IHEP Experiments and Software Development

Xiaomei Zhang

IHEP Computing Center

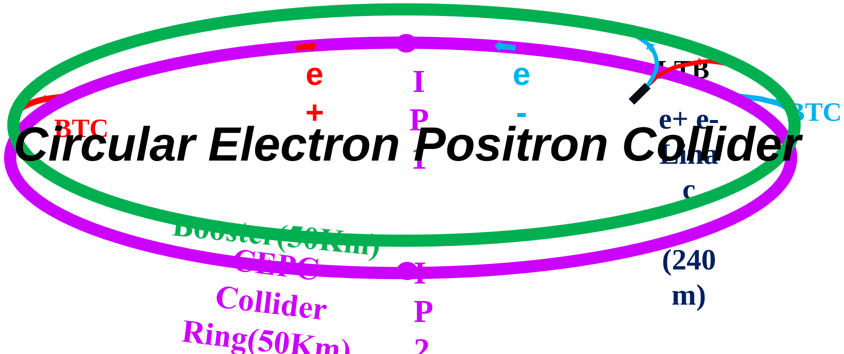
中国科学院高能物理研究所
Institute of High Energy Physics
Chinese Academy of Sciences

Outline

- Overview of IHEP experiments
- General Challenges
- Challenges and practice of software
  - JUNO Neutrino experiment
- Status and challenges of computing
- Summary

IHEP and IHEP experiments （1）

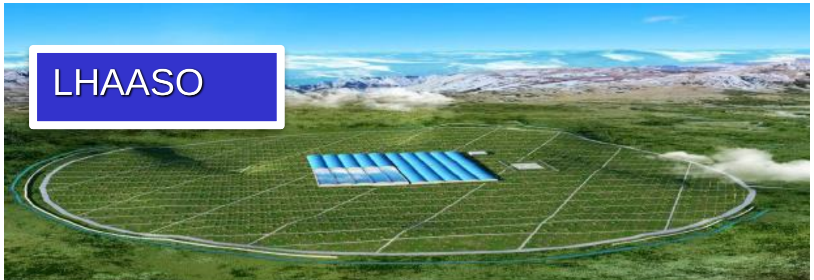
- Institute of High Energy Physics, CAS, Beijing, China
- The largest fundamental research center in China
  - Particle physics, AstroParticle physics, Accelerator and synchrotron radiation technology and applications
- The major institute conducts the high energy physics experiments in China
  - Collider :
    - BESIII – Study physics on tau-charm energy region
    - CEPC – Precision measurement of the Higgs/Z boson



