Neutrinotelescopes
Antares & KM3NeT @ CC-IN2P3

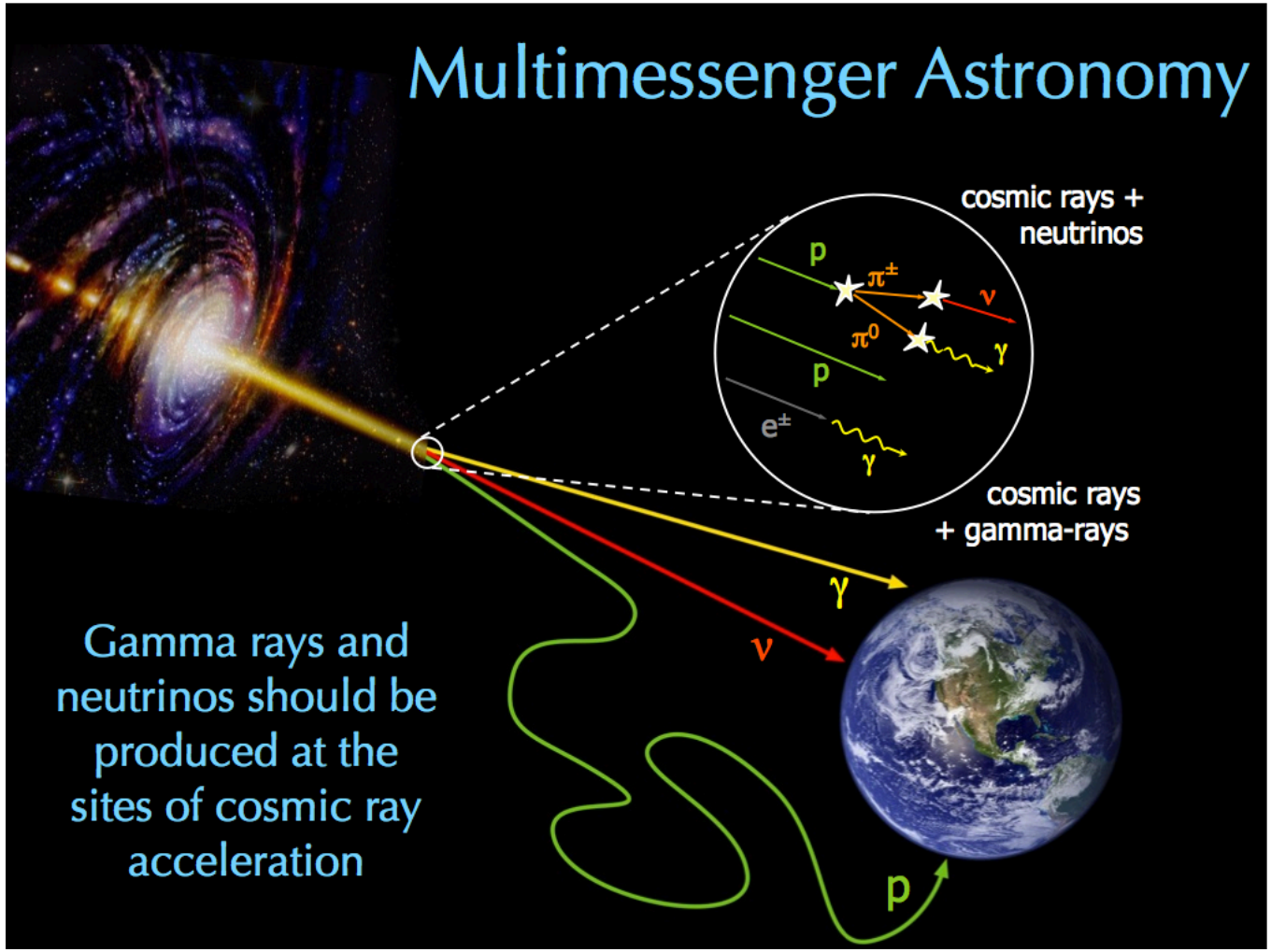
Kay Graf, Jürgen Brunner
CC-IN2P3 User Days

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

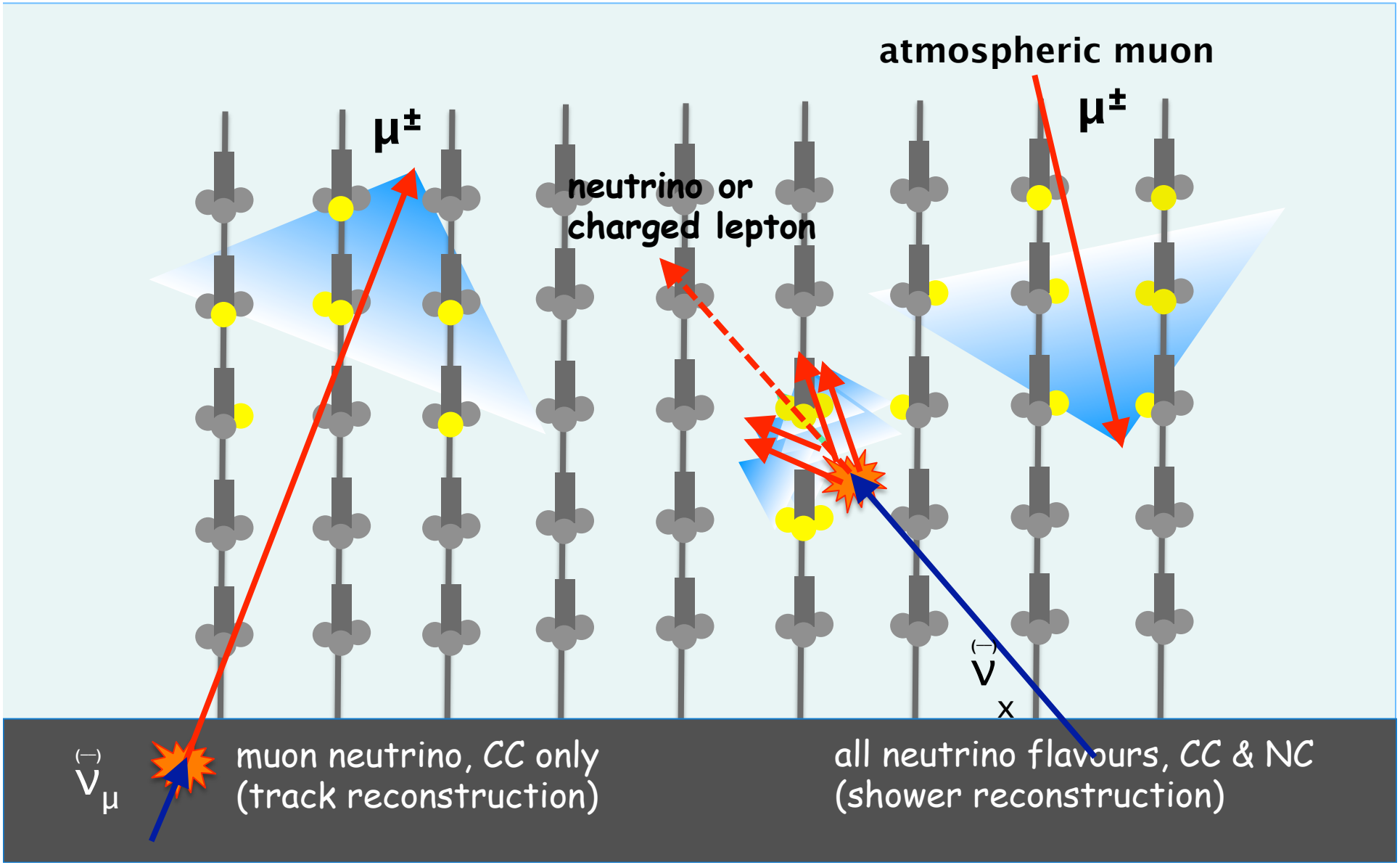
FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

