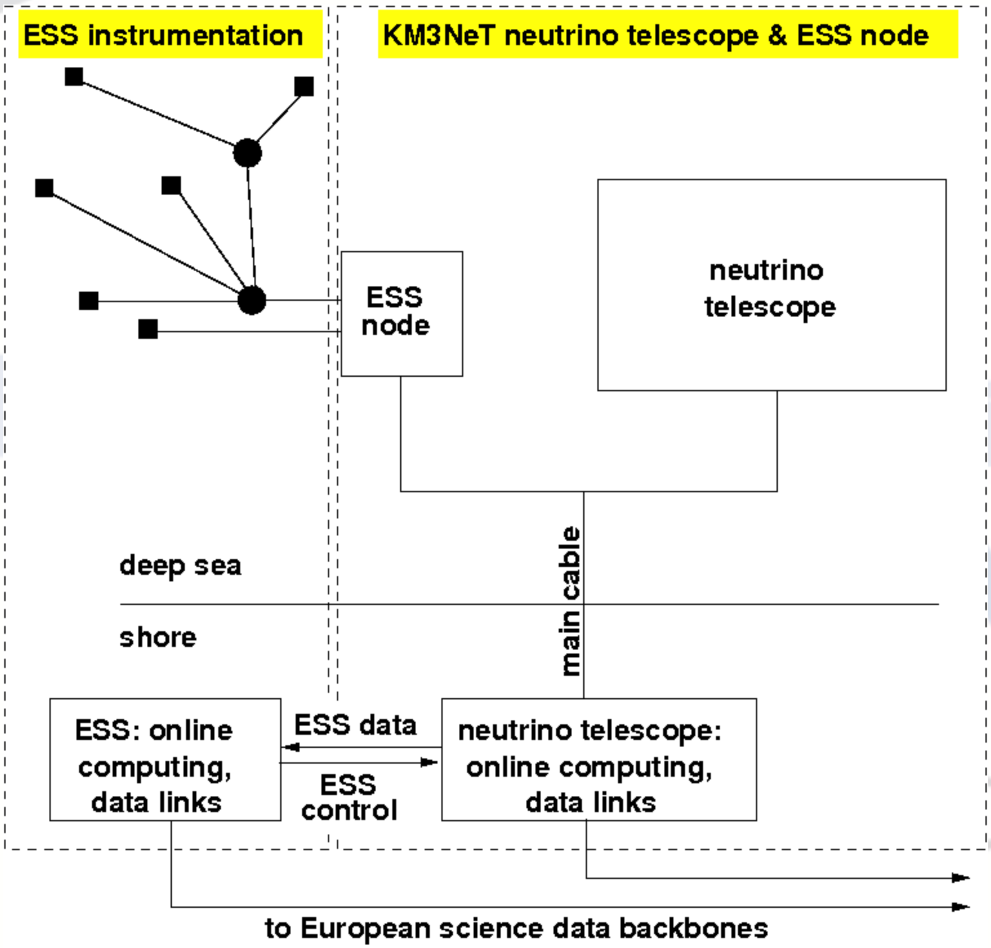
Il calcolo per KM3NeT

Agnese Martini
INFN LNF

The KM3NeT infrastructure (I)

- The KM3NeT research infrastructure will comprise a deep-sea neutrino telescope (ARCA) and the neutrino-mass-hierarchy detector ORCA at different sites and nodes for instrumentation for measurements of earth and sea science (ESS) communities.
- The cable(s) to shore and the shore infrastructure will be constructed and operated by the KM3NeT collaboration.

The KM3NeT infrastructure (II)



Both the detection units of ARCA/ORCA and the ESS nodes are connected via a deep-sea cable network to shore.

Note that KM3NeT will be constructed at multiple sites (France (FR), Italy (IT) and Greece (GR)) as a distributed infrastructure. The set-up shown in Figure will be installed at each of the sites.

