Enabling Grid Computing resources within the KM3NeT computing model

Christos FILIPPIDIS
on behalf of the KM3NeT collaboration

The KM3NeT Project

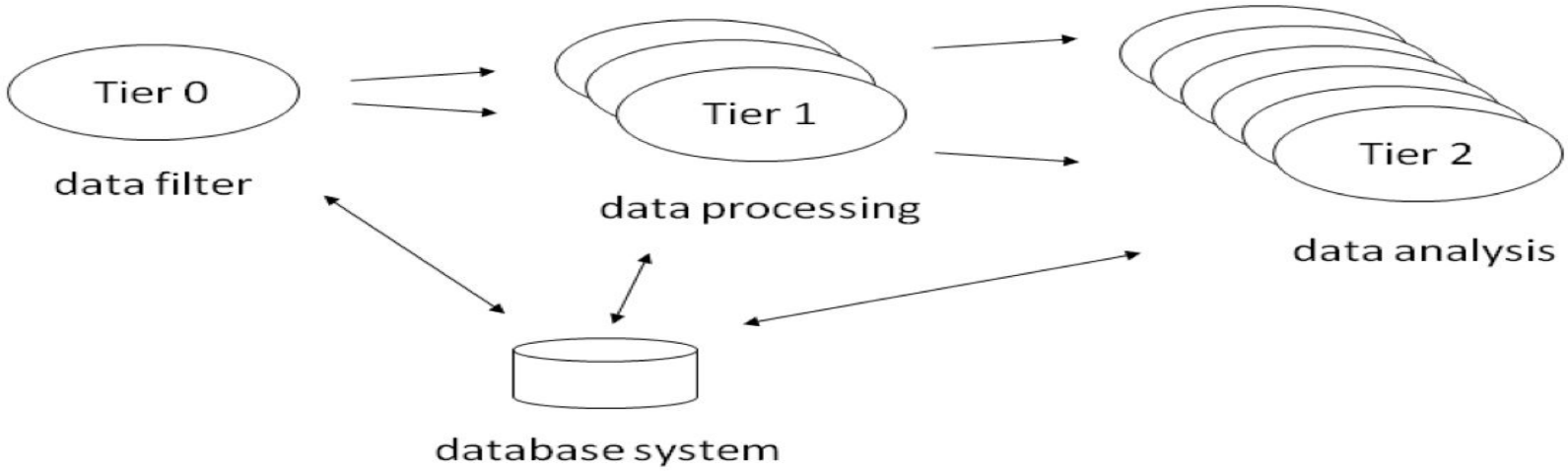
- KM3NeT is a future European deep-sea research infrastructure, which will host a distributed network of neutrino telescopes with a total volume of several cubic kilometers at the bottom of the Mediterranean Sea.

- KM3NeT is part of the ESFRI roadmap and will open a new window on the Universe.