IHEP and IHEP experiments （2）

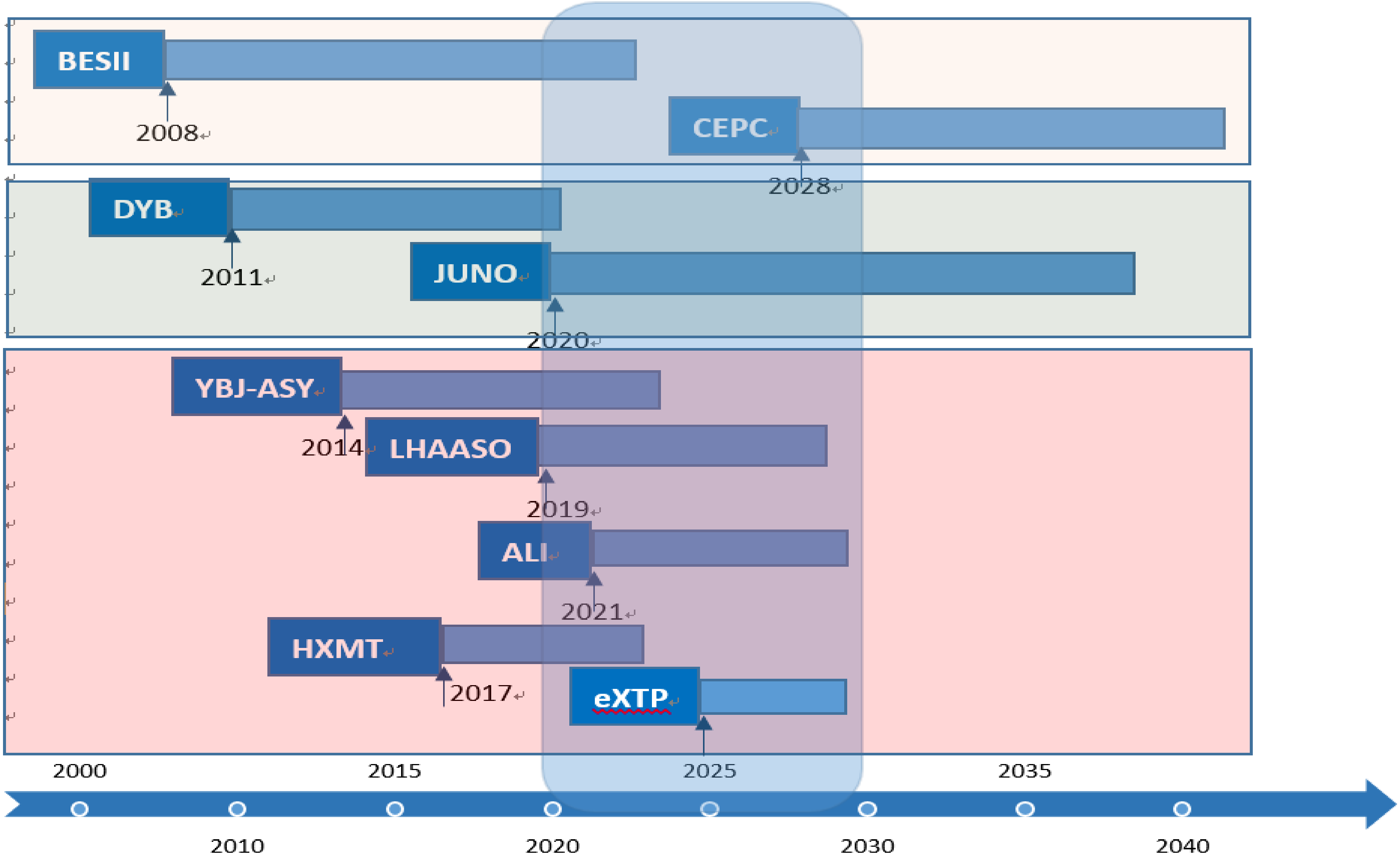
- Neutrino :
  - DayaBay – measurement of the missing angle
  - JUNO -- determine neutrino mass hierarchy and precisely measure oscillation parameters
- Astrophysics:
  - LHAASO -- detector array for high energy gamma ray and cosmic ray detection （after YBJ)
  - Ali -- Observation of Gravitational Waves
  - HXMT/eXTP -- X-ray astronomy satellite for high energy astrophysical space observations

