Perspectives of the computing/software demanding of CEPC and SppC

Gang Li for the CEPC study group

li.gang@mail.ihep.ac.cn

Institute of High Energy Physics, CAS

Outline

- Introduction to the projects
- CEPC computing/software
- Perspective of SppC
- Summary

CEPC-SppC

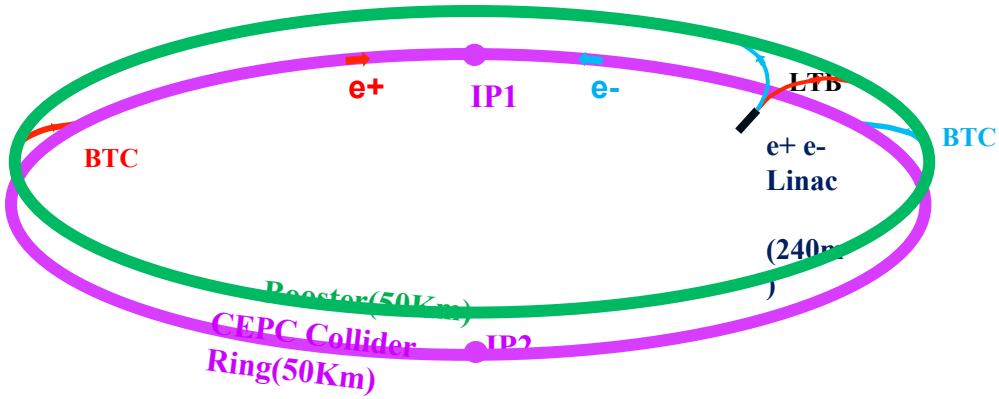
Phase 1: e^+e^- Higgs (Z) factory **two detectors, 1M ZH events in 10yrs**

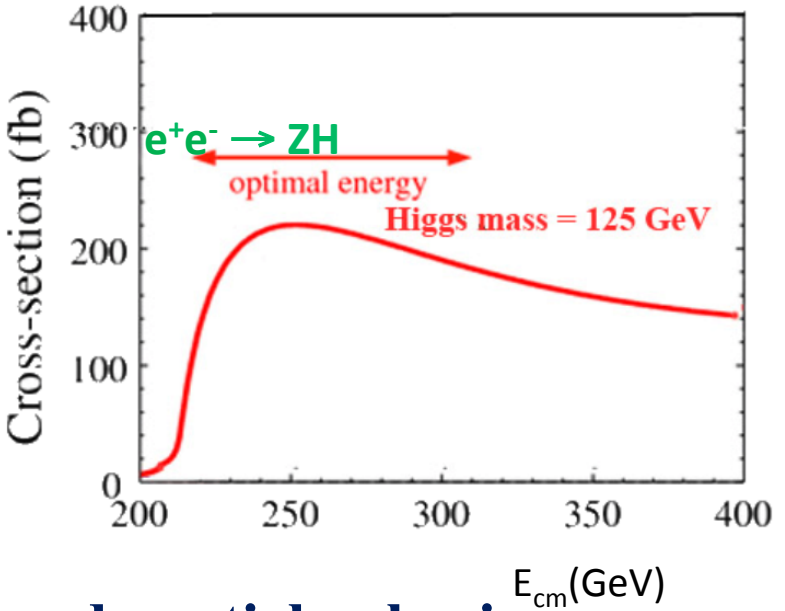
E_{cm}≈240GeV, luminosity ~2×10^{34} $cm^{-2}s^{-1}$, can also run at the Z-pole

Precision measurement of the Higgs boson (and the W/Z boson)

Phase 2: a discovery machine; pp collision with E_{cm} ≈ 50-100 TeV; ep,HI options

Discovery machine for BSM



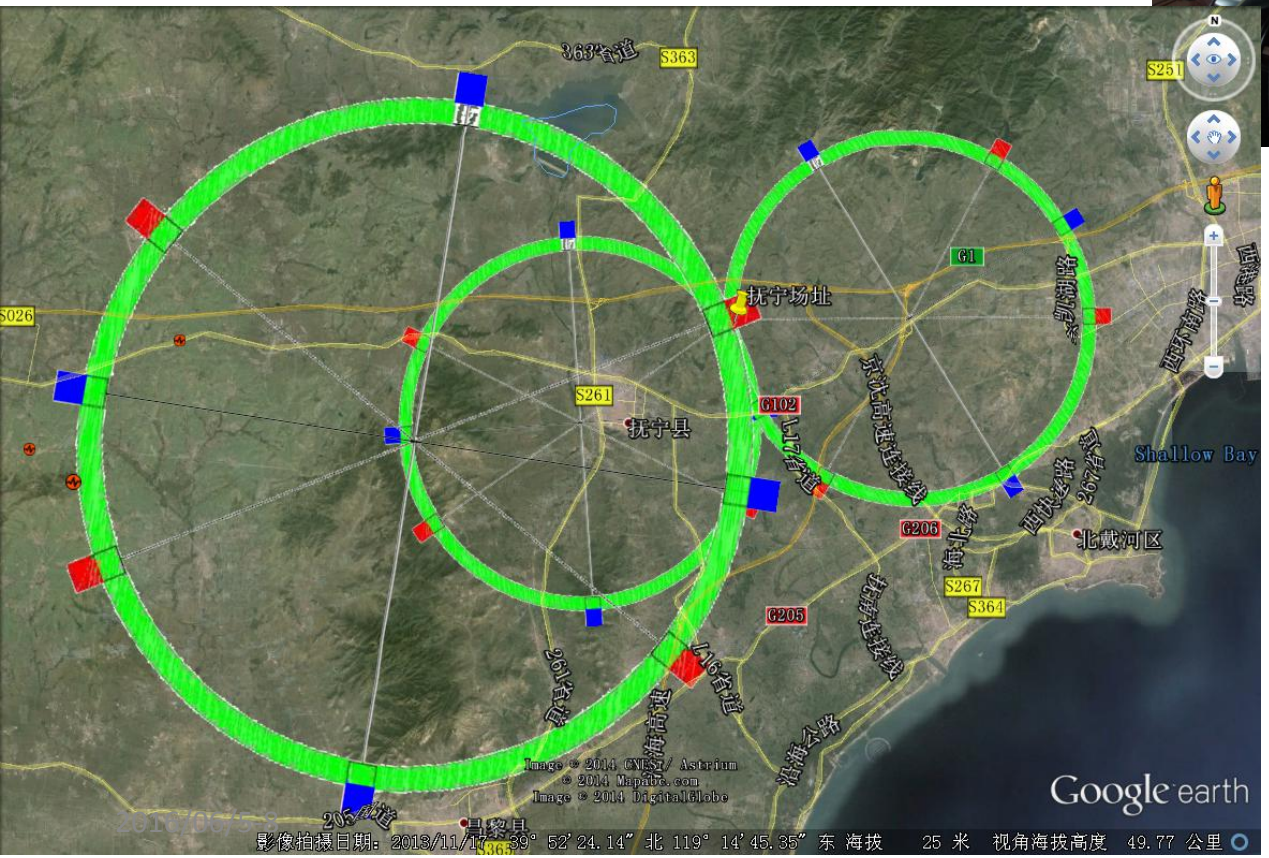


favored post BEPCII accelerator based particle physics program in China

Office of Engineering and Support at IHEP

- appointed by IHEP director Y. F. Wang on April 15, 2014
- veterans and very experienced professionals