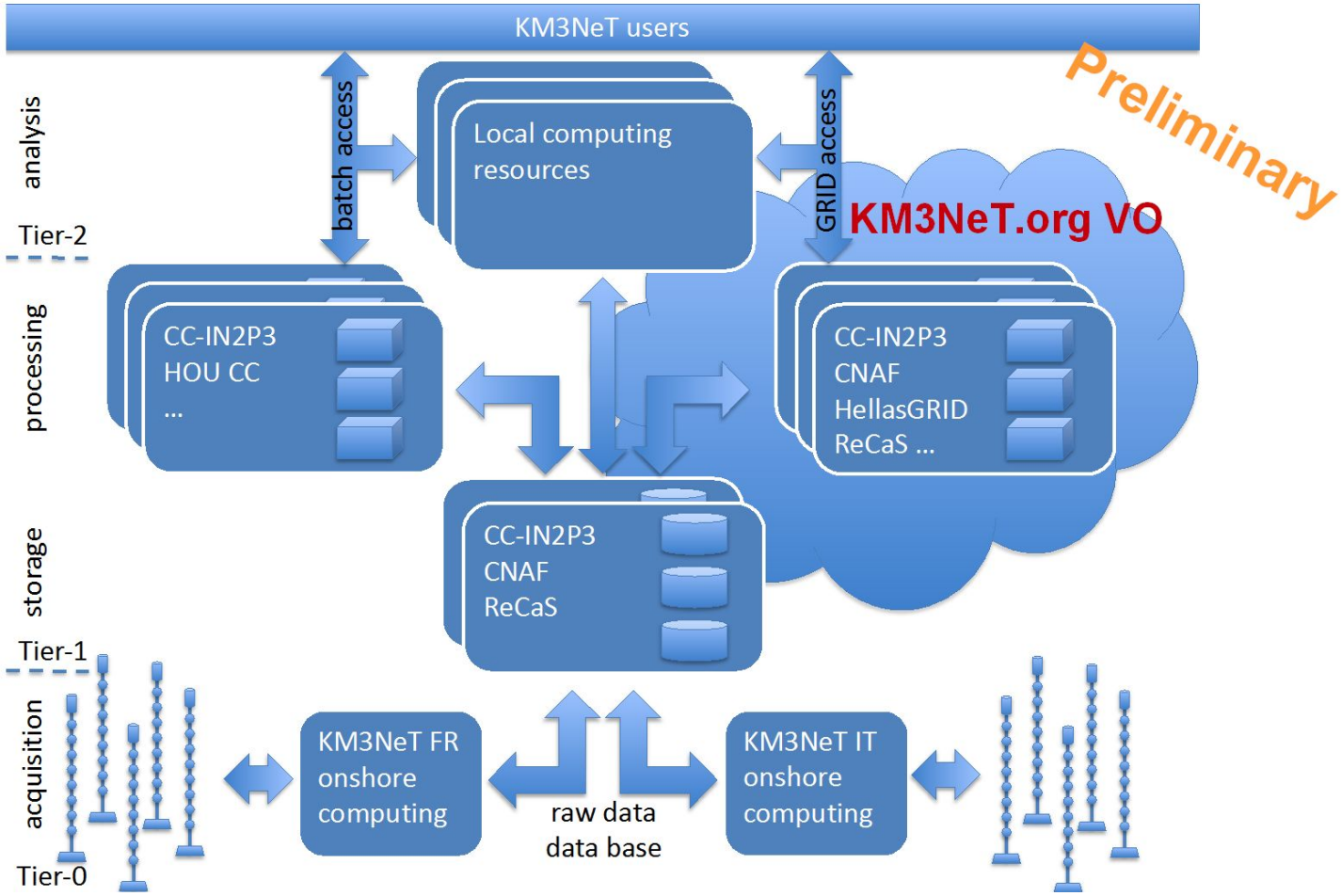
- The telescope will search for neutrinos from distant astrophysical sources like gamma-ray bursts, supernovae or colliding stars and will be a powerful tool in the search for dark matter in the Universe.

KM3NeT Computing Model

•The KM3NeT computing model is similar to the one of the LHC experiments with different levels of tiers (hierarchical data processing system).
•Serial to multi-parallel or GPU-optimized jobs.
•Very frequent WAN data transfers and data sharing among individuals, groups.