Hard X-Ray Moderate Telescope



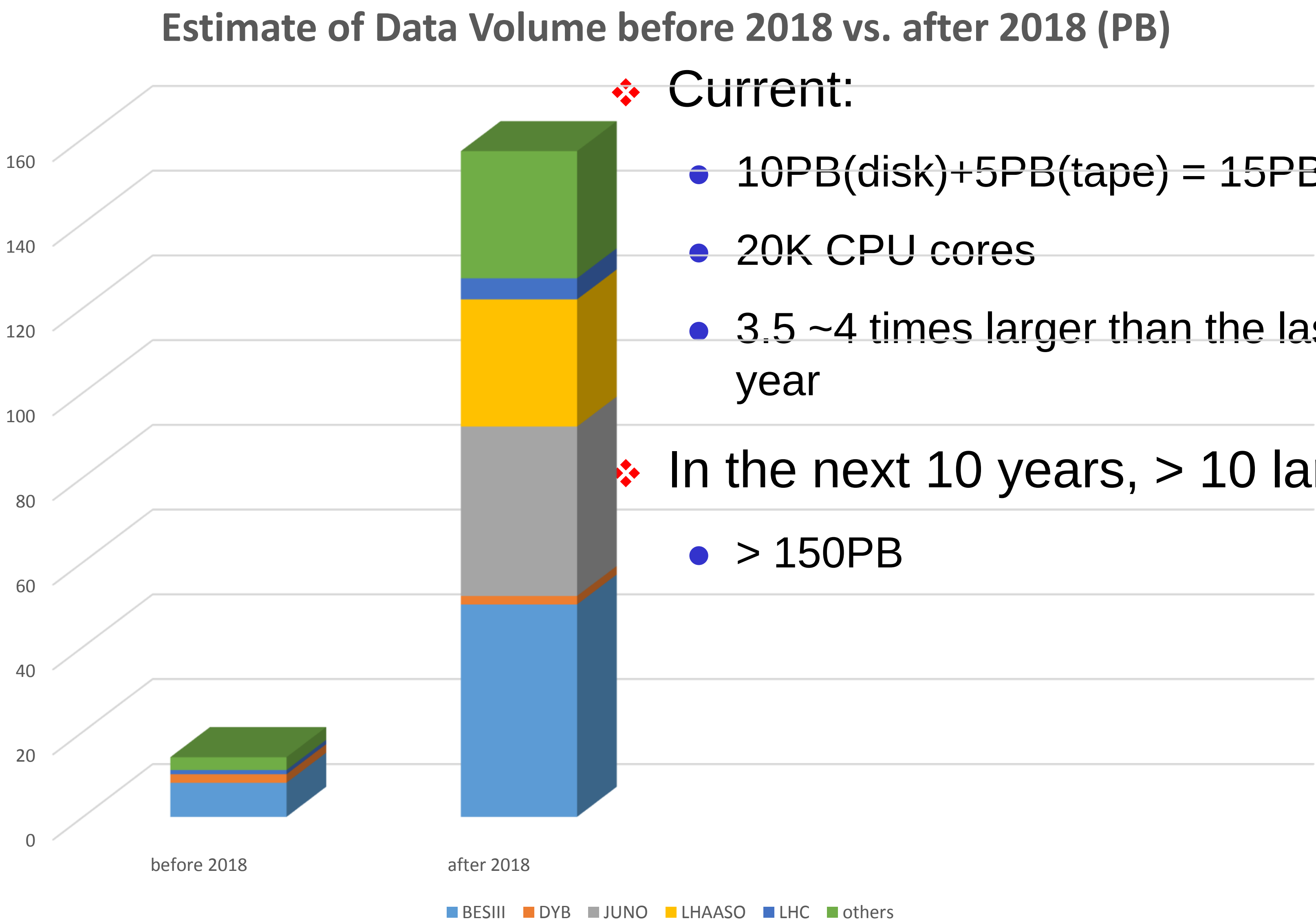
Large High Altitude Air Shower Observatory

Timeline



BESII
2008
CEPC
2028
DYB
2011
JUNO
2020
YBJ-ASY
2014
LHAASO
2019
ALI
2021
HXMT
2017
eXTP
2025
2000
2015
2025
2035
2010
2020
2030
2040

Data Volume (roughly)