Neutrino Telescopes

- Main goal : Find Acceleration Sites of Cosmic Rays

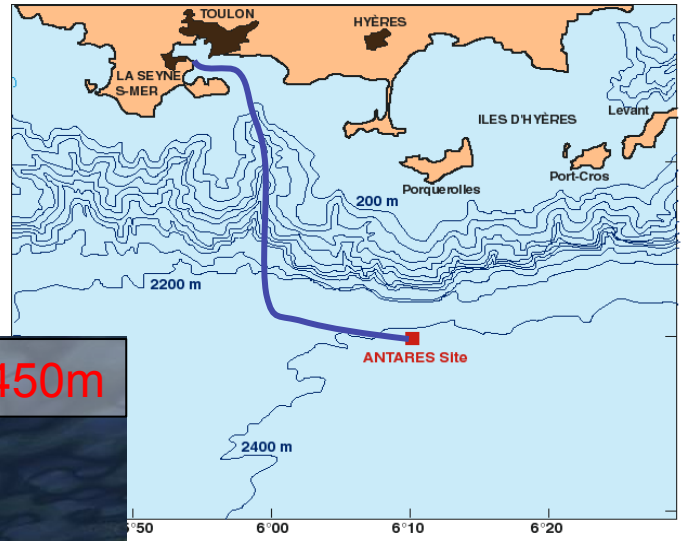


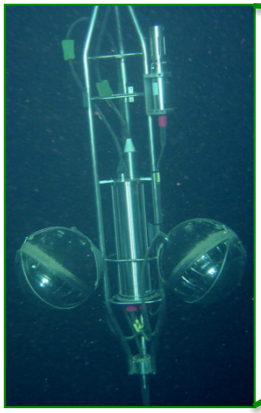
Principles of neutrino detection

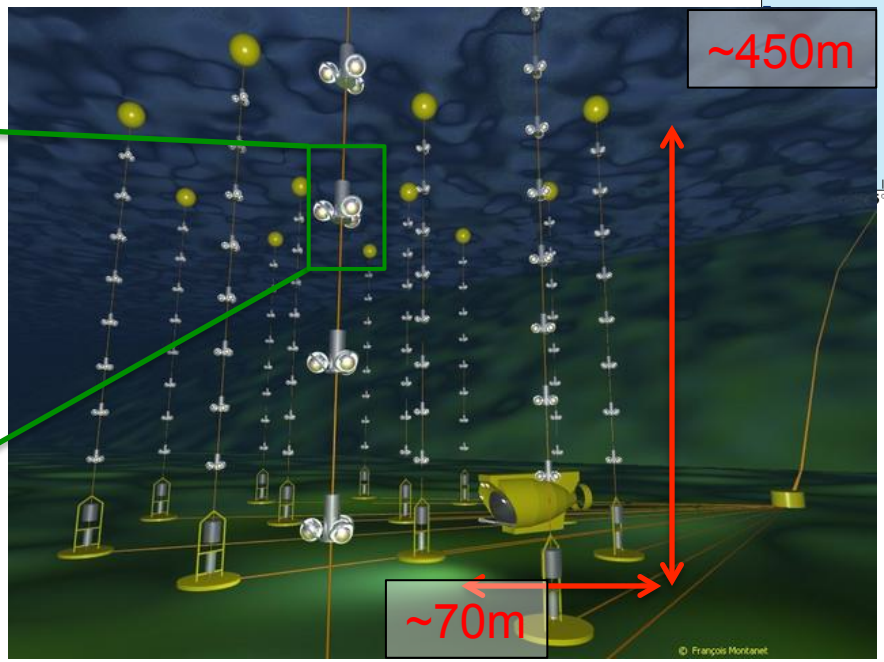


The ANTARES detector

- 12 lines, ~70m spacing
- 25 storeys per line
- 3 x 10-inch PMTs per storey
- Construction 2006-2008
- Still data taking