KM3NeT Computing Model Flowchart



Computing Requirement Estimates

Preliminary

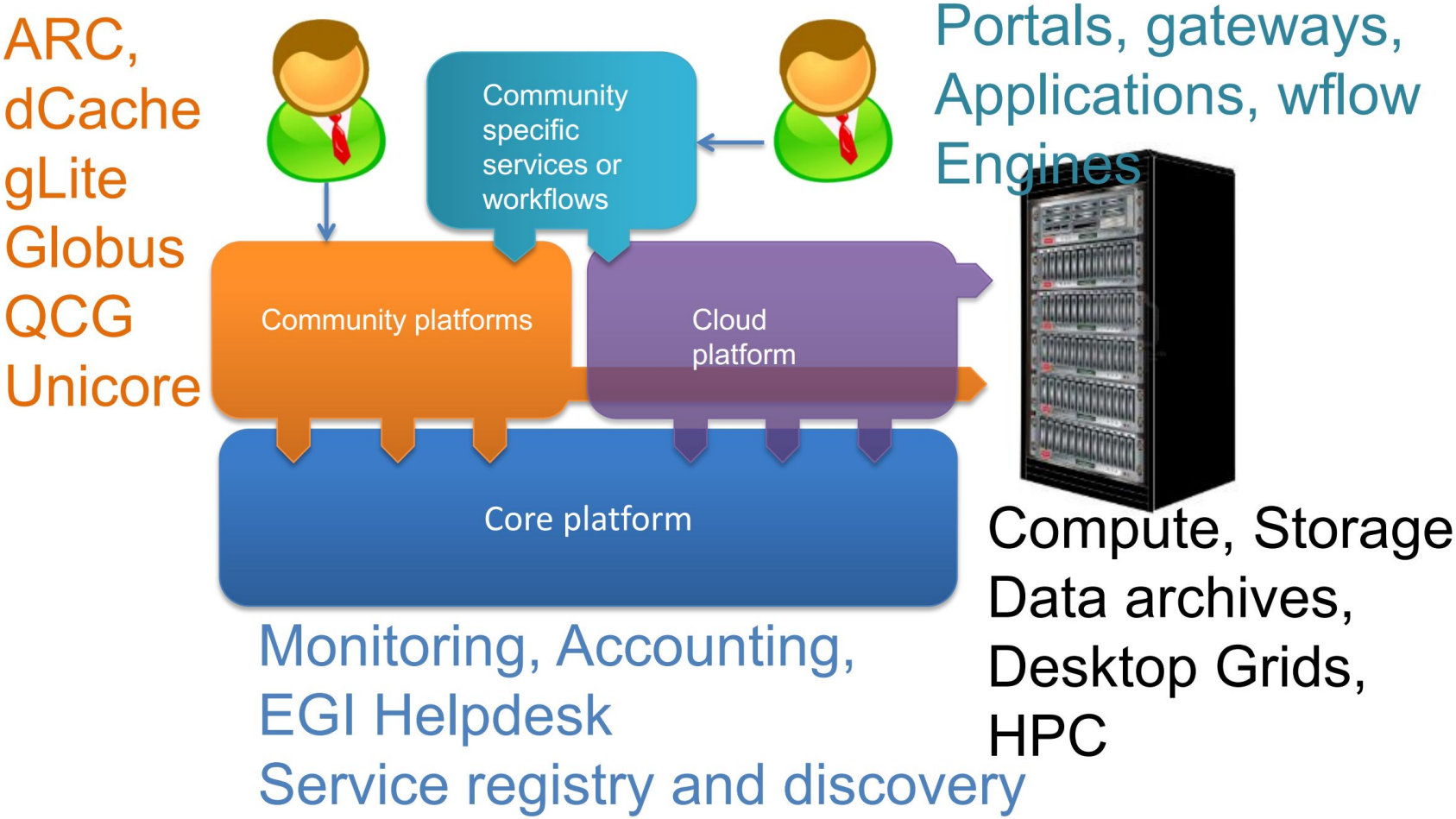
per year	size (TB)	computing time (HS06.h)	computing resources (HS06)
One Building Block	1000	350 M	40 k
Phase 1	300	60 M	7 k
- first year of operation	100	25 M	3 k
- second year of operation	150	40 M	5 k
Phase 2	2500	1 G	125 k
Final Phase	4000	2 G	250 k

Grid Computing Systems Definition

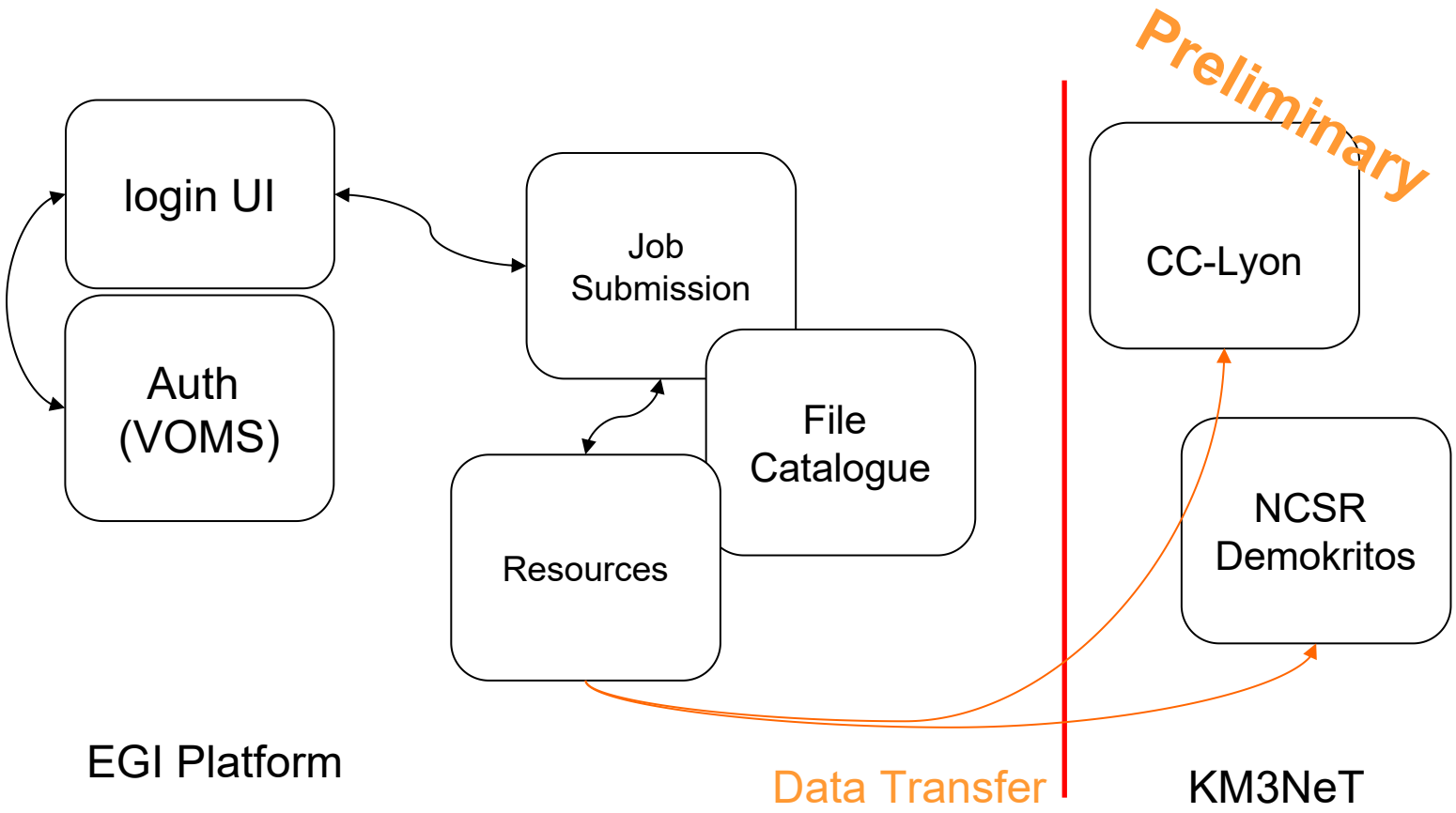
- Collection of geographically distributed heterogeneous resources.