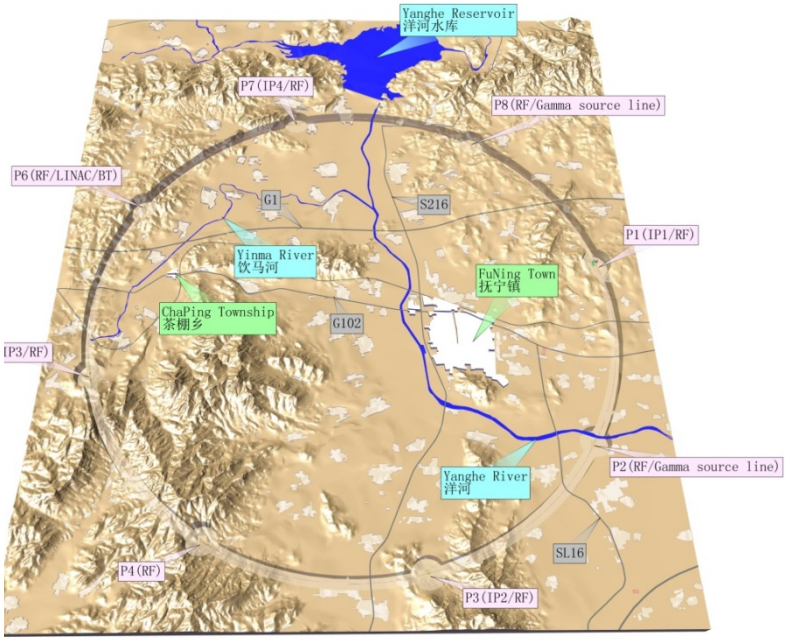
"enormous amount of effort & progress"

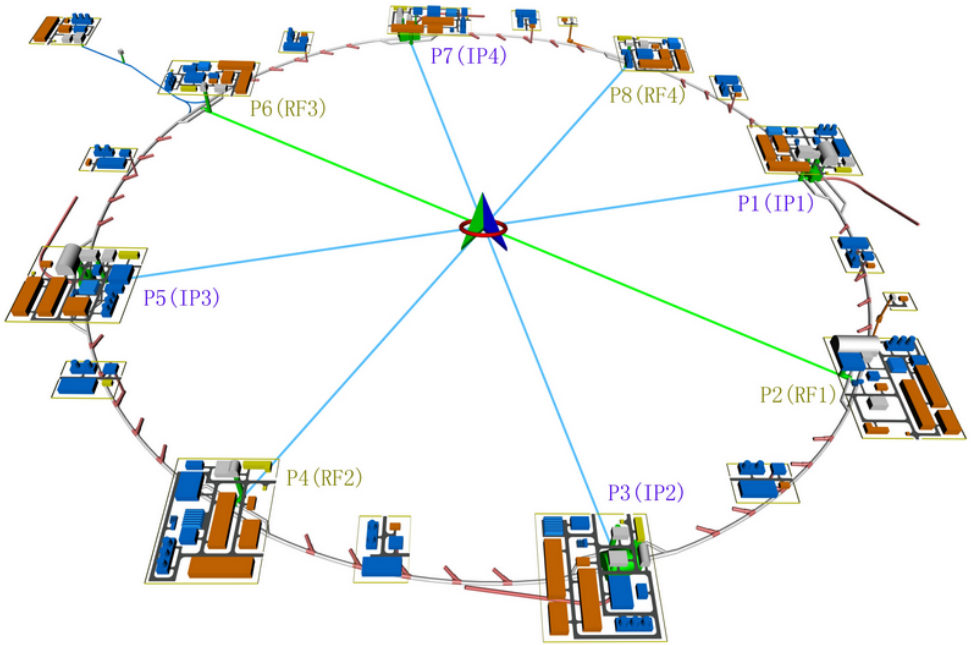


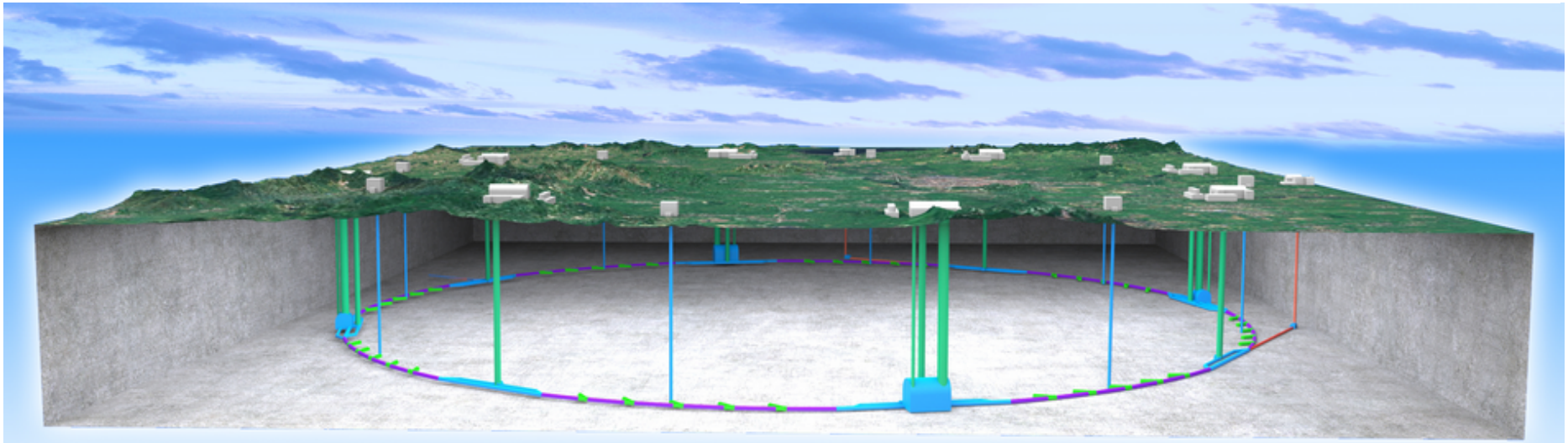
QingHuangDao site Investigation

- 300km from Beijing
- Geo well suited
- Great environment

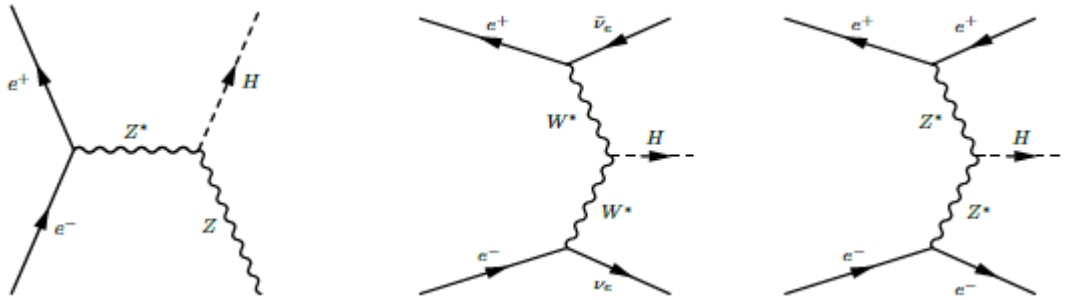
CEPC “Qinghuandao Site” Investigation



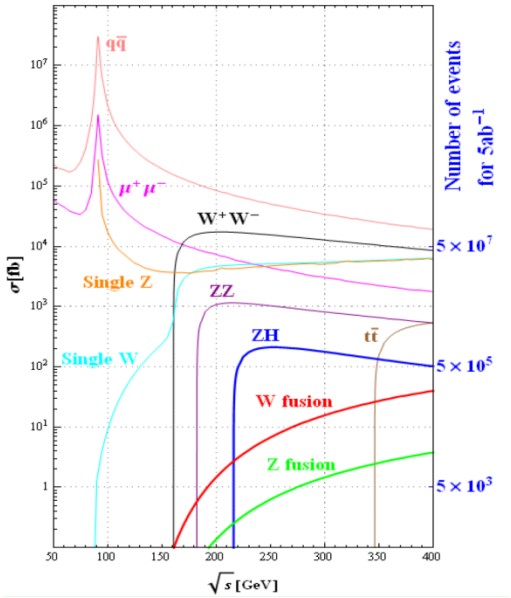




- Precise measurements of the Higgs properties as a Higgs Factory (similar to ILC@250 GeV)
  - Mass, J^{PC}, couplings, etc. → reach (sub-) percentage accuracy