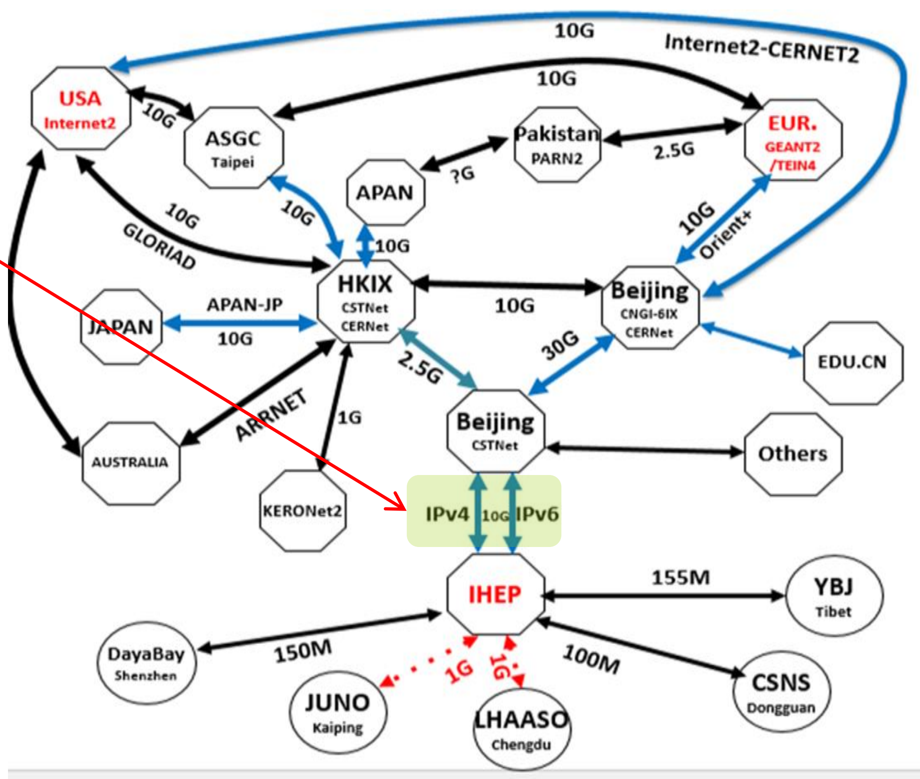


- Current:
  - 10PB(disk)+5PB(tape) = 15PB
  - 20K CPU cores
  - 3.5 ~4 times larger than the last 5 year
- In the next 10 years, > 10 larger
  - > 150PB

Network

- **Network capacity is >10 times faster in 3 years**
  - In 2012, 10Gbps LAN, 1Gbps to Europe and USA
  - Since 2015, 160Gbps LAN, 20Gbps outbound，10Gbps to Europe/USA

- **In 2018 IHEP joined LHCOne, with CSTNet and CERNET**
  - Two dedicated 20Gbps are added
  - LHC Beijing T2 are switching to LHCOne
  - Other Non-LHC experiments are applying to benefit from this link
- **Network provides good env for international cooperation**