"An infrastructure that enables flexible, secure coordinated resource sharing among dynamic collections of individuals, institutions and resources." - *Ian Foster and Carl Kesselman*

European Grid Infrastructure (EGI)



EGI-KM3NeT Workflow



KM3NeT.org VO Central Services

Service	Site
Authentication/authorization system (VOMS)	RECAS-NAPOLI
User Interface (UI)	RECAS-NAPOLI, HG-08-Okeanos
Logical File Catalogue (associate a logical file name to a physical file)	RECAS-NAPOLI
Job submission and management system (WMS)	HG-03-AUTH

KM3NeT.org VO Resources

Site Name	CPU slots (not pledged)	Storage Space (pledged)
INFN-BARI	3338	
INFN-LNF	1892	
UNINA-SCOPE	64	
UNINA-RECAS	1052	100 TB
HG-03-AUTH	232	
HG-08-Okeanos	55	
GR-12-TEIKAV	1152	

KM3NeT SW Installation Procedure

1. Upload source code and binaries @ the CERN Virtual Machine File System (CVFMS) repo: /cvmfs/km3net.egi.eu: Grid + Local access (CC-Lyon)
2. If we need to compile the SW localy:

- Submit a test job to each site in order to check the enviroment.
- Submit a job as a Software Manager (SGM), to each site.The job "copies" the sources from the CVFMS repo and compiles the SW under the /opt/soft_exp/km3net.org directory to each site (done once per site from the SGM)

Corsika 1% Production on the GRID

•Investigates correlation of atmospheric neutrino and cosmic ray flux.
•Used to simulate detailed muon distribution (test mupage) esp for muons at E>50 TeV.
•Produces muons and neutrinos at sea level.

•Submited: 1K jobs with 1000 showers each
•Time per job: ~ 12 – 20 h.
•Total time: ~ 40 h.
•Output per Job: 22 – 35 MB
•Total output ~27GB

# of jobs	Events	Real Time	Previous production	Output per job in MBs
			CPU days @ HOU	
20K	100k	3 weeks	~40k	100

KM3/MUPAGE on the GRID

•Mupage used to produce single muons and muon bundles at detection level.
•km3 implements MUSIC to propagate muons through the can, and calculates photon hits on PMTs.
•Current mupage/km3 production target is to test different ORCA spacings:
- km3 uses 0.5m track segments - O~1000 segments / muon
- tests hits for each PMT: O~50,000 tests (number / time of hits)
~ 50 million hit tests per event
~ O~10^6 muons/day: 5 x 10^13 tests / day of simulated data!
•Submited :3 test jobs, 100k events each (MUONs).
•output: 2 files 4MBs and 39 MBs each.
•Duration: 3.8-5.2 Days.

# of jobs	Events	Real Time	Previous production	Output per job in MBs
			CPU days @ Lyon	
7K	1000	3 weeks	~35k	300

Conclusions

•By integrating Grid Computing resources in the KM3NeT computing model we provide a flexible and secure way for users to reach more resources and fulfill the KM3NeT computing requirements.

•Our biggest challenge for the near future is to finalize the VO policy and develop data management tools and utilities that will create a smooth connection between the EGI platform and the KM3NeT computing model.