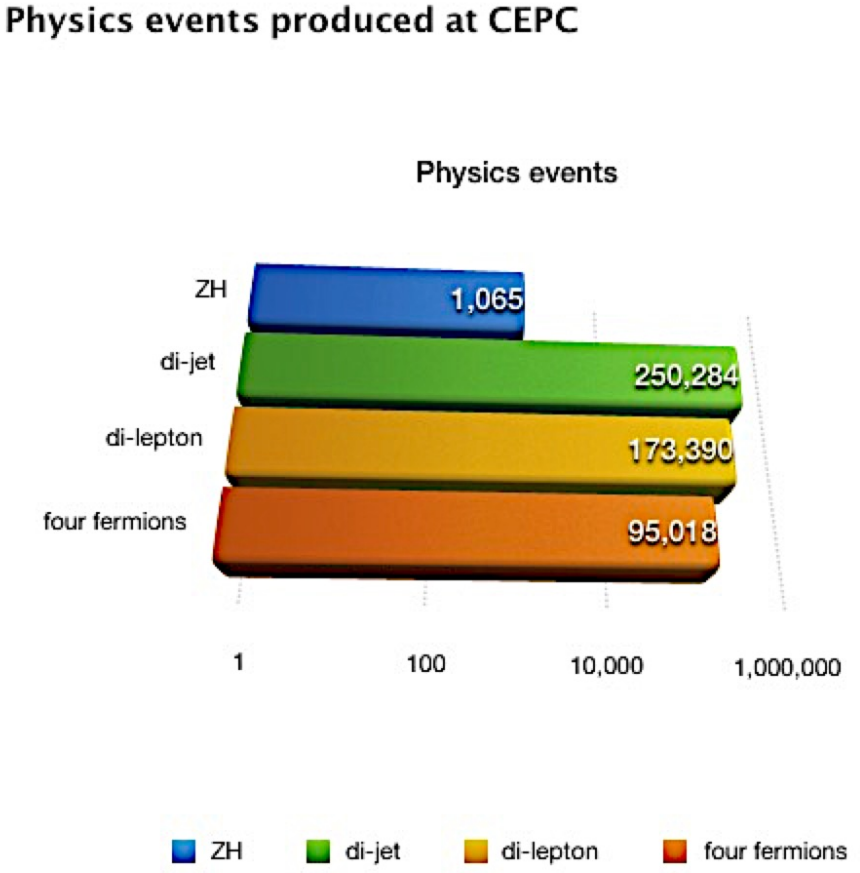
Process	Cross section (fb)	Nevents in 5 ab^{-1}
	Higgs boson production	
$e^+e^- \to ZH$	209	1×10^6
$e^+e^- \to \nu\bar{\nu}H$	6.9	3.5×10^4
$e^+e^- \to e^+e^-H$	0.6	3.0×10^3



- Precise measurements of Electroweak Symmetry-Breaking parameters at Z-pole and WW threshold
  - $m_Z, m_W, \Gamma_Z, \sin^2\theta_W^{\text{eff}}, \alpha_S$, *etc. + searches for rare decays*