KM3NeT

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

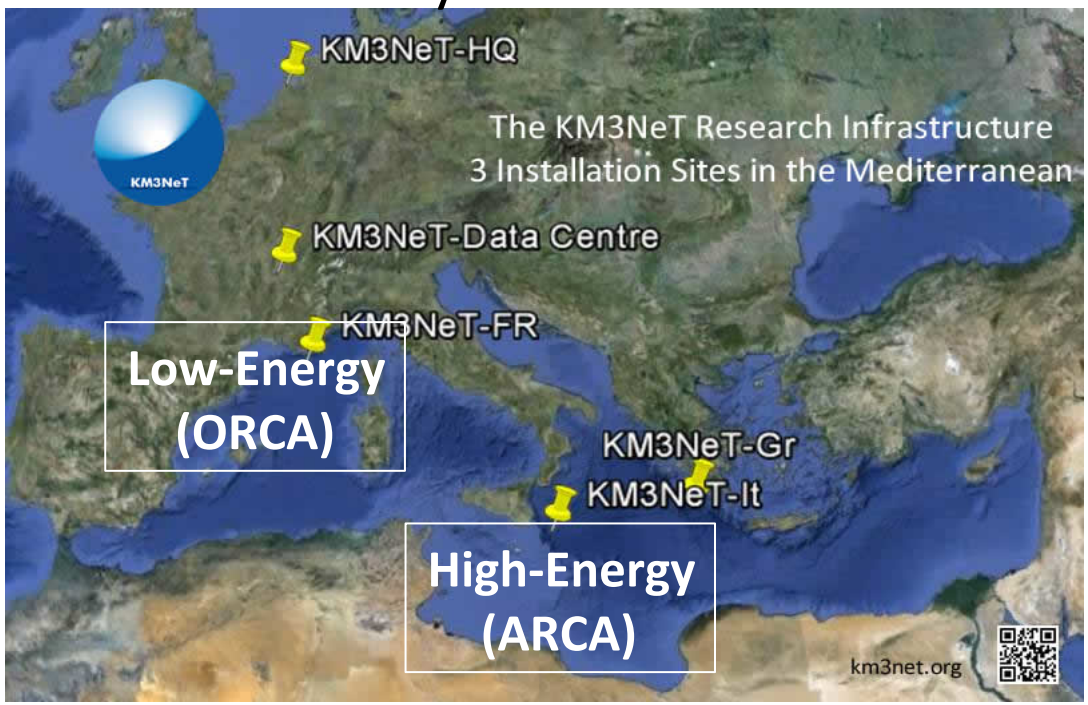
Research infrastructure with 2 main physics topics:

- The origin of cosmic neutrinos (high energy)
- neutrino physics (low energy)
- Deep Sea Observatory