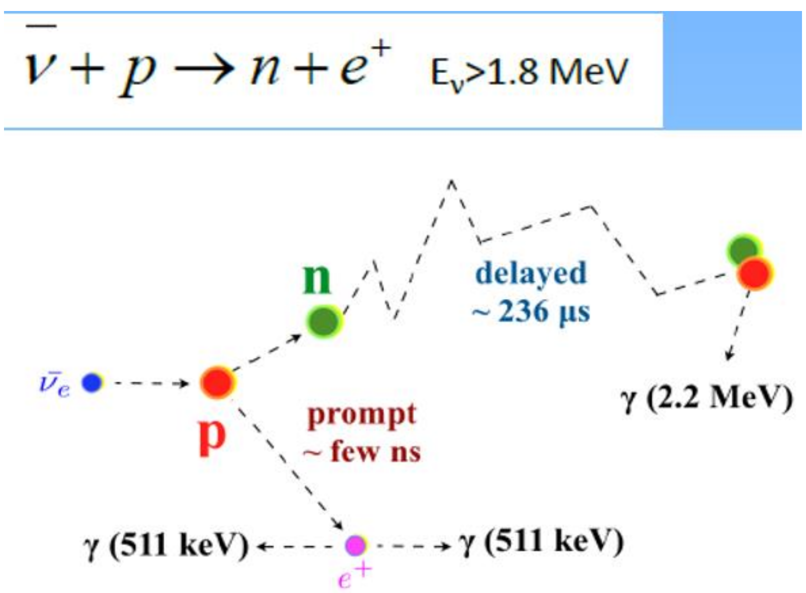


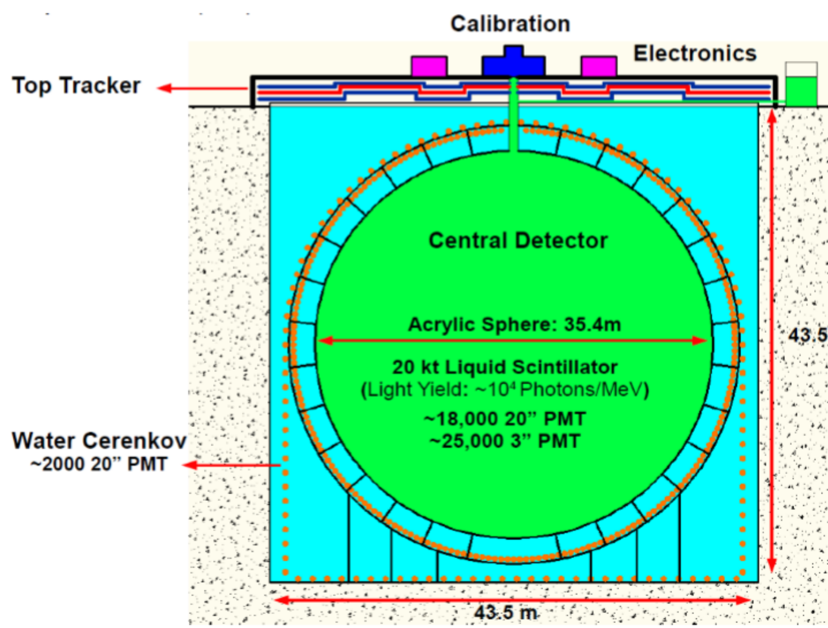
General Challenges

- Increased scale of experiments, number of experiments, complexity of experiments, data volume and resource need
  - fully exploit the advantage of modern hardware architectures
  - integrate available heterogenerous resources, benefit from the fast development of network
- Manpower on software and computing not grow accordingly
  - Software and computing work not well recognized among physics community
  - Eg. Increased resource supports from domestic collaborators, but still weak site support
- To make things "easier"
  - Common infrastructures to cover needs of all experiments on both software and computing, one effort to benefit all
  - "Easy" grid to fit into the need without too much efforts of sites

Deep view of JUNO experiment

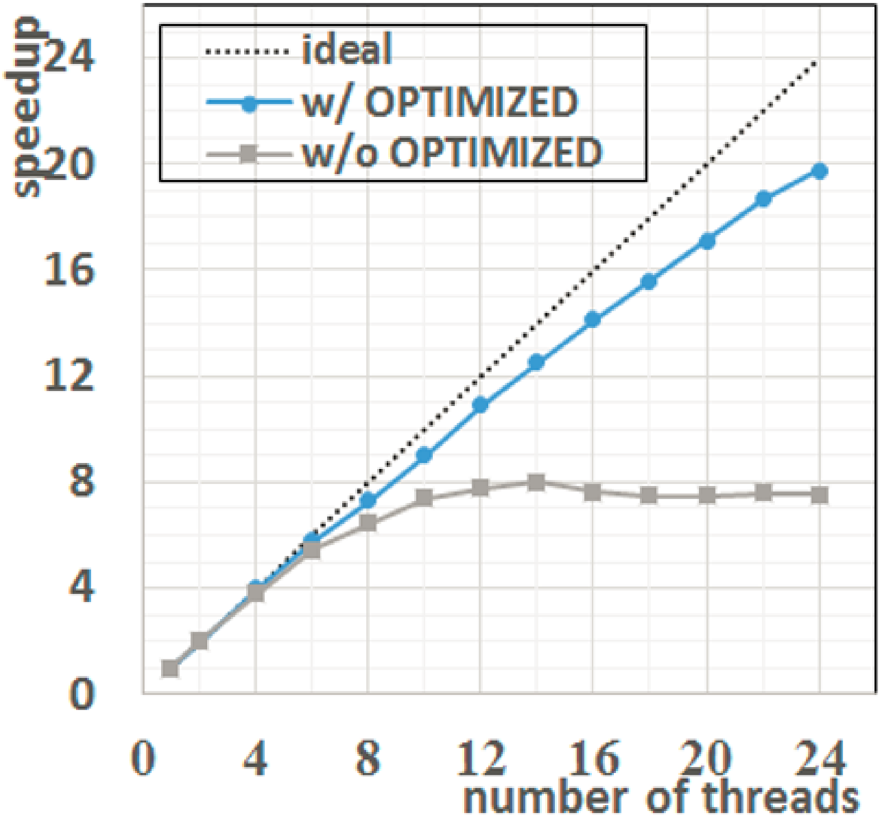
- The central detector is made of liquid scintillator (LS) and ~18,000 PMTs
- For such a large LS detector, the rate of cosmic muons reaching the inner detector is about 3 Hz
- Understanding muon induced background is crucial part of the experiment
  - One 100Gev comic Muon create several millions of optical photons
- Event correlation, analysis has to cover more than one event in one loop