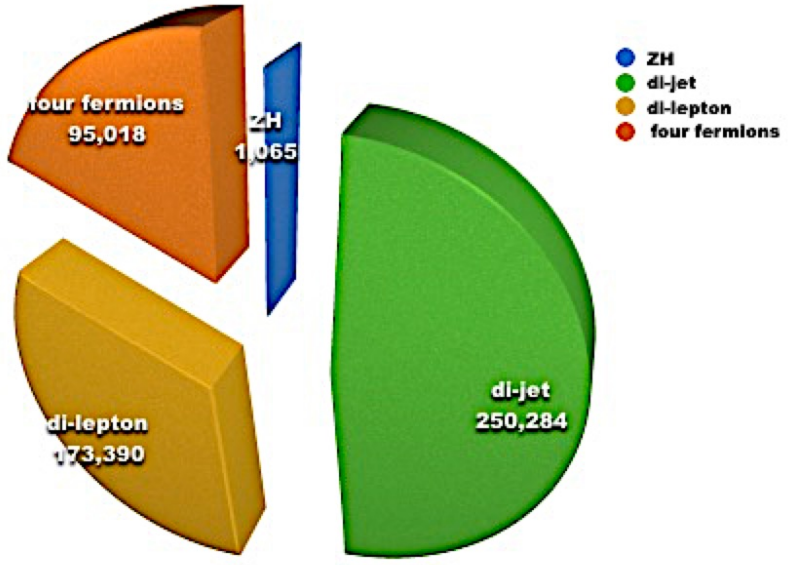
Typical event at CEPC~50-100 particles

CEPC上预期物理事例：
5亿@240GeV +300亿@91GeV（不包含γγ*事例）

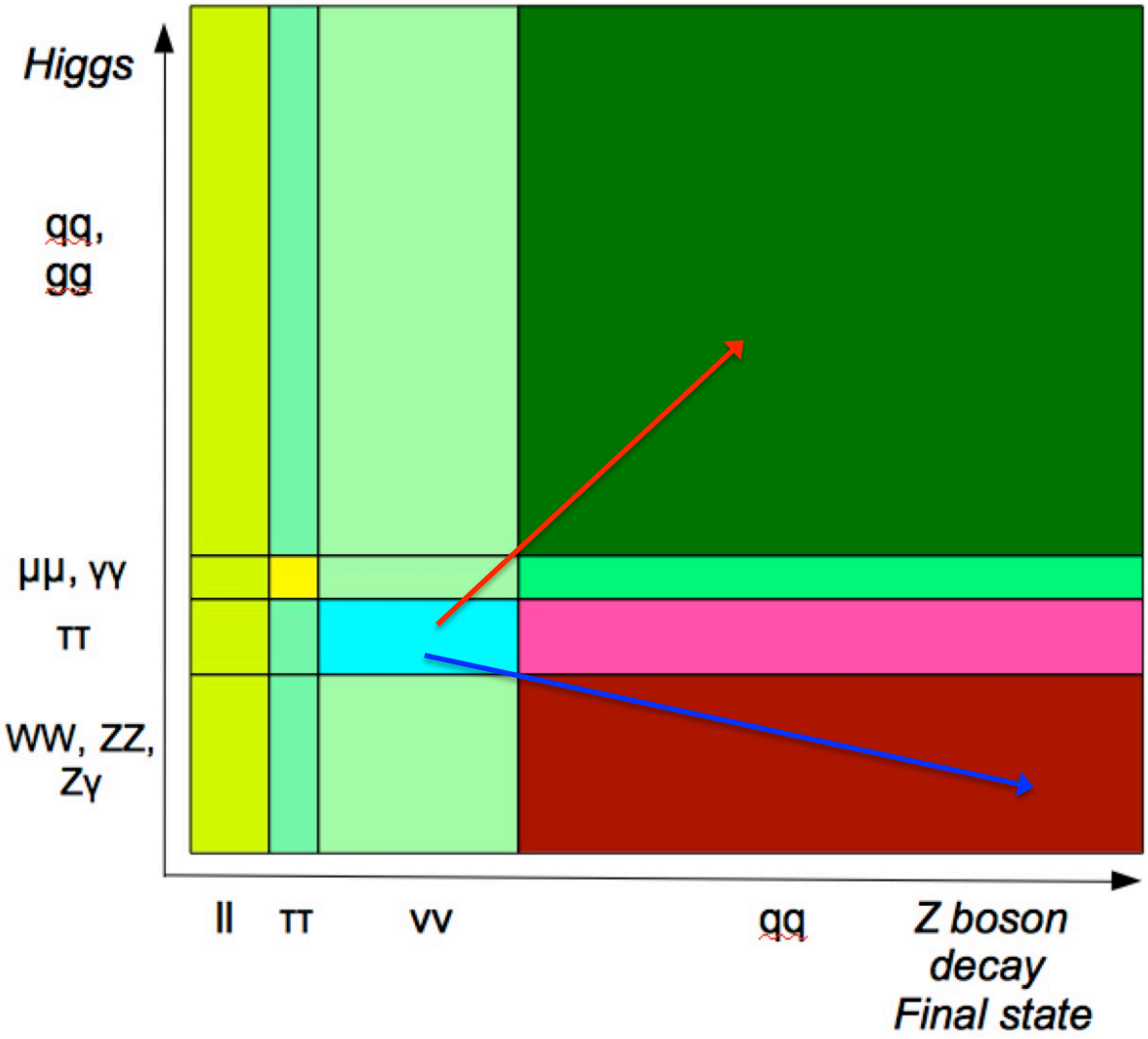


Number of interested physics events

CATEGORY	# of events (k)
ZH	1,065
di-jet	250,284
di-lepton	173,390
four fermions	95,018



Higgs 物理

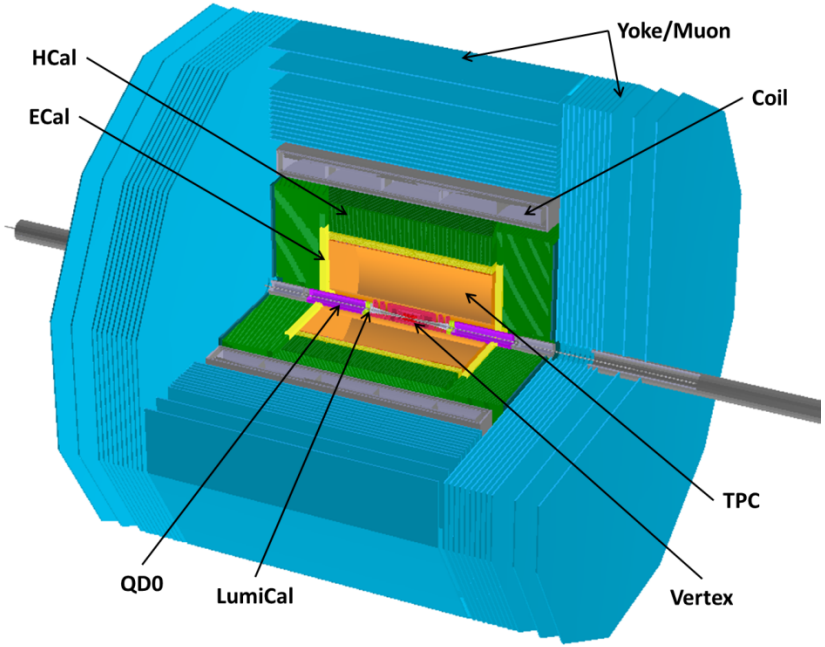


复杂的物理对象

挑战算法

大量计算、存储

CEPC Detector : needs optimization



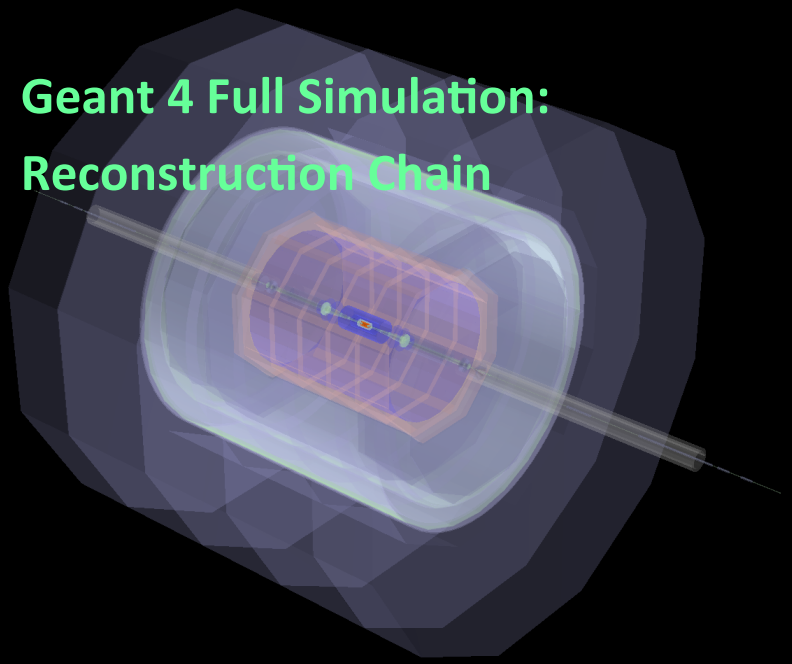
ILD-like detector with additional considerations:

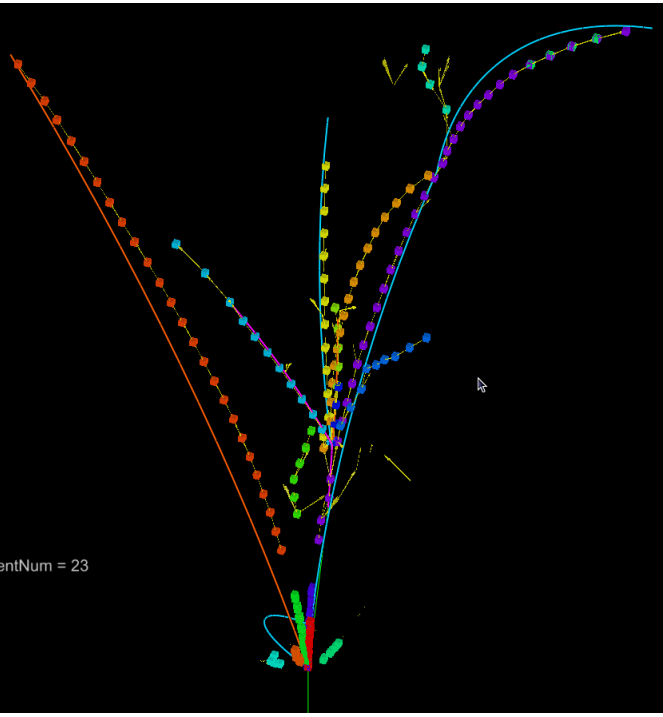
- Shorter L* (**1.5**/2.5m)
- No power-pulsing
- Limited CM (up to 250 GeV) → calorimeters of reduced size
- Lower radiation background → vertex detector closer to IP
- …