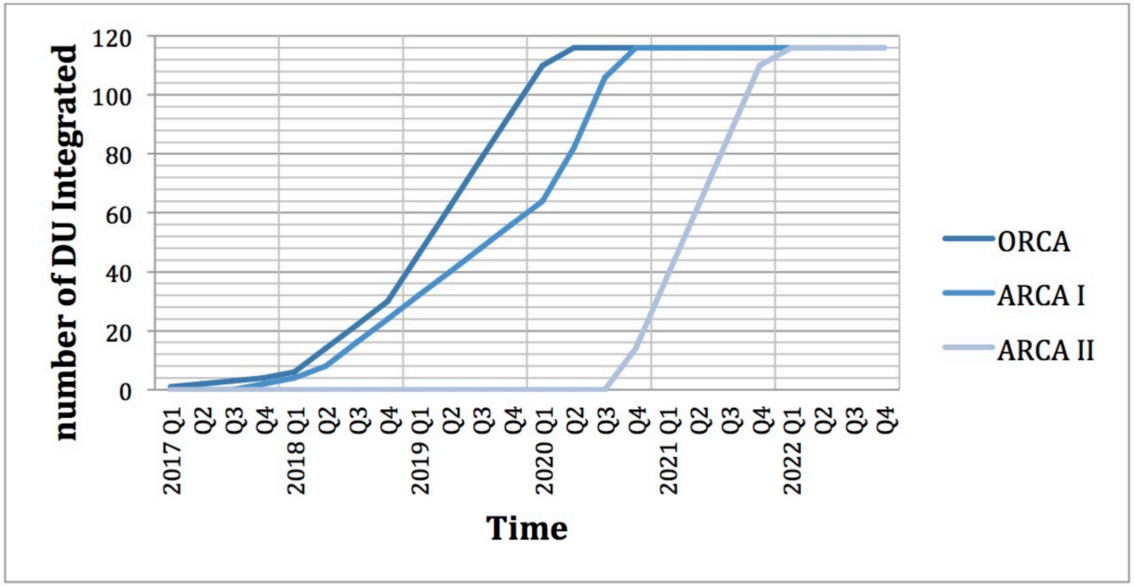
The detector

- The ARCA/ORCA will consist of building blocks (BB) containing 115 detection units (vertical structures supporting 18 optical modules each).
- Each optical module holds 31 3-inch photomultipliers together with readout electronics and instrumentation within a glass sphere. One building block thus contains approximately 65,000 photomultipliers.
- Each installation site will contain an integer number of building blocks.
- The data transmitted from the detector to the shore station include the PMT signals (time-over-threshold and timing), calibration and monitoring data.

Phase	Detector Layout	No. of DUs
Phase 1	approx. ¼ BB	31 DUs
Phase 2	2 BB ARCA/ 1 BB ORCA	345 DUs
Final Phase	6 building blocks	690 DUs
Reference	1 building block	115 DUs