Simulation (1)

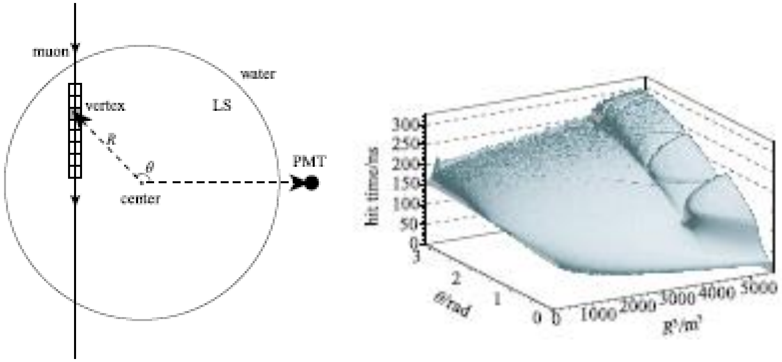
- Current performance efforts are focused on parallelism
- Benefit from the multi-threading support and thread safe of Geant4 10.x, performance optimization with parallelism in simulation have been done
  - Got a good speed-up
- In ACAT 2017, “Parallelized JUNO simulation software based on SNiPER” (poster Lightning Talk)



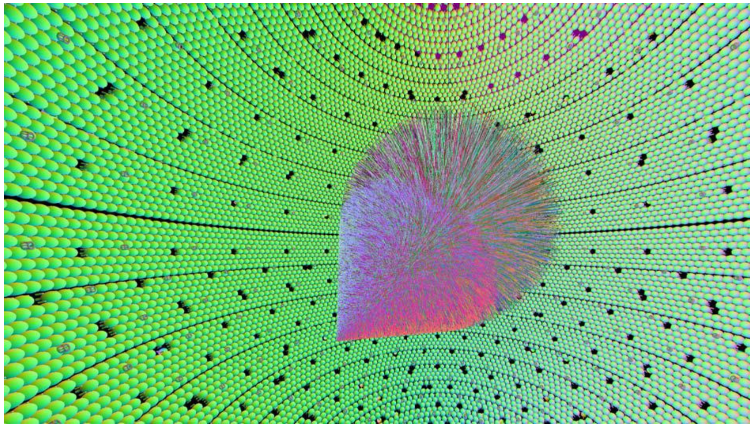
Simulation (2)

- Simulation of optical photons produced by Muon in LS poses severe constraints on both CPU time and Memory
  - CPU time >95%, memory>2GB, > 2 hours/event
- Push us to explore ways for optical photons simulation
  - Massive parallelism with GPU , achieving 1000 times speed-up
  - Fast simulation, and improve it with Machine Learning in future

Tao Lin, **fast simulation**. J.Phys.Conf.Ser. 898 (2017) no.4, 042001



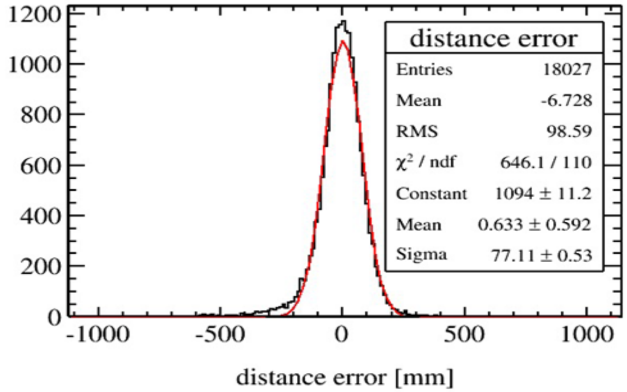
Simon Blyth, Opticks, **GPU full simulation**. J.Phys.Conf.Ser. 898 (2017) no.4, 042001



