detectors, including KM3NeT with e-Needs comparable to particle physics experiments again
    ⇒ “high” e-Needs, branch into the (almost post-) grid time
- KM3NeT solution: base on available e-Infrastructures (ANTARES) and include new developments in eCommons

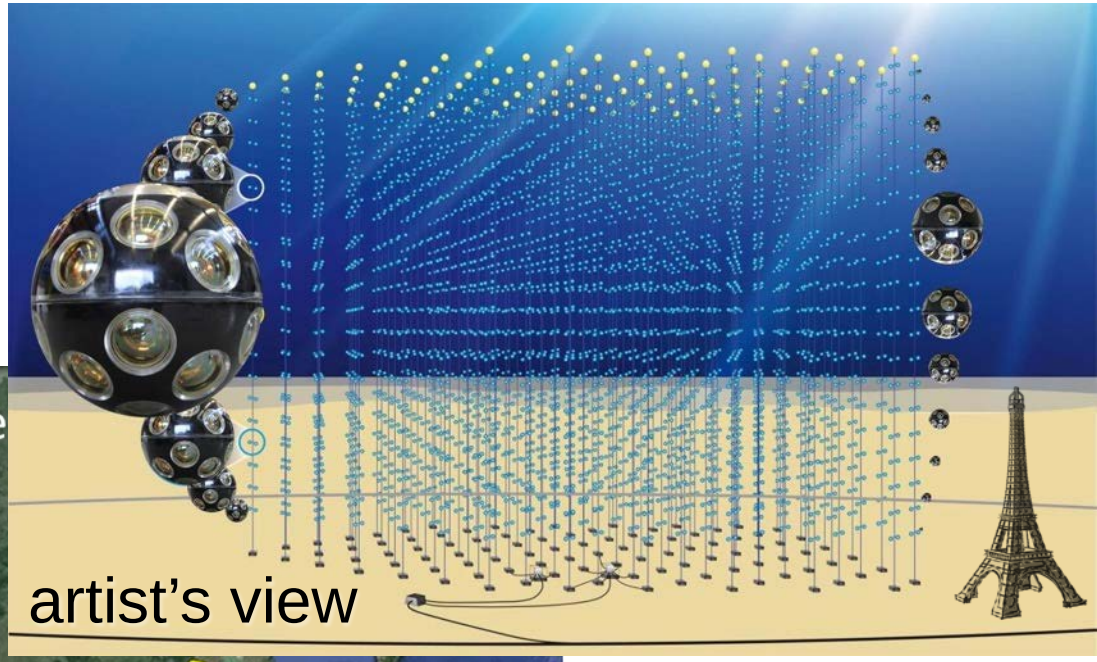
What is KM3NeT?

- deep-sea research infrastructure in the Mediterranean Sea
- includes
  - ARCA: large telescope for neutrino astronomy/astrophysics
    ⇒ exceeds existing telescopes by substantial factor in sensitivity
  - ORCA: dense detector for neutrino physics
    ⇒ first measurement of neutrino mass hierarchy
  - provides node for earth and marine sciences



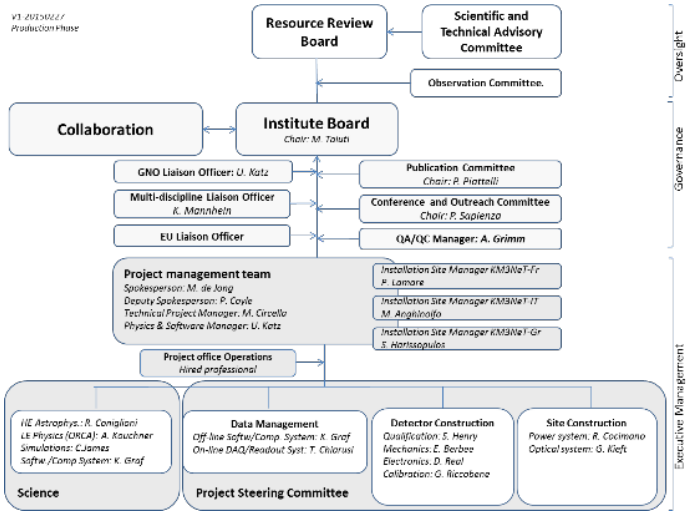
Detector and Collaboration

- European collaboration
- 3 installation sites
- starting installation of first phase



Status and Stages of the Project

- Phase-1 _____ *now*
  - ✓ transition from consortium to collaboration
    - ✓ Organisation and administration set up (management, steering committee, MoU)
    - ✓ Oversight in place (Resources Review Board, Scientific & Technical Advisory Committee)
  - ✓ acquired funds for 0.2 building blocks
  - ✓ started construction
  - start science (size = 3 x precursor experiment ANTARES)



- 2.0 _____ *next*
  - science capitalisation
  - funding requests on-going for 3 building blocks
- 3 _____ *future*
  - possible expansion

Application to ESFRI Roadmap 2016

Advantages w.r.t. original proposal (2006)?