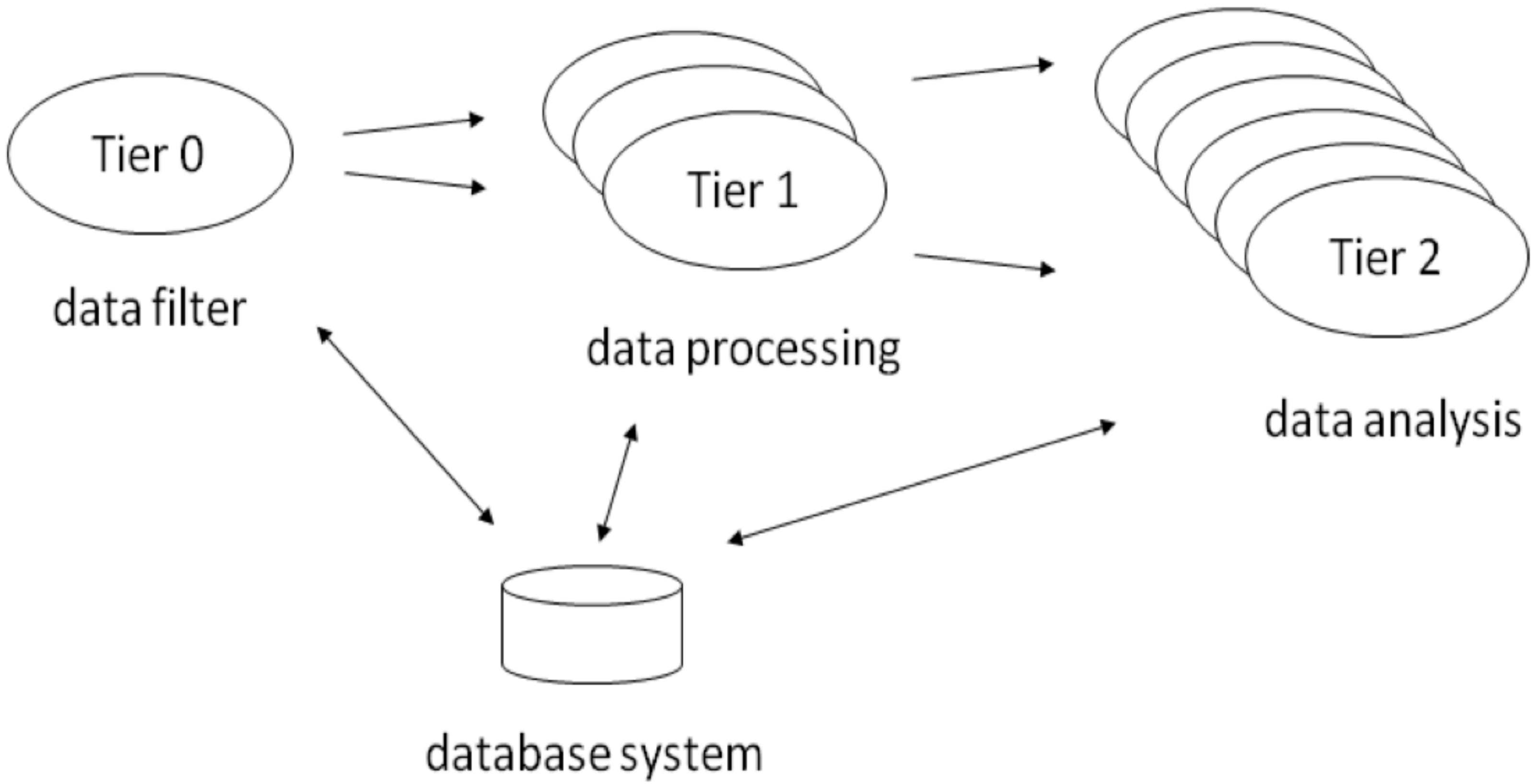
The detector status

- Status Italian Seafloor Network: down since 16th of April for power failure. After investigations The network will be powered off until a ROV inspection that we are investigating the possibilities
- Status of the French seafloor Network: A repair operation of the cable that connects the infrastructure is currently ongoing. The fault was located in a joint close to the node and the repair should be completed in the next few weeks.
- 2 DU are being prepared for deployment in France this summer.
- DOM/DU integration: The overall integration of DOM's and DU's is not affected by any of the network problems and will be pursued without delay.



The KM3NeT Computing Model

The KM3NeT computing model (data distribution and data processing system) is based on the LHC computing model. The general concept consists of a hierarchical data processing system, commonly referred to as Tier structure