Single Collaboration

Single Technology

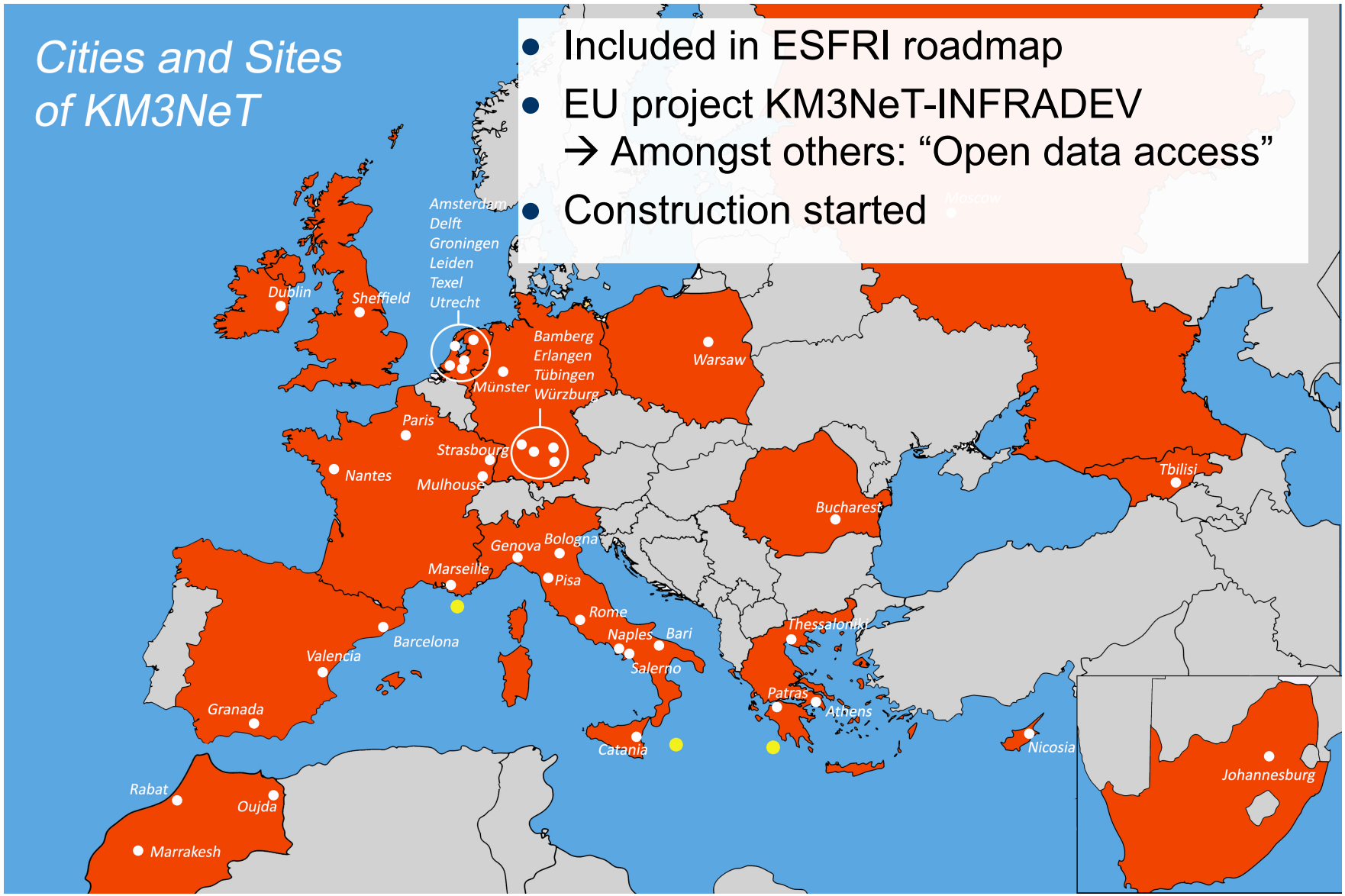
Single Management



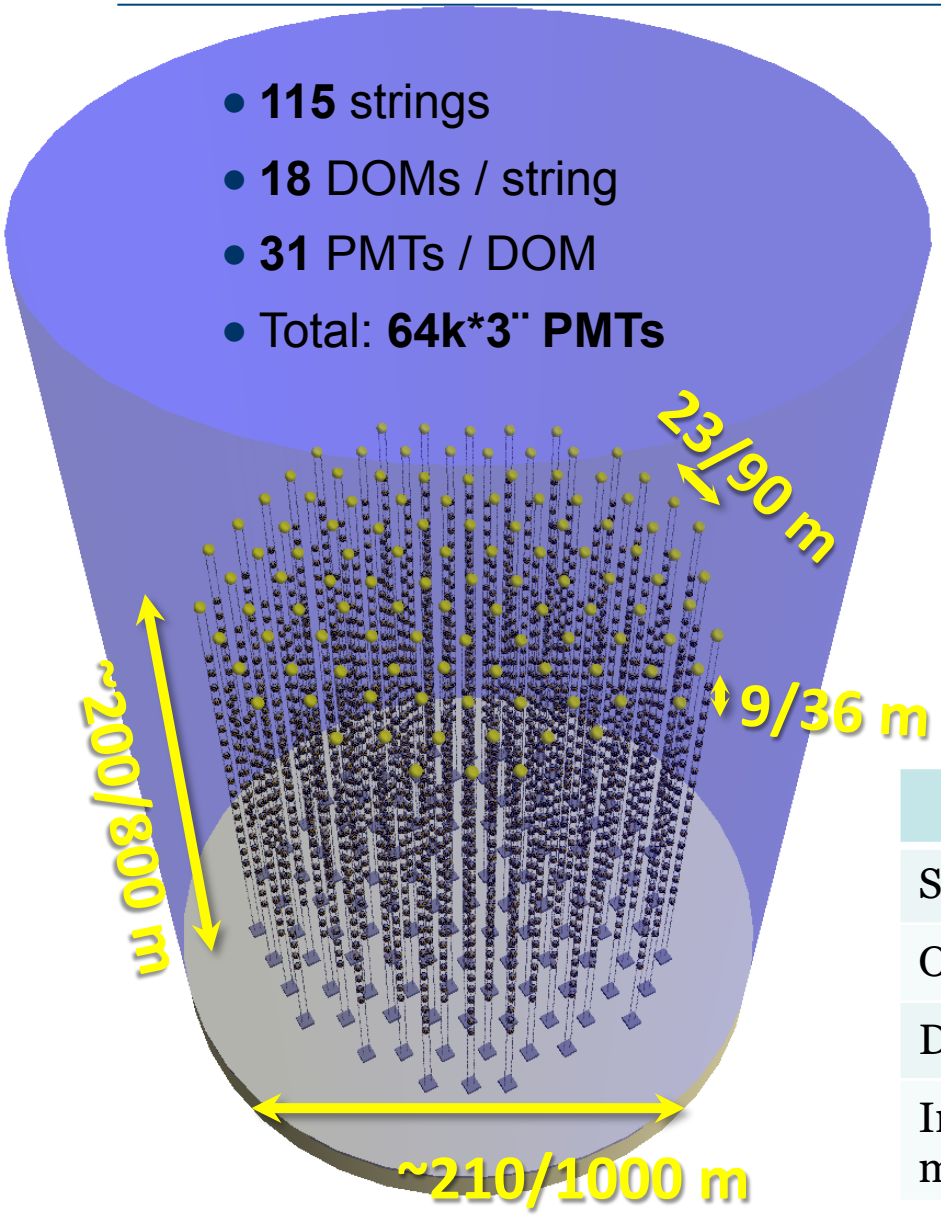
ARCA- Astroparticle Research with Cosmics in the Abyss

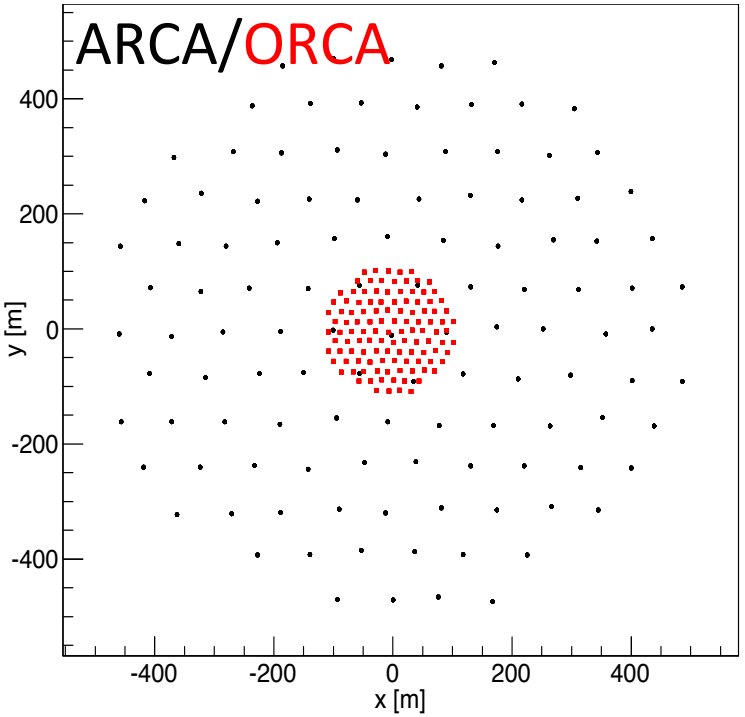
ORCA- Oscillation Research with Cosmics in the Abyss

The KM3NeT Collaboration



KM3NeT Building Block





	ORCA	ARCA
String spacing	23 m	90 m
OM spacing	9 m	36 m
Depth	2470 m	3500 m
Instrumented mass	8 Mton	0.6*2 Gton