- science _____ *augmented*
  - cosmic neutrinos exist (was hypothetical)!
  - significantly improved discovery potential for neutrino point sources
  - first determination of mass hierarchy of neutrinos
- technology _____ *proven*
  - interest from other experiments
  - e.g. Neutrino Telescopes Workshop, 2-6 March 2015, Venice
    - Tommaso Dorigo's blog during about KM3NeT optical module: *"I think these are quite cool objects."*
- costs _____ *reduced*
  - spending profile matched with realistic funding scenario
- implementation _____ *agreed*
  - coherent support from all parties

Science – Flashlights

- Origin of cosmic rays is – for more than 100 years – one of the outstanding scientific questions (old)
  - neutrino astronomy is at the edge to make a break-through (new)

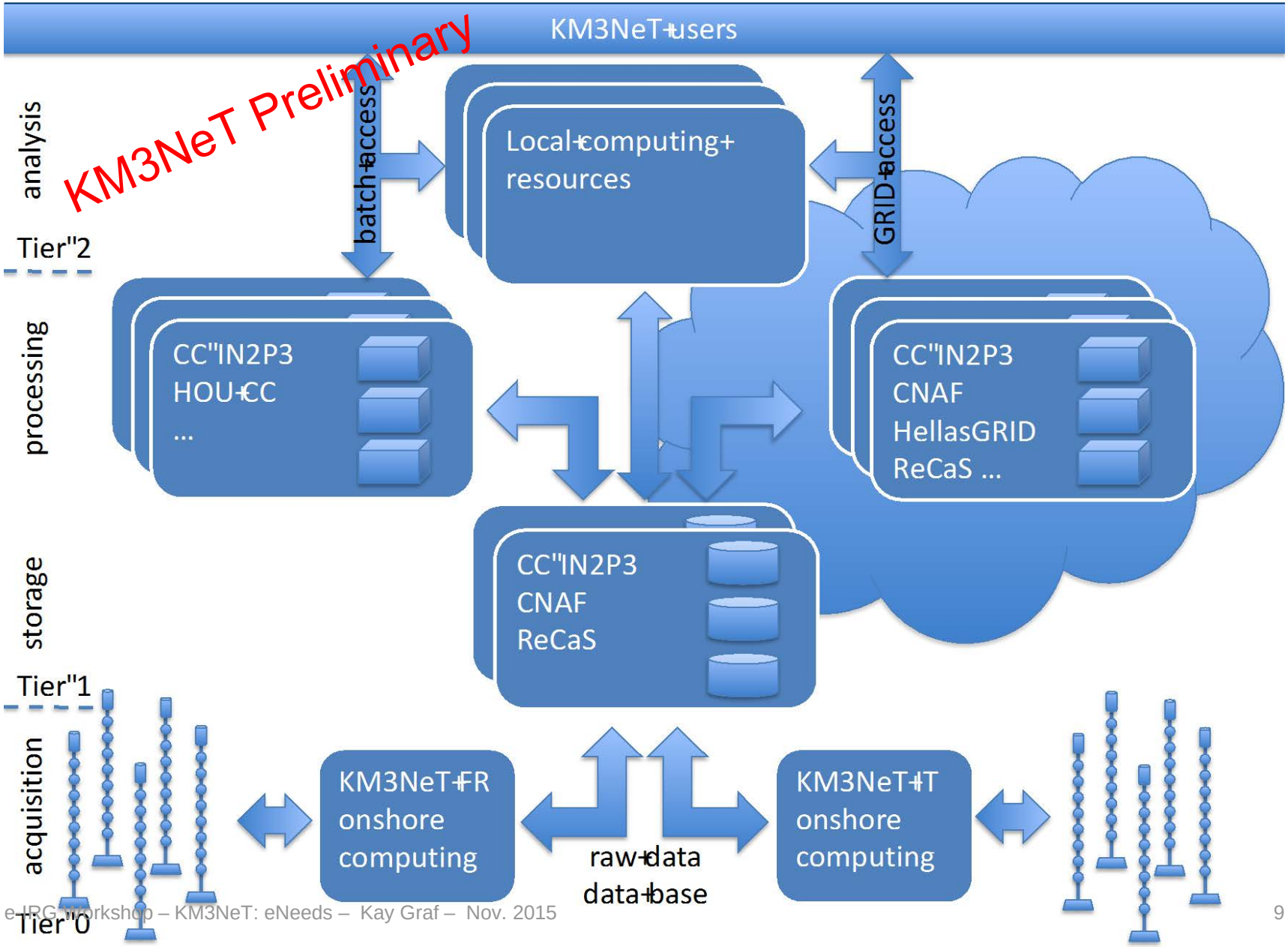
- CP-violation needed to explain [anti-]matter imbalance in the Universe is not yet found (old)
  - neutrino physics provides opportunity to find missing link between particle physics