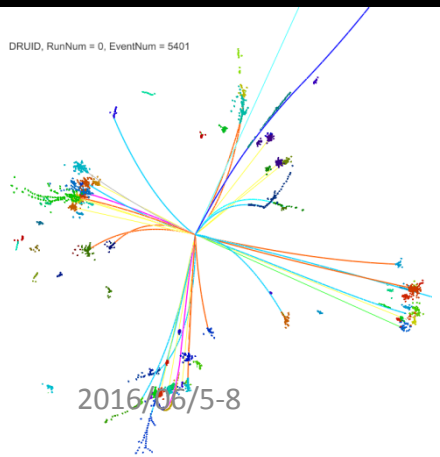
- **Similar performance requirements to ILC detectors**
  - Momentum: $\sigma_{1/p} < 5\times10^{-5}\ \mathrm{GeV}^{-1}$ ← recoiled Higgs mass
  - Impact parameter: $\sigma_{r\phi} = 5 \oplus 10/(p\cdot\sin^{\frac{3}{2}}\theta)\ \mu\mathrm{m}$ ← flavor tagging, Br
  - Jet energy: $\frac{\sigma_E}{E} \approx 3-4\%$ ← W/Z di-jet mass separation

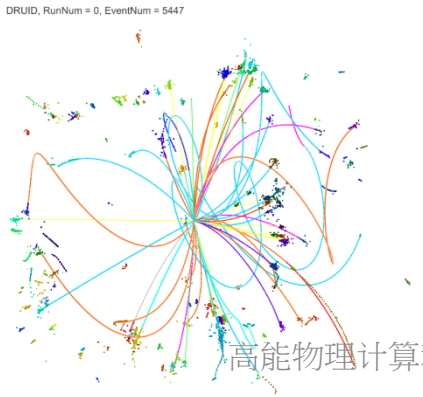
Simulation & optimization

- Geant 4 Full Simulation:
- Reconstruction Chain



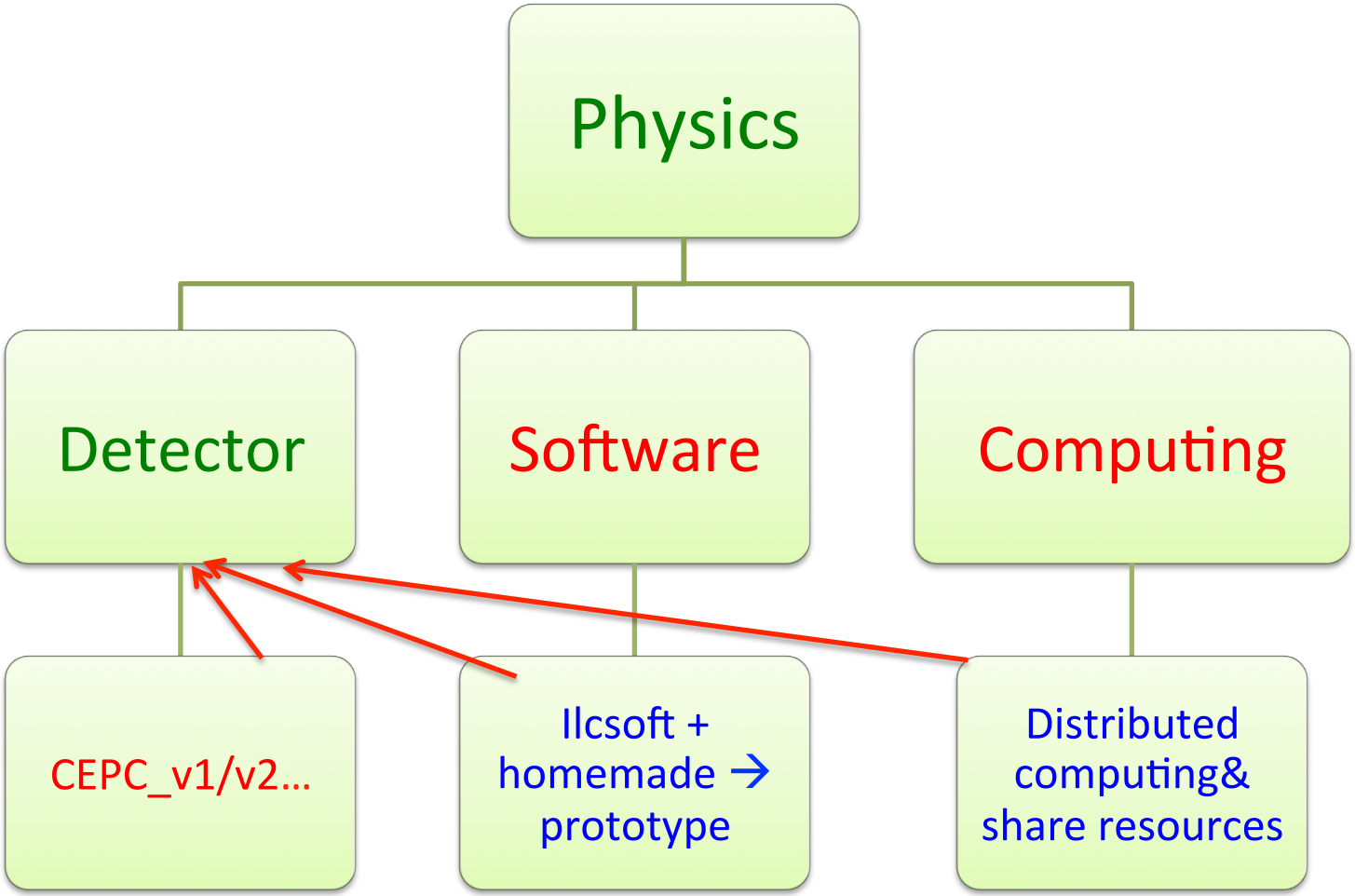






仅一次全模拟样本的产生过程就需要2000个核0.5年的计算量

Computing


Physics
Detector
Software
Computing
CEPC_v1/v2…
Ilcsoft + homemade → prototype
Distributed computing& share resources

Demand analysis

- R&D phase : data volume evaluation
  - Demand： 1PB storage, 2000 CPU cores, DB servers
  - Current： 采购0.5PB，借用 500 cores + 分布式~300（晓梅，颜田）
- Experiment phase : data volume evaluation
  - Higgs工厂（0~10年）
    - 原始产生300TB/年，10年积累大约3PB
      - 假设取1*10^6事例，1MB/event
    - BESIII数据量（每年）＊3
  - Z工厂（10~11年）
    - 原始产生至少100PB/年
      - 假设取10^11事例，0.5MB/event
    - BESIII数据量（每年）＊1000
- Need more details with software design

Software framework consideration

Use an existing one  vs Develop from beginning 

- Consideration of the choice for CEPC
  - Enough services and functionalities
  - Easy to use
  - Future supports
- Almost all widely used frameworks can satisfy our requirements
- Several potential candidates are investigated and compared

Framework candidates investigation

- Marlin: currently used by CEPC(with uncertain official support?)
- Gaudi: very popular for collider physics experiments, most familiar to us, very comprehensive but a bit heavy
- ROOT: very flexible and powerful, but need more manpower for some service functionalities development
- ART: optimized for high intensity physics experiments and a little complex
- Sniper: lightweight and optimized for non-collider experiments

	Marlin	Gaudi	ROOT	ART	SNiPER
User Interface	XML	Python, TXT	Root script	FHiCL	Python
Adoption	ILC ???	Atlas, BES3, DYB	Phenix, Alice	Mu2e, NOvA, LArSoft, LBNF	JUNO, LHAASO

Computing considerations

- It is still far to confirm the computing technology now used for 30 years more
- But we believe the technology is evolving step by step
- Now the main computing task is to study and follow the latest computing technology to prepare for the future, including
  - Cloud computing
  - Distributed computing
  - Multi-cores computing
  - High performance computing
  - Unified distributed data management and access
  - “Smart” network, high bandwidth future network
  - …….