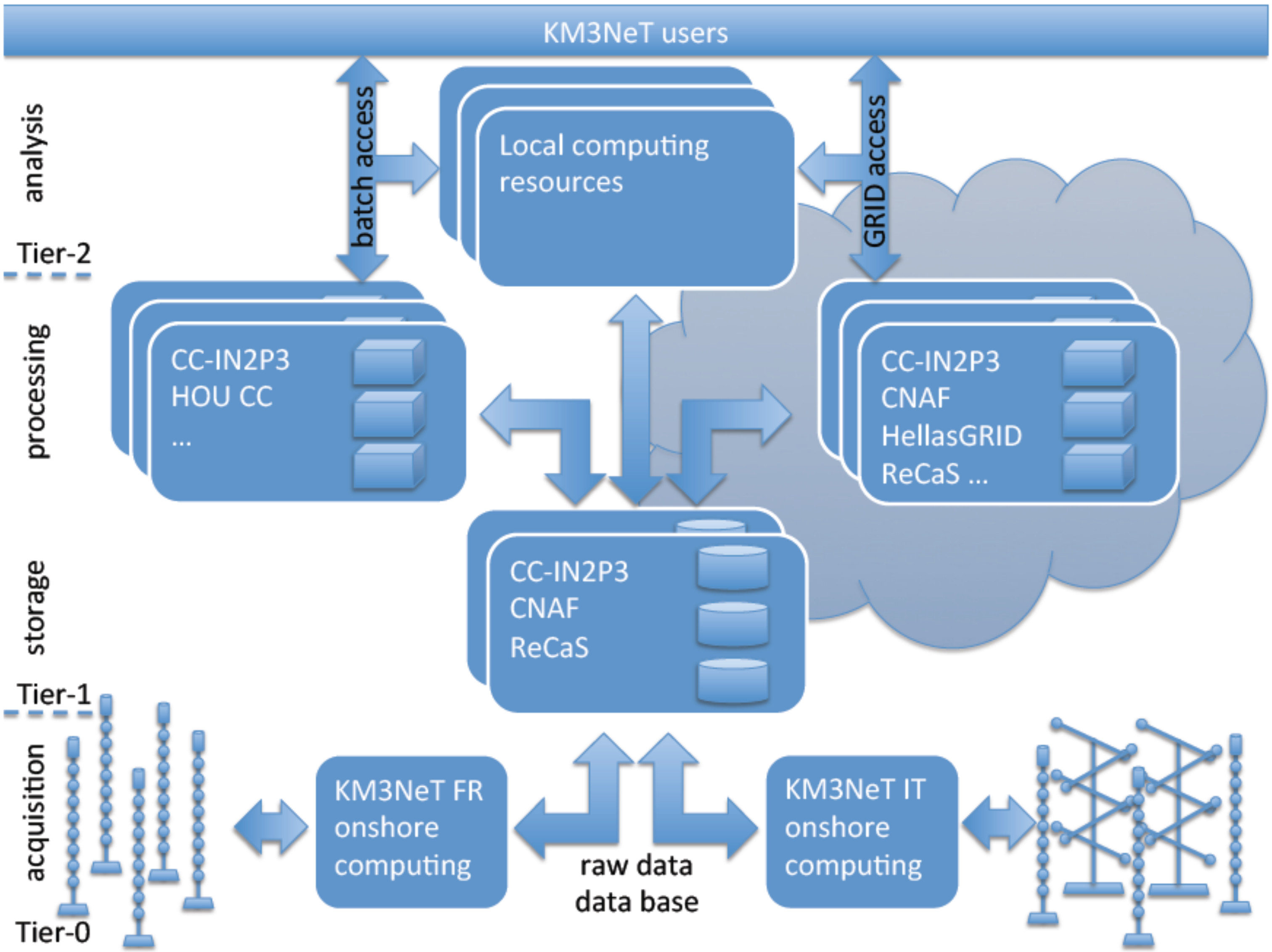
Data processing steps at the different tiers

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

Computing centres and pools provide resources for the KM3NeT

Tier	Computing Facility	Main Task	Access
Tier-0	at detector site	online processing	direct access, direct processing
Tier-1	CC-IN2P3	general offline processing and central data storage	direct access, batch processing and grid access
	CNAF	general offline processing and central data storage	grid access
	ReCaS	general offline processing, interim data storage	grid access
	HellasGrid	reconstruction of data	grid access
	HOU computing cluster	simulation processing	direct access, batch processing
Tier-2	local computing clusters	simulation and analysis	varying

Detailed Computing Model



Data distribution

- ➢ The raw data must be transferred from Tier0s (detector sites) to Tier1s: CC-IN2P3 and CNAF. The 2 storages must be mirrored.
- ➢ The data resulting from all the processing tasks have to be transferred to the Tier1s

The main task is to implement an efficient data distribution and mirroring system between these 2 centers that implement different data transfer protocol

- iRODS (Integrated Rule-Oriented Data System) at CC-IN2P3
- GridFTP (File Transfer Protocol for grid computing) at CNAF

Detector Deployment (DUs at beginning of the year)

year	Detector Layout	No. of DUs
2018	4 Dus ORCA 2 Dus ARCA	6 DUs
2019	0,3 building block ORCA 0,25 building block ARCA	65DUs
2020	1building block ORCA 1/2 building block ARCA	170 DUs
2021	1 building block ORCA 1,3 building block ARCA	260 DUs
2022	1 building block ORCA 2 building block ARCA	345 DUs