Computing Model and Data Management: General Scheme

- Tier-like structure, mixed access: GRID + direct (batch)

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 processing and/or grid access
Tier-2	local computing clusters	simulation and analysis	varying

- Data transfer between the computing centers based on GRID access tools (where applicable)
- Central services funded through CNRS and INFN, additional services by the collaboration institutes


KM3NeT users
KM3NeT Preliminary
analysis
batch access
Local computing resources
GRID access
Tier 2
processing
CC IN2P3
HOU CC
...
CC IN2P3
CNAF
HellasGRID
ReCaS ...
storage
CC IN2P3
CNAF
ReCaS
Tier 1
acquisition
KM3NeT FR
onshore
computing
raw data
data base
KM3NeT IT
onshore
computing
Tier 0

Data Management Plan

- Data formats and meta-data following common practise (root, xml, ascii) → allow for integration in eCommons
- use of existing eCommons (e.g. from ANTARES) and specific development via ASTERICS
  - contacts with e.g. EGI, GÉANT, GAR, RENATER, surfNET
- Data storage:
  - essential: reproducibility and usability of all scientific results over full time of experiment (+ 10 years after shut-down)
  - parallel storage of low- and high-level data at CC-Lyon and CNAF (long-term commitments, pledged resources)
  - central services like software repository, central software builds
  - additional resources available (e.g. 500TB HDD disk space at ReCaS via INFN)
- Data processing by specialized service group
- Data access: via WAN/GRID access tools (xrootd, iRODS and gridFTP)

Open Data Policy

- central goal: prompt dissemination of scientific results, new methods and implementations
- public access
  - summary data (event information plus quality information) after fixed latency (typically 2 years)
    - defined in MoU
  - web based downloads of data and software
    - harmonize with GNN, ASTERICS (as well as for other eCommons)
- on request
  - more (detailed) data, earlier releases, etc.
- observer in KM3NeT collaboration (free of charge)
  - access to all data, meetings, etc. (but no voting rights)
    - co-author if he/she contributed to publication

eCommons in Neutrino Astronomy

- Integration of IceCube data?
  - Yes, under umbrella of

    - joint publications
      - 1st ANTARES/IceCube paper to be published
      - book on neutrino astronomy in preparation (World Scientific)
    - exchange of sky maps
    - exchange of detector performance data
    - discussions on costs
    - collaboration between working groups from both experiments (including yearly workshop – MANTS)
- Integration of data from other neutrino experiments?
  - published data
    - neutrino oscillation data (values + uncertainties) used in combined fits
  - example Memoranda of Understandings of ANTARES
    - real-time alerts (e.g. TaToo)
    - pipe-lined data streams (e.g. Virgo)

Training Services

- Interface to VO, including training tools and training programs
- Training services for scientists and students
  - central software boot camps, etc.)
  - cooperation with (Astro-)Particle Physics Masterclasses (COSMICS)
- disseminate data from detector operation (acoustics, environment monitoring) – incl. documentation and training courses
  ⇒ outside community

Questions from the ESFRI QNR – personal comments

- 4.1. What will be the data management and open data policy of the RI?
  - clear question, answer see above
- 4.2. What is the plan for supporting advanced data management and how will it be funded?
  - clear question, answer: committed partners
- 4.3. What is needed from external e-infrastructure services?
  - does not fully apply as there is large project-internal support – see 4.2
- 4.4. Will the RI contribute to the development of e-infrastructure commons in the field or in general?
  - certainly, eCommons highly important
- 4.5 Will the RI policy on data include training services for “data practitioners” to enable the effective use of data repositories and data analysis tools by non-scientists?
  - certainly for scientists, why also for non-scientists?

Summary

- KM3NeT currently in Phase-1
- Data management plan and computing model established
  ⇒ sustainable also for the next project phases
- KM3NeT collaboration
  - actively uses eCommons
    (taken over from ANTARES and other experiments)
  - contributes to their further development
- Discussion, feedback and advice is always welcome!

Backup Slides

Computing Needs for Phase 1 (per year)

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	85	-	85	-	1
Monitoring and Minimum Bias Data	43	-	43	-	1
Experimental Data Processing					
Calibration (incl. Raw Data)	213	3 M	425	6 M	2
Reconstructed Data	43	15 M	85	31 M	2
DST	21	4 M	43	8 M	2
Simulation Data Processing					
Air showers	10	14 M	10	14 M	1
atm. Muons	5	1 M	5	1 M	1
neutrinos	2	22 k	20	220 k	10
total:	**208**	**37 M**	**290**	**60 M**	

Computing Needs for 1 Building Block (per year)

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					
Calibration (incl. Raw Data)	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**	

KM3NeT Preliminary