Strengthen international cooperation

- Establish relationships with international organizations
  - HEP software foundation
  - DIRAC consortium
- Take part in conferences
  - CHEP
  - Hepix
- Cooperate with international HEP experiments
  - LHC, BELLEII
- Involved in the development of advanced technology

Software

Why we need a dedicated software?

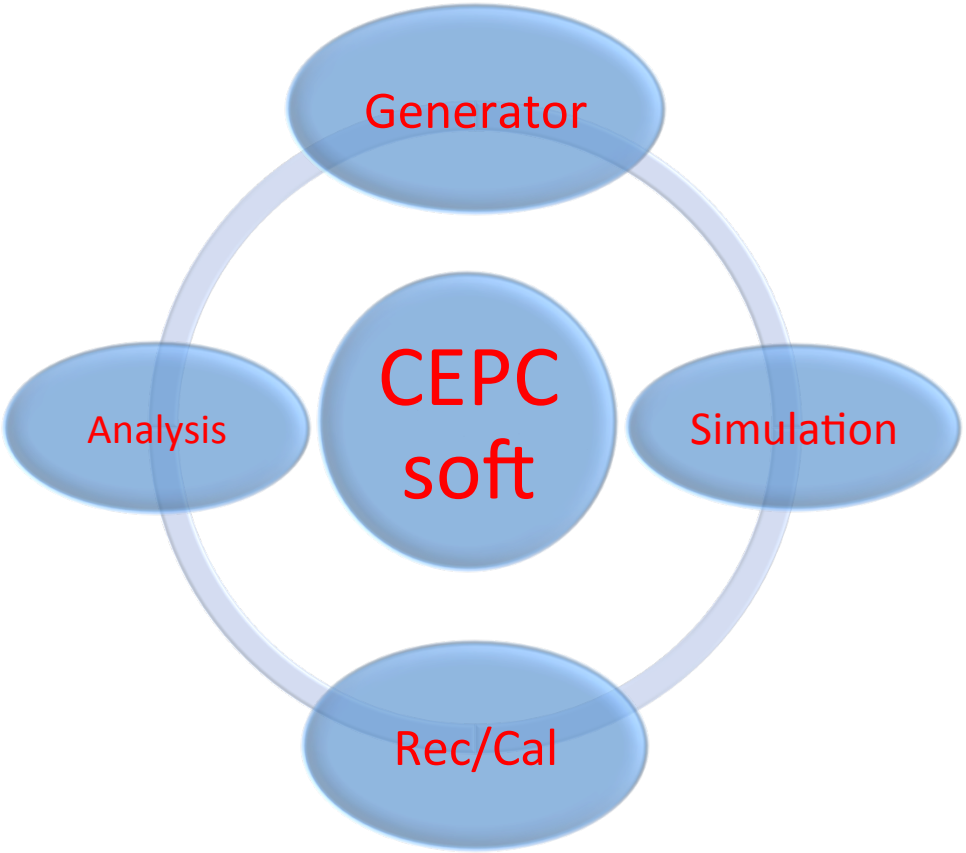
- CEPC: HZW (top?) physics
  - H (Higgs): first priority
  - Z and W (electroweak): large FREE data @ 250GeV
  - W@160GeV and Z@91GeV necessary?

- Answer 1: Demonstrate and evaluate the physics potential of HZW(top)

Why we need a dedicated software?

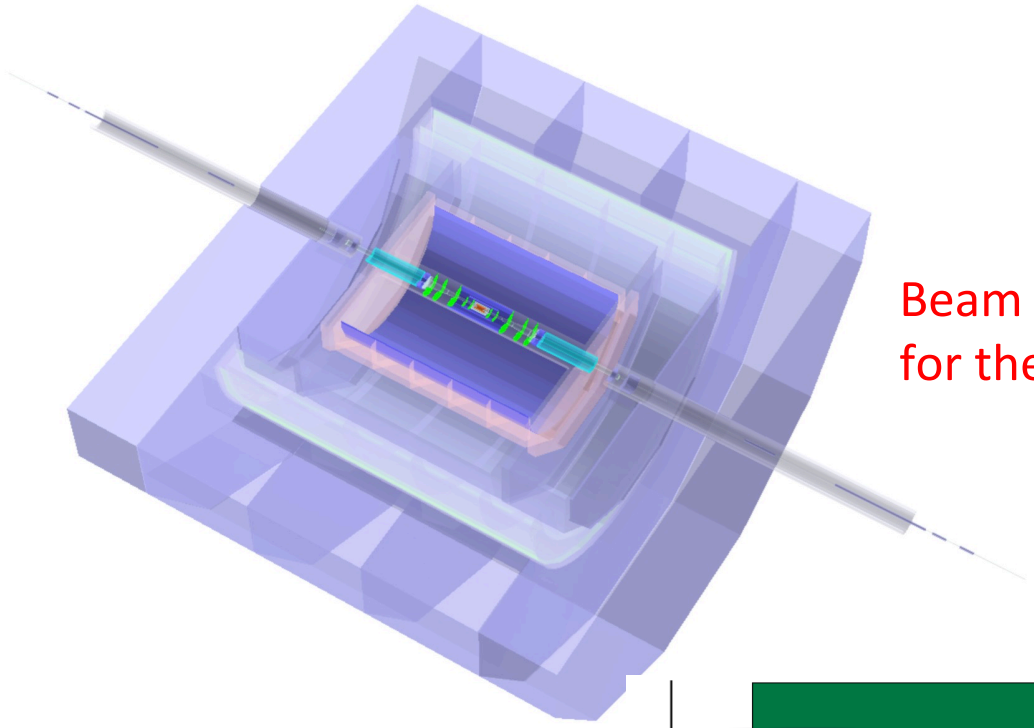
- Pre-CDR: Detector model from ILD with some modifications
- Next CDR&TDR:
  - Alternative choice: silicon
  - Detector geometry: smaller for less expense
  - Key technical problems: MDI, active cooling, B …
  - More precise vertex for jet flavor identification
  - …
- Answer 2: Optimize the detector design to balance physics and expense

Software: a chain of (G)SRCA



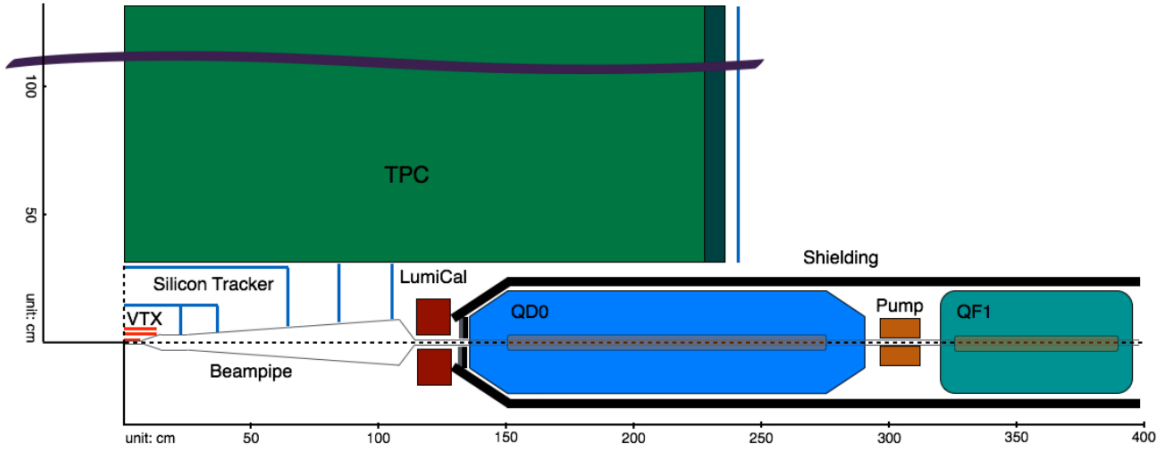
- Generator: usually independent
- Simulation: flexible to edit/ change geometry
- Rec/Cal: cope with the changes of detector and maximize the performance
- Analysis: precision