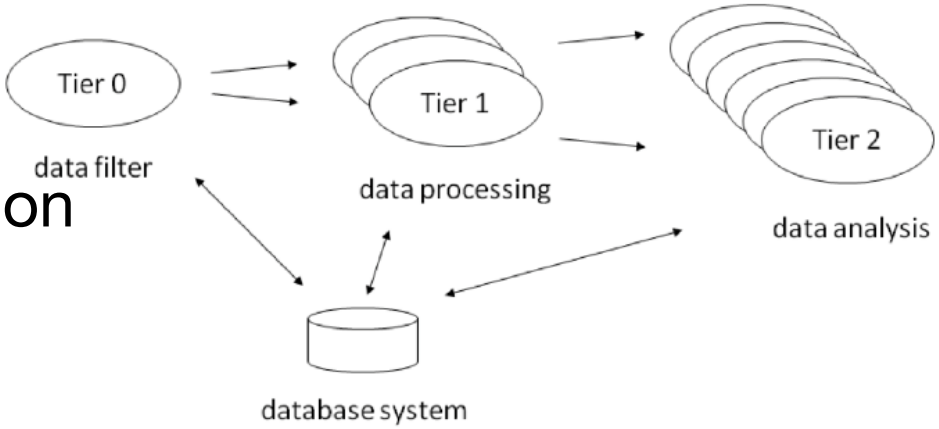
KM3NeT detector and data

Data rate

- Altogether: 3 building blocks
- 192500 PMTs, 5-10 kHz single-photon rate each,
  all data to shore ➔ some GB/second

Data processing

- Online data filter (on shore)
- Calibration and reconstruction
- Data analysis



+ Monte Carlo
+ Online analysis for alerts etc.
+ Earth and sea sciences

Requires advanced data management

The data management plan

- Computing model
- “FAIR” data:
  - findable
  - accessible
  - interoperable
  - re-usable

KM3NeT INFRADEV – H2020 – 739560

KM3NeT Data Management Plan	
KM3NeT-INFRADEV GA DELIVERABLE: D4.1	
Document identifier:	KM3NeT-InfraDev-WP4-D4.1_v1.4
Date:	22 June 2017
Work package:	WP4
Lead partner:	FAU
Document status:	Endorsed by PMB and KM3NeT IB
Dissemination level:	Public

Computing Model: General Scheme

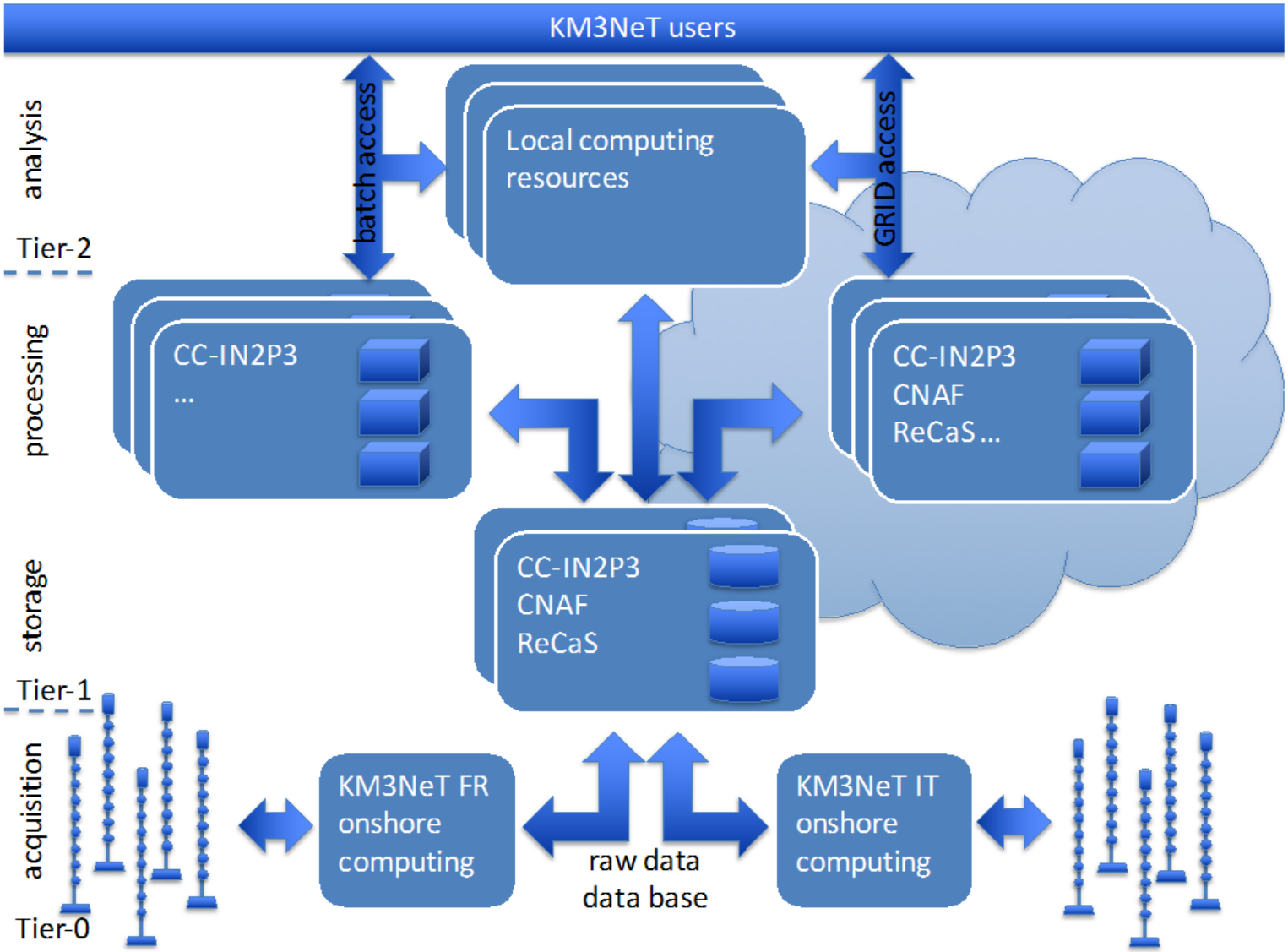
ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- 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