1 Building block = 115 DUs

Overview on computing requirements at T1s per year

per year	size (TB)	computing time (HS06.h)	computing resources (HS06)
One Building Block	1000	350 M	40 k
2018	250	25 M	3 k
2019	750	350 M	40 k
2020	500	150 M	20 k
2021	2000	700 M	80 k
2022	3000	700 M	80 k

KM3NeT
arca&orca

Detailed expectations of necessary storage and computing time for **one building block** (per processing and per year)

processing stage	size per proc. (TB)	time per proc. (HS06.h)	Size per period (TB)	time per year (HS06.h)	periodicity (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. Myons	50	1 M	25	638 k	0.5
neutrinos	2	22 k	20	220 k	10
total:	**827**	**188 M**	**995**	**353 M**	

First building block deployed in 2 years

Networking

phase	connection	average data transfer (MB/s)	peak data transfer (MB/s)
Building Block:			
	Tier-0 to Tier-1	25	125
	Tier-1 to Tier-1	50	500
	Tier-1 to Tier-2	10	50
Final Phase:			
	Tier-0 to Tier-1	200	1000
	Tier-1 to Tier-1	500	5000
	Tier-1 to Tier-2	100	500

Rough estimate of the required bandwidth.
Note that the connection from Tier-1 to Tier-2 has the largest fluctuation, driven by the analyses of data (i.e. by users)

KM3NeT on the GRID
VO Central Services

Service	Site
Authentication/authorization system VOMS	RECAS-NAPOLI
User Interface	RECAS-NAPOLI, HellasGrid-Okeanos, CNAF, Frascati
Logical File Catalog	RECAS-NAPOLI
Job submission and management system (WMS)	HellasGrid-Afroditi

- KM3NeT is starting on the GRID
- use case: CORSIKA simulation

GRID sites supporting the VO KM3NeT

Site name	Storage (TB)	CPU (not pledged)
HG-03-AUTH		440
HG-08-Okeanos	-	120
INFN-BARI		654
INFN-FRASCATI		2016
INFN-T1	200	25504
RECAS-NAPOLI	320	1500
UNINA-EGEE		64

arca&orca

HS06.h the last year

Tier1 — Sum CPU Work HS06 Hours by Tier 1 Node and Month (Custom VOs)

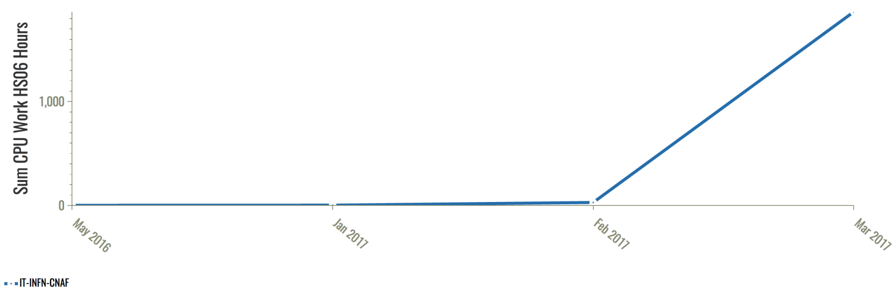
Tier 1 Node	May 2016	Jan 2017	Feb 2017	Mar 2017	Total	Percent
IT-INFN-CNAF	0	2	28	1,865	1,895	100%
Total		2	28	1,865	1,895	
Percent	0.00%	0.11%	1.48%	98.42%		

1 - 1 of 1 results ‹ 1 › Number of rows per page 30

Download JSON Data / Download CSV Data

The information in the previous table is also shown in the following graph.