Detector: CEPC_v1

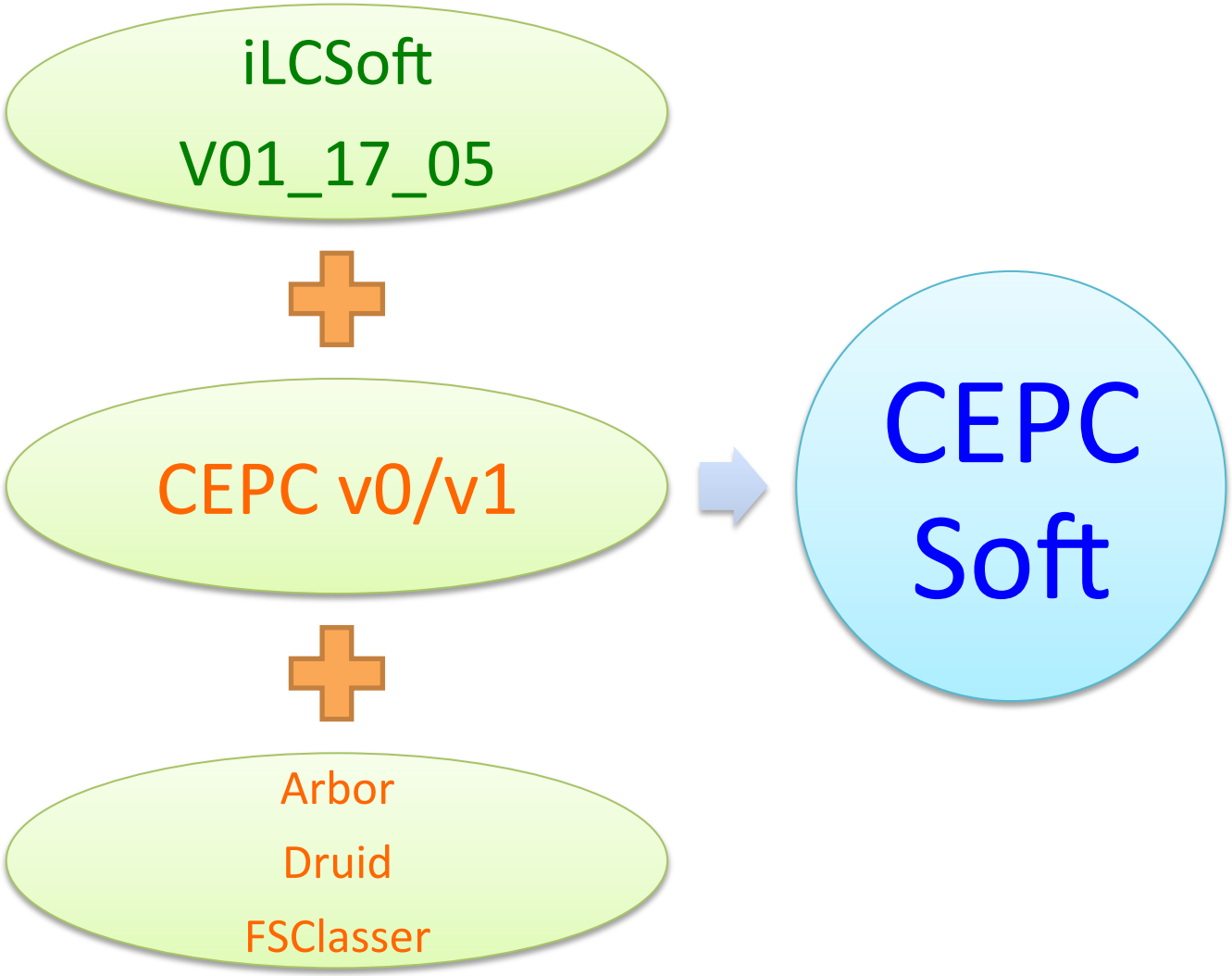


Beam backgrounds studied carefully for the MDI design

ILD reference, necessary changes



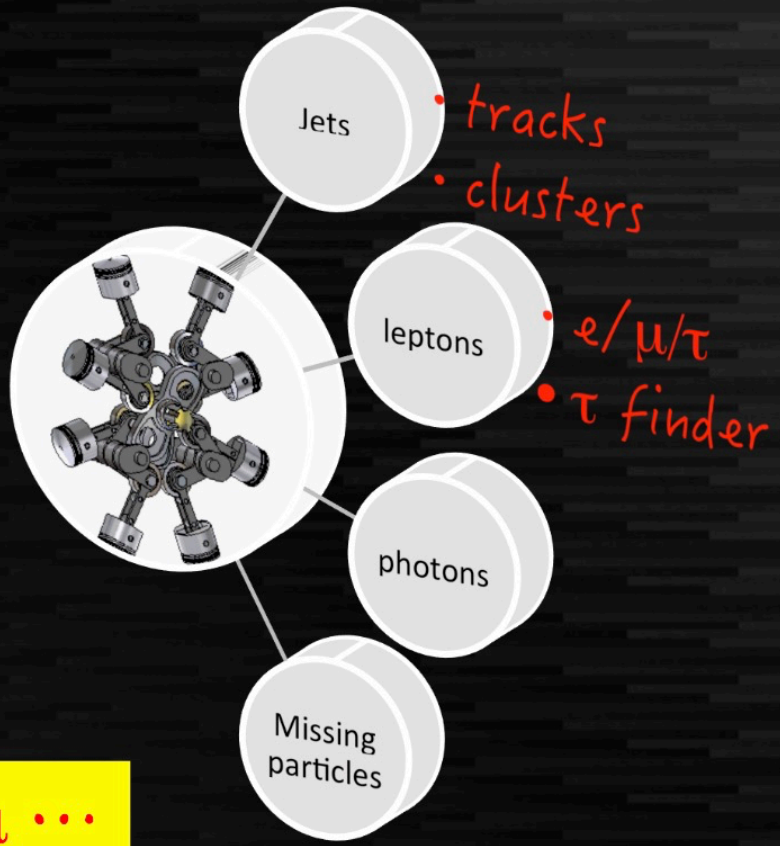
Dedicated homemade tools developed for CEPC conceptual design