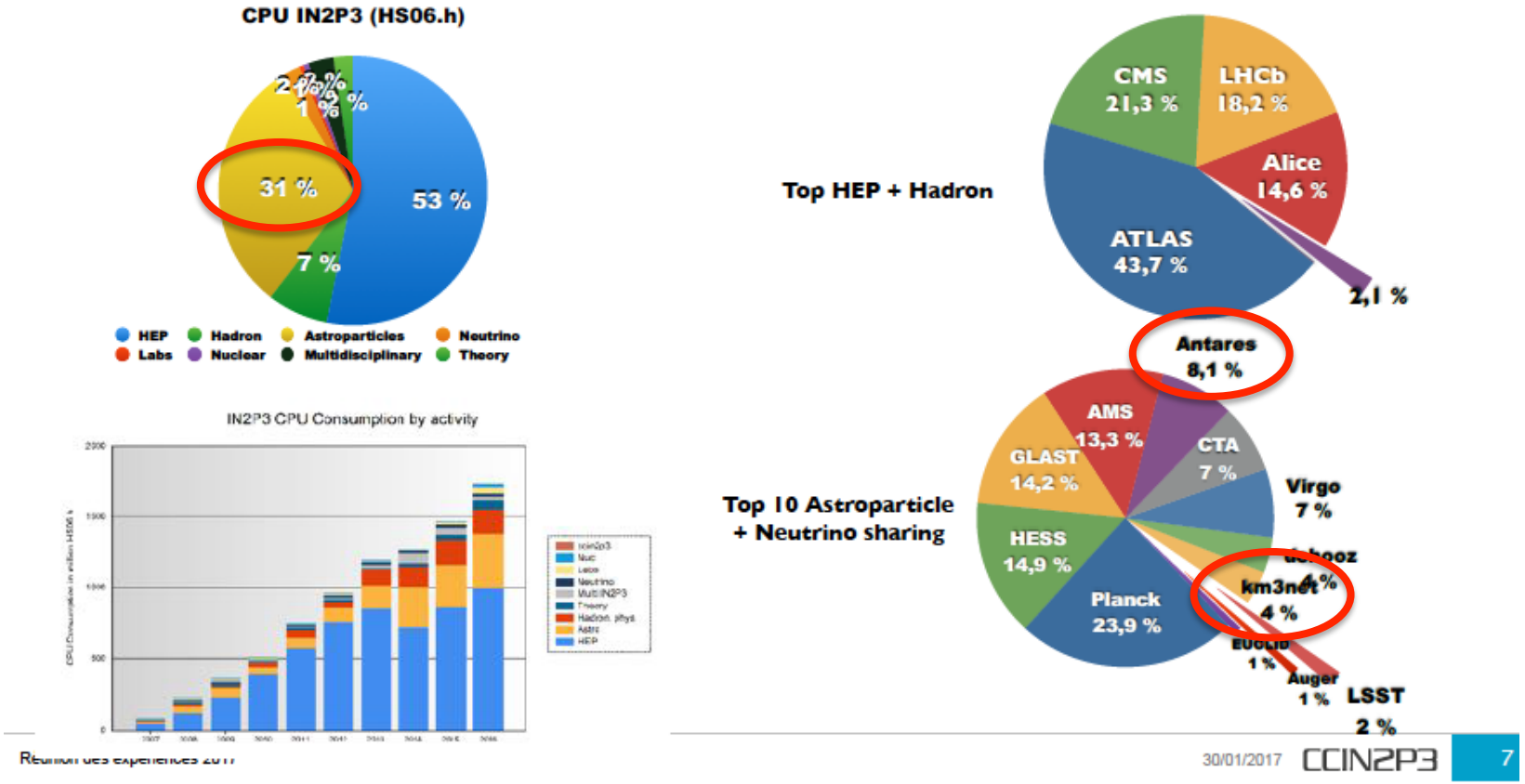
Computing model



Ressource usage @ CC

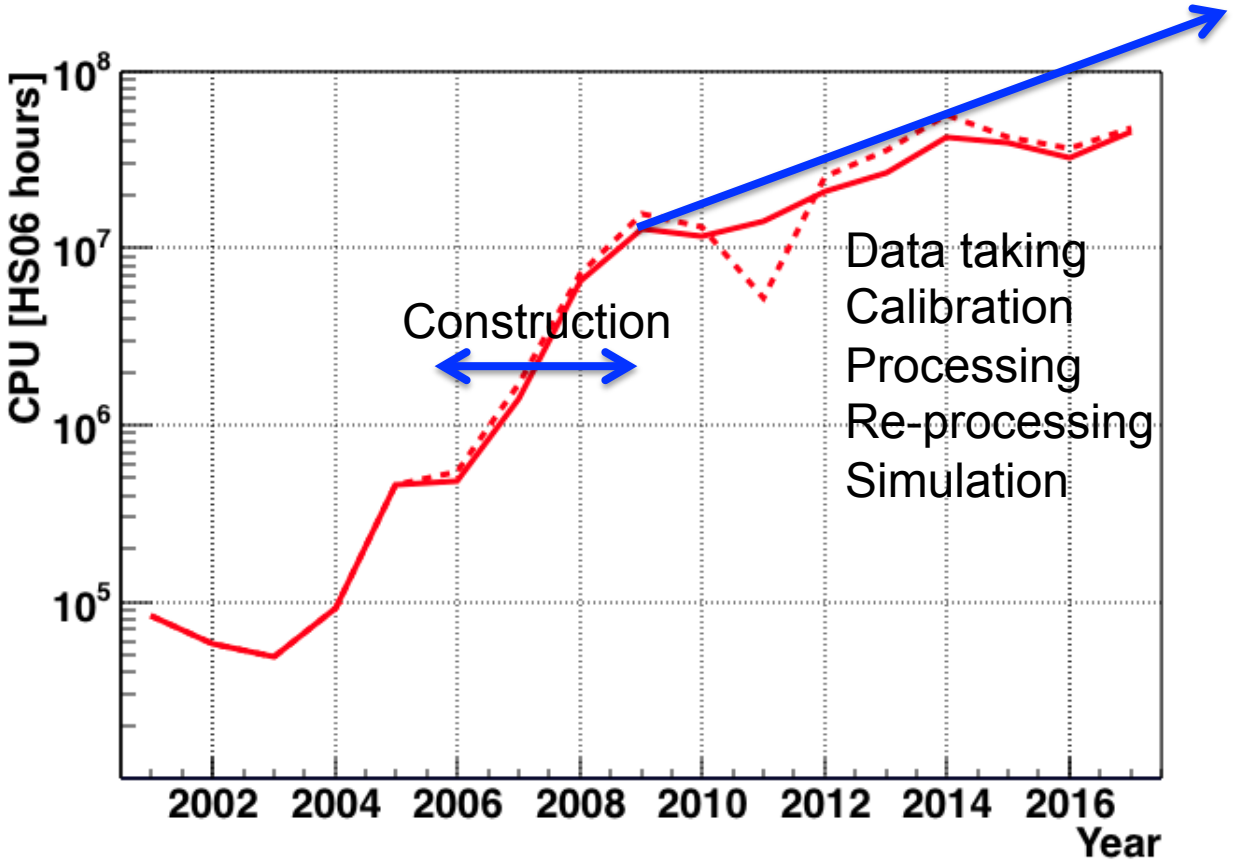
Values for 2016 from last years user meeting