Reconstruction

- Precisely rejecting muon induced background is the hard part due to complexity of effects of optical photons in LS
- Traditional way with optical model has difficulties to make corrections with such complexity
- We are looking into Deep Learning which help us learn the relationship between the event signals and the track parameters automatically
- Preliminary study has shown very promising results

"Muon reconstruction of JUNO experiment with convolutional neural networks", ACAT talk, 2017



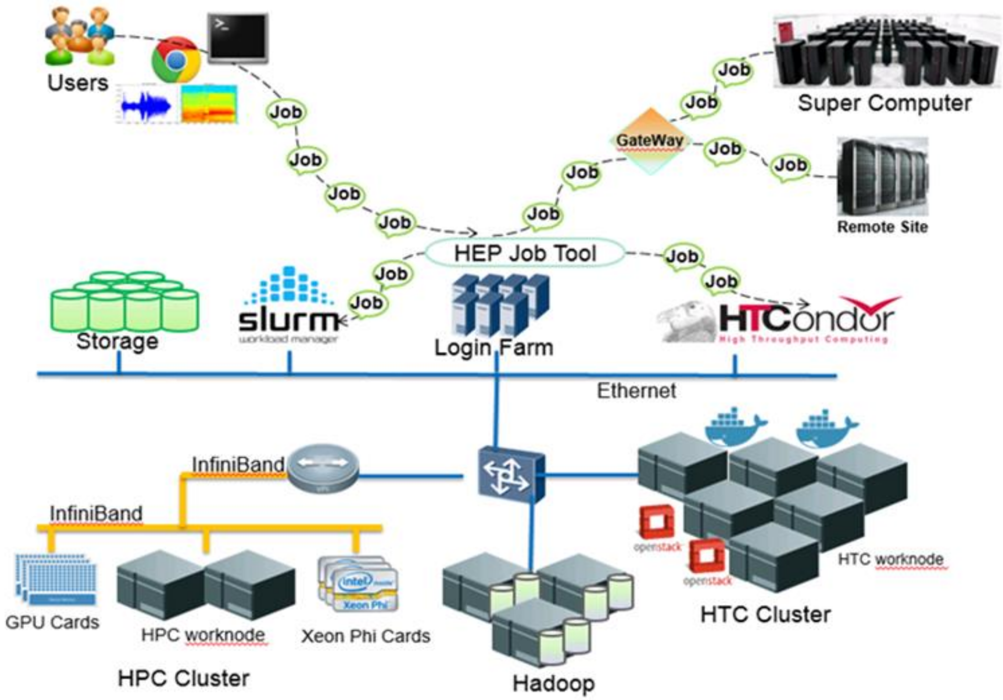
Analysis

- In 2020s, BESIII enters in the last part of its life cycle
- With large aggregated data, analysis will become a significant part of BESIII data processing
- Partial Waves Analysis with GPU could become more important
  - Got experience a few years ago, but only used by small groups due to complexity
  - How to extend it to the whole group in an easy way?
- I/O could also be bottleneck for large scale of analysis

Data-flow processing framework

- Gaudi was adopted in BESIII since 2008
  - Work fine in collider experiments
- SNiPER was developed to meet special requirements of Neutrino program
  - Multi I/O stream, Event correlation and Hits mixing
  - Light-weighted and simple, becoming a common framework for IHEP medium and small experiments
- SNiPER-MT is being developed to support parallelism
  - 2017 ACAT talk “Parallel computing of SNiPER based on Intel TBB”
- Future plans and challenges
  - Full parallelization, Support Machine Learning from infrastructure
  - Adapt to heterogeneous hardware to allow usage of more resource

Computing



- HTCondor with Lustre provides HTC platform for all the IHEP experiments
- IHEPCloud based on OpenStack was built in 2015 to ease the way of managing large scale of resources
- With parallelism and machine learning becoming hot among IHEP experiments, HPC farm is being built up last year in IHEP using SLURM
- How to efficiently organize and share resources in mixture of HTC and HPC environment?