Sum CPU Work HS06 Hours by Tier 1 Node and Month



HS06.h at Tier1 CNAF

Tier2 — Sum CPU Work HS06 Hours by Country and Month (Custom VOs)

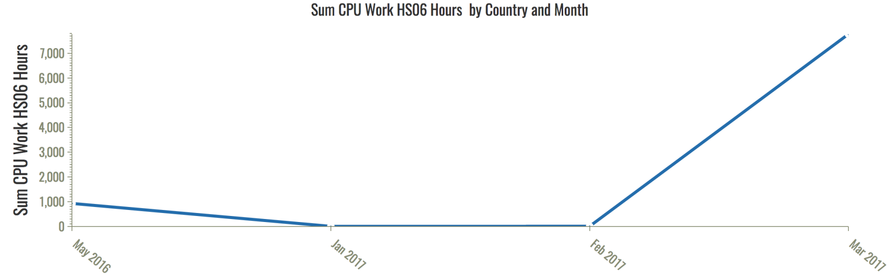
	May 2016	Jan 2017	Feb 2017	Mar 2017	Total	Percent
Italy	930	4	8	7,774	8,715	100%
Total	930	4	8	7,774	8,715	
Percent	10.67%	0.05%	0.09%	89.20%		

1 - 1 of 1 results ‹ 1 › Number of rows per page 30

Download JSON Data / Download CSV Data

The information in the previous table is also shown in the following graph.

Sum CPU Work HS06 Hours by Country and Month



HS06.h at Tier2s Italy

INFN for OBELICS 3.4

Activities in WP 3.4 (1/3)

- **CORELib: COsmic Ray Event Library**
  - Background to many experiments
  - Also a tuning benchmark
  - Potentially useful to other communities
  - Currently using CORSIKA as generator

- **Status of production**
  - **Proton-induced showers:**
    - *HEMODELS: QGSJET01 with CHARM, QGSJET01 with TAULEP, QGSJET-II with TAULEP, EPOSLHC with TAULEP*
    - *LEMODEL: GHEISHA*
    - *about 21M Evts per HEMODEL*
    - *7 energy bins (2×102GeV-103GeV+equally logarithmically spaced from 1TeV to 109GeV)*
    - *power-law spectrum with -2 spectral index*
    - *zenith angle from 0 to 89 degrees*

  - **Nuclei-induced showers:**
    - HEMODELS: QGSJET01 with CHARM, QGSJET01 with TAULEP, QGSJET-II with TAULEP, EPOS-LHC with TAULEP
    - LEMODEL: GHEISHA
    - about 21M Evts per HEMODEL
    - 7 energy bins (A×2×102GeV-A×103GeV+equally logarithmically spaced from A×1TeV to A×109GeV)
    - power-law spectrum with -2 spectral index
    - zenith angle from 0 to 89 degrees

INFN for OBELICS 3.4

Activities in WP 3.4 (1/3)

- **CORELib: COsmic Ray Event Library**

Status of production

Energy range (GeV)	Number of events
200-1000	10^7
10^3-10^4	10^7
10^4-10^5	10^6
10^5-10^6	10^5
10^6-10^7	10^4
10^7-10^8	10^3
10^8-10^9	10^2

Production done with and without Cherenkov radiation

High energy model	Low energy model	Option	
		TAULEP	CHARM
QGSJET01	GHEISHA		X
QGSJET01	GHEISHA	X	
QGSJETII-04	GHEISHA	X	
EPOS LHC	GHEISHA	X	

INFN for OBELICS 3.4

Activities in WP 3.4 (2/3)

- **ROAst: ROot extension for Astronomy**
  - Classes to access astronomical catalogues
  - Coordinate transformation
  - Moon position and motion
  - Generators of primary particles (neutrinos will be implemented, others will be supported only as placeholders)

Status

Catalogue name	Status
UCAC4	Supported
URAT1	Supported
GSC-II (Guide Star Catalog)	Supported
Fermi-LAT 3FGL	Supported
TeVCat	Supported

LUNAR MOTION: DONE

Astronomical coordinate system	Geographical coordinate system	Time coordinate
Equatorial	N/A	N/A
Galactic	N/A	N/A
Horizontal	Lat-Long/UTM	Unix time/UTC/Local Sidereal
Eclyptic rectangular	N/A	N/A

Summary

- The data distribution model of KM3NeT is based on the LHC computing model. The estimates of the required bandwidths and computing power are well within current standards. A high bandwidth Ethernet link to the shore station is necessary for data archival and remote operation of the infrastructure.
- KM3NeT-INFN already addressed requests to CCR
- KM3NET is already operative on GRID
- We are also active on Big Data future challenges, i.e. Asterics