- CPU IN2P3 Usage



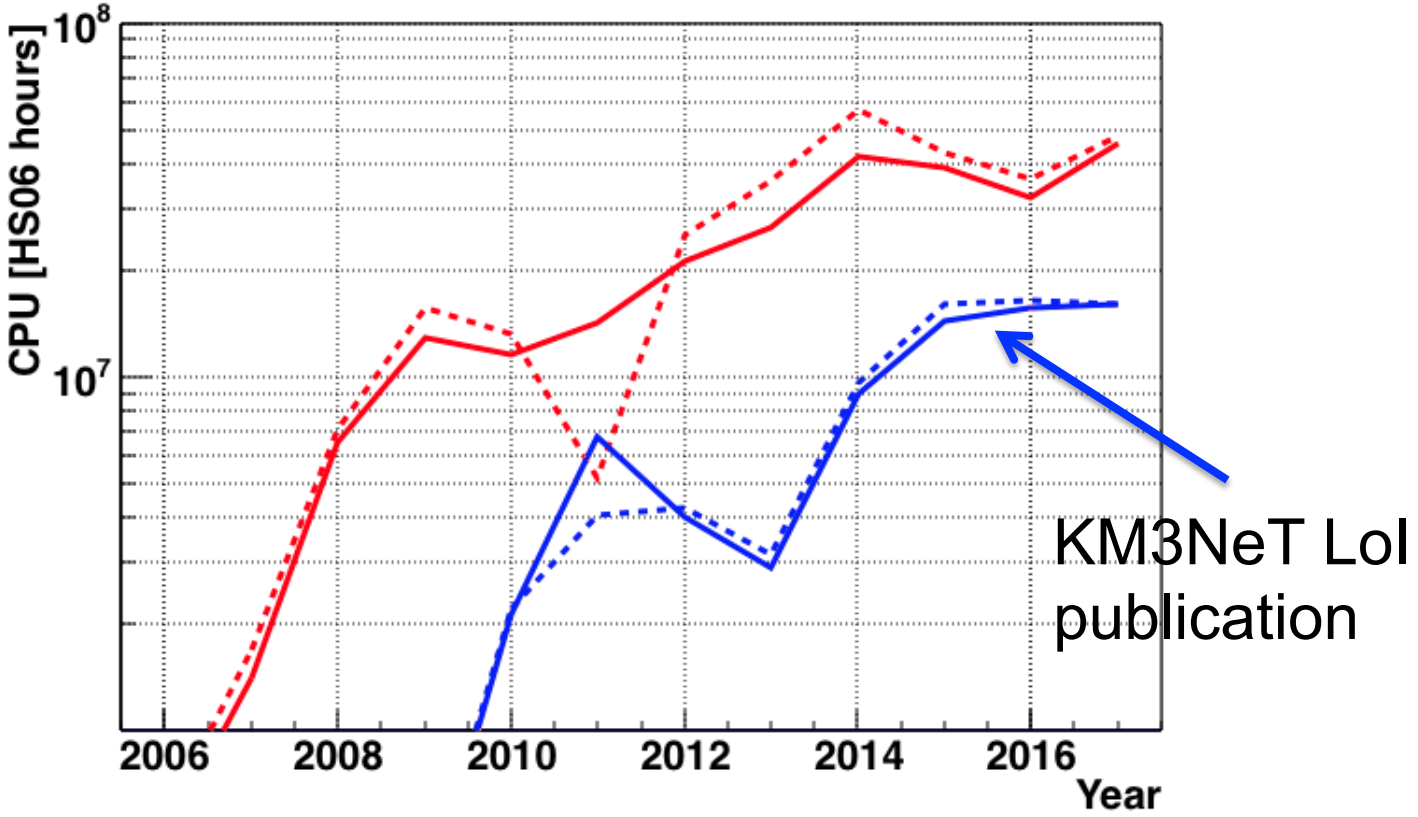
CPU Consumption

- Antares : probably one of oldest “clients” of the CC



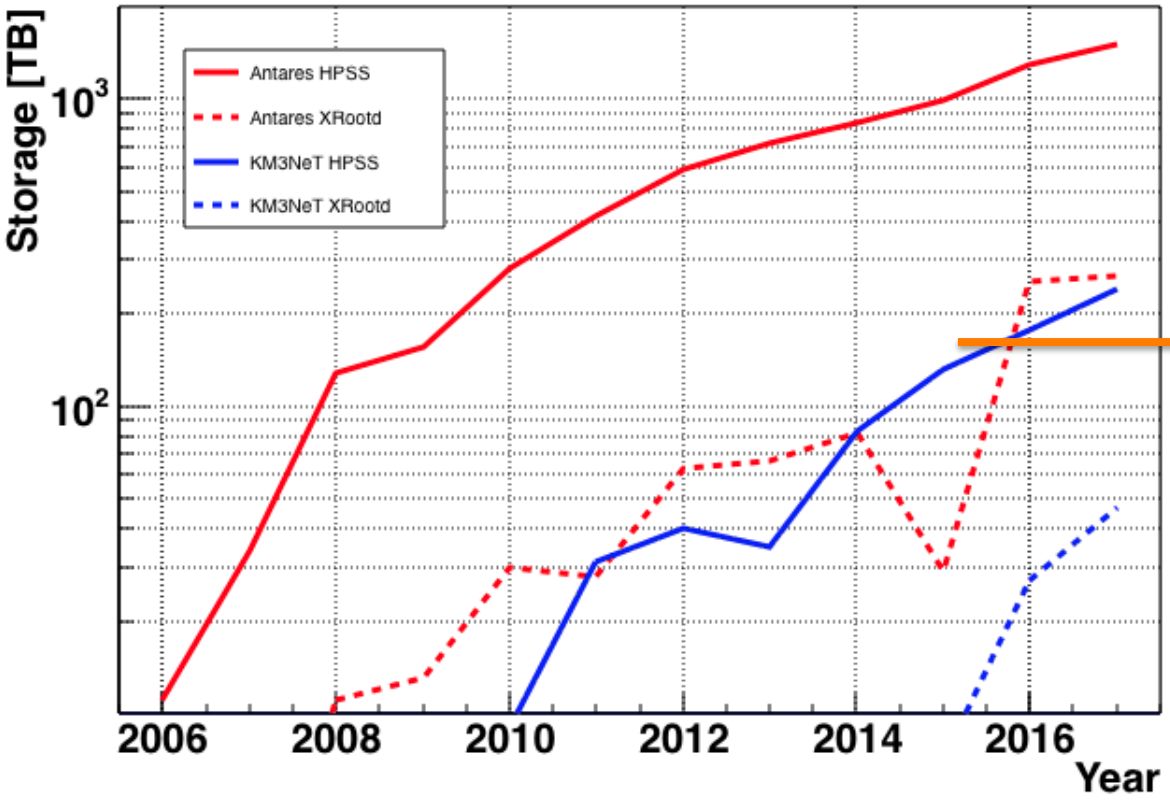
CPU Consumption

- Antares versus KM3NeT : only factor ~3 difference
- KM3NeT construction just starting : strong increase expected



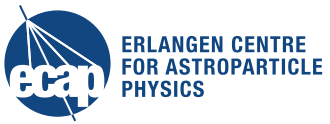
Storage : HPSS / XRootD

- Antares versus KM3NeT : factor ~10 difference
- Every two years : amount of stored data doubles



Semiper
Very important for fast analysis
Shared between Antares & KM3Net

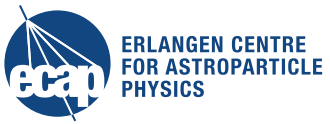
1 building block, 1 year



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

Data Management Plan



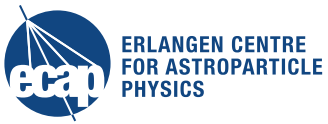
- Data formats and meta-data following common practice (root, xml, ascii) → allow for integration in eCommons
- Use of existing eCommons (e.g. from ANTARES) and specific development via ASTERICS
- Data storage:
  - essential: reproducibility of all scientific results and data usability over full time of experiment (+ at least 10 years after end of operation)
  - 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
- Data access: via WAN/GRID access tools
  - (xrootd, iRODS and gridFTP)

Open-access data policy

- Public access to data
  - Summary data (event information plus quality information) after fixed latency (typically 2 years)
  - Web-based downloads of data and softwares
  - Also: Simulation data
- On request: More (detailed) data, earlier releases, etc.
- Harmonise with wider community (→ GNN and ASTERICS),
  use common tools/platforms (→ Virtual Observatory, VO),
  link to other eCommons

GNN = Global Neutrino Network;
ASTERICS = Astronomy ESFRI and Research Infrastructure Cluster;

Oracle Database



- Informations by Cristiano Bozza

- Used for Metadata, slow control, monitoring both Antares & KM3Net
- Current use in KM3NeT
  - 700 GB detector monitoring
  - 300 GB Photomultiplier tests