Distributed computing(1)

- IHEP distributed computing is built in 2014
  - Integrate resources from collaborations and commercial resources
- With limited manpower, make things as easy as possible
  - Adopt DIRAC as WMS, simple and flexible
  - Allow sites to join just as clusters from the beginning
  - One big central SE, sites allow to join without SE
- Good cooperation with DIRAC community
  - Join DIRAC consortium in 2016
  - Benefit a lot from DIRAC
  - Join the efforts on common need

Distributed computing(2)

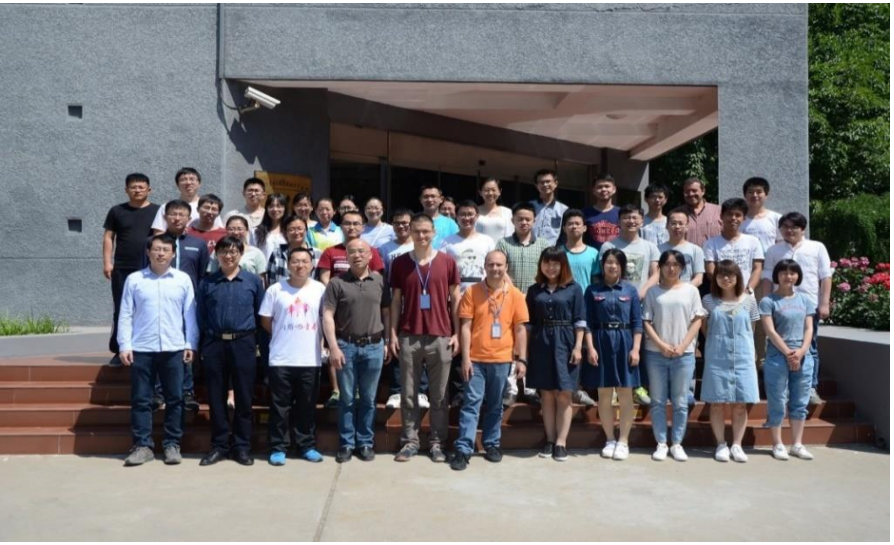
- Recent efforts
  - Become a common infrastructure for multi-experiments
    - Work have been done on WMS and file catalog
    - Need more efforts on condition database, bookkeeping, monitoring, data transfer……
  - Integrate more available resources
    - Done with Cloud, Cluster……
    - Use Singularity to avoid differences from OS versions
  - Multi-core supports for parallelized experiment software
    - Efficiency need to be further study with real use cases
  - All the above work have the related CHEP and ACAT posters and talks

Distributed computing(3)

- Challenges
  - With larger scale and modern network, need more “smart” storage
    - Hierarchy storage? Data lake?
  - Integration of HTC and HPC resources
    - not easy with complexity of HPC resources
  - With larger scale, site maintenance is a pain
    - Current grid infrastructure need too much efforts, Downtime and Resource information, Security, monitoring, operation groups…….

cooperation

- We already have some good experience on mini-workshops with international software and computing groups
  - BESIII Cloud computing Summer School with INFN, Sep 2015
  - A four days hands-on Geant4 Mini-workshop with INFN, May 2017
  - IHEP network security workshop with CERN, Sep 2017
- Looking forward to more chances on cooperation and sharing efforts



Summary

- IHEP experiments cover wide fields, including collider, Neutrino, Astrophysics…..
- Have common concerns with HSF about future challenges
- In 2020s, our software challenges mainly come from Neutrino experiments
  - Push our computing and software to be on the road of parallelization
- Building common computing and software infrastructure are also our plans to serve all experiments
- Look forward to closer cooperation with other experiments