- Estimated extrapolation
  - 5-10 TB per year for full KM3NeT detector

- Recently issues with access speed
  - Critical for data taking as CC-Oracle DB is accessed directly from shore station at La seyne and even form Sicily

KM3NeT challenges and ideas

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

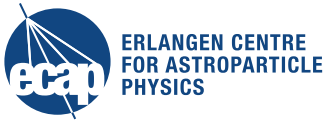
- Centralised Tier-1 services (data access, data analysis) provided by CC-IN2P3 and CNAF (and GRID infrastructure)
- Data processing on batch farm and GRID
- typically I/O limited → best service (iRods, xrootd, rftools) change over time, guidance from experts welcome
  - data synchronisation between two computing centres is a challenge
  - Direct access to storage resources via gridFTP possible?
- Data storage
  - ORACLE DB → had performance issues
  - SPS – imprtant for individual analyses, difficult to keep “clean”

KM3NeT challenges and ideas II

ERLANGEN CENTRE FOR ASTROPARTICLE PHYSICS

- Software
  - Change of OS always a challenge
  - Seems to become more and more complicated → why ?
  - Software/analysis preservation using singularity under development → solution or yet another level of complexity ?
- /afs phaseout issues
  - GROUP_DIR : smooth, no problem
  - THRONG_DIR
    - connection to web server
    - User pages, log-files, shift-reports, editable cgi-bin
  - @sys functionality : platform transparency apparently lost
- Platform for machine learning (DL) approaches would be interesting
- Support for public data access and data conservation (services under development, implementation not yet settled)
- Training sessions or online courses possible?

Summary



- KM3NeT construction started, detector with 3 blocks by ~2023
- Data management plan and computing model established
- KM3NeT adopts tier-model for data management
  (large data volumes, significant CPU requirements)

- Thanks to CC-IN2P3 for the excellent support!
  - Some ideas for improvements in the talk