A dedicated analysis framework

Novices can start from root …

Feed all types of particle object to the combination engine for further processing



ee+X, μμ+X, jj+ee, jj+μμ …

Data → ntupes → plots

Towards CDR&TDR

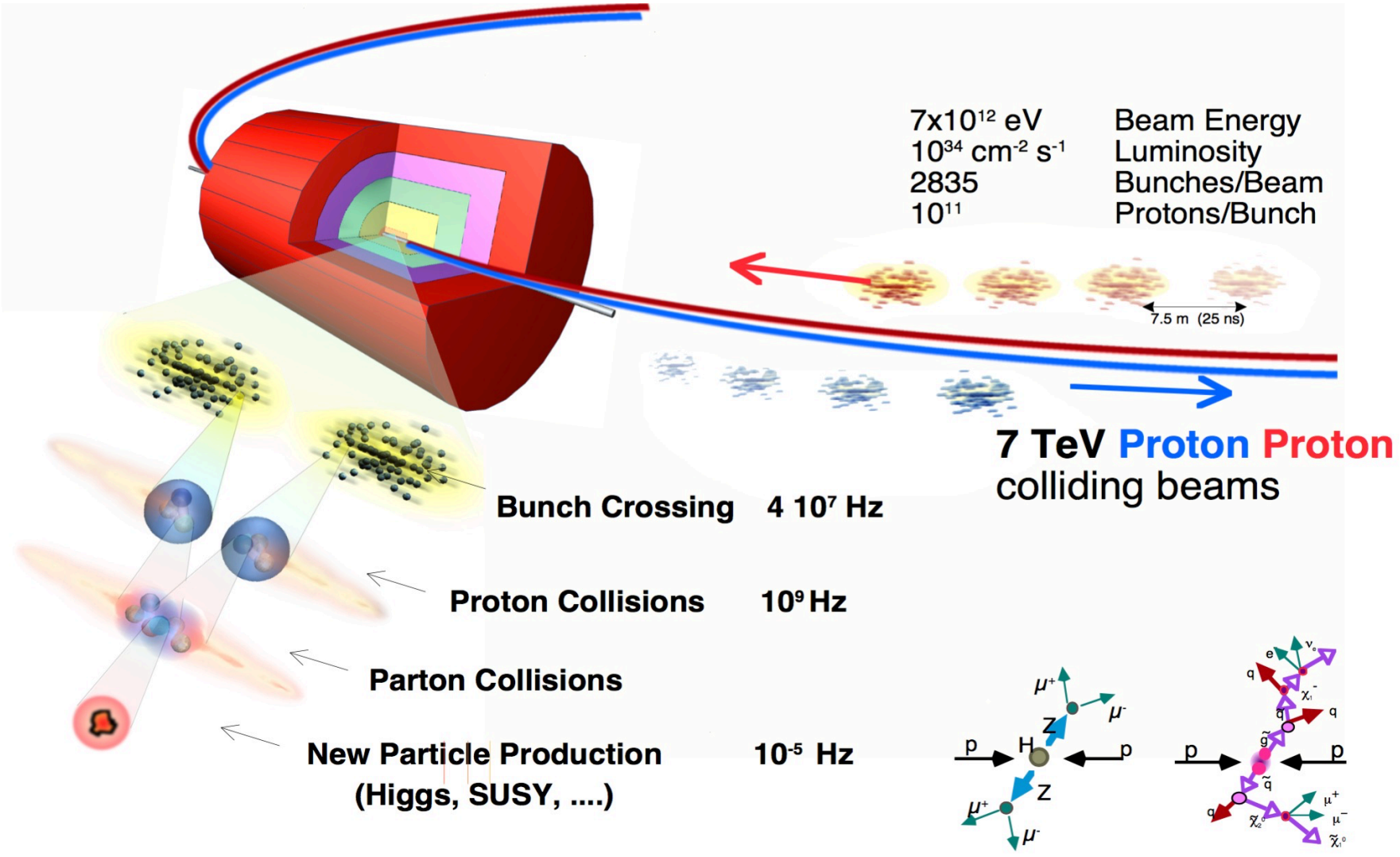
Two tasks of software

- Designing/optimizing detector and answering key questions
- Systematics control
  - Calibration
  - Dedicated physics object algorithms: e, μ, τ, γ, jet
  - MC/theoretical inputs

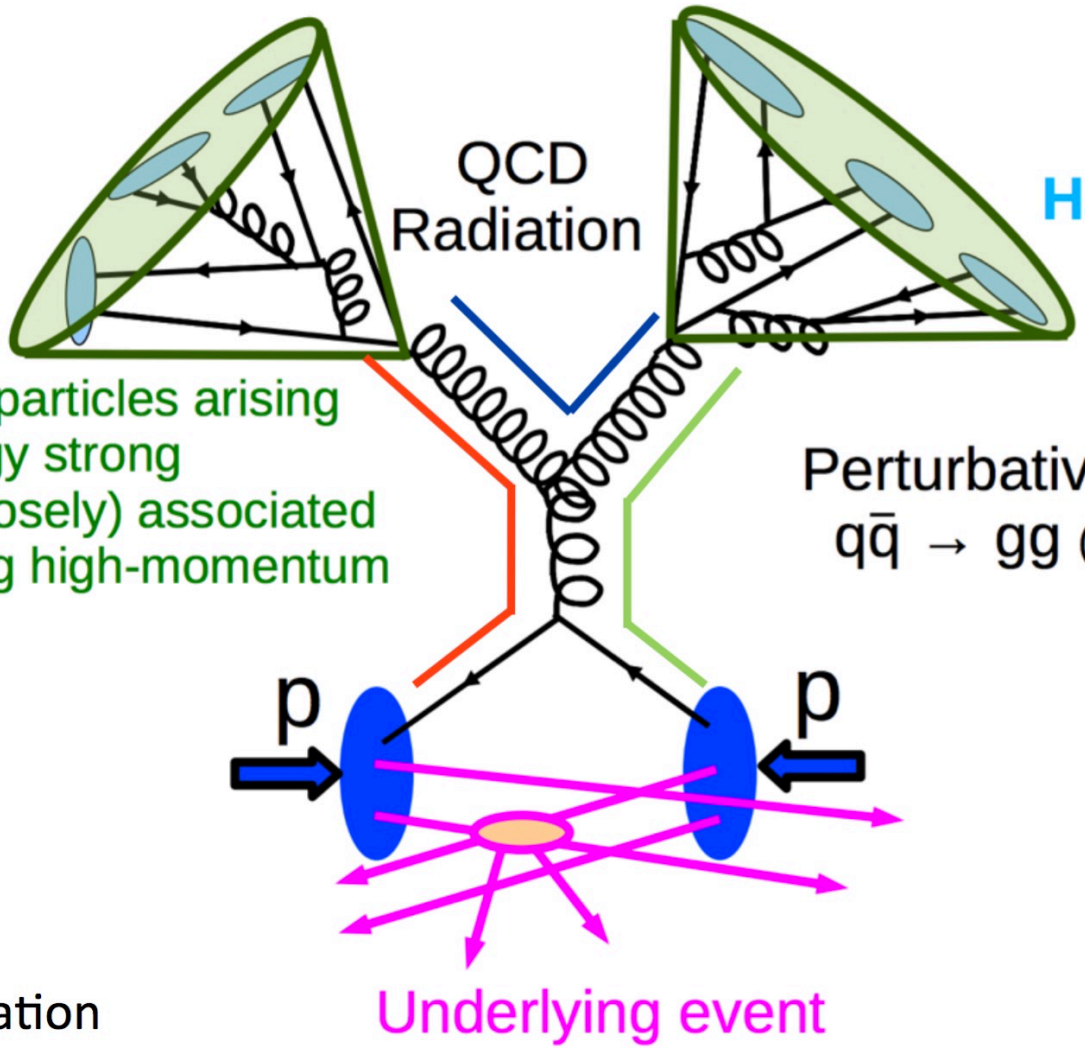
New framework: a team formed

- **CEPC Software framework:**
  a. **Developed from iLCSoft**
  b. **Sufficient for R&D & optimization studies**
  c. **Has difficulties to support experimental data taking & processing**
- **Future requirements**
  a. **Parallel computing**
  b. **Data base handling**
  c. **User friendly, efficiency, etc**
  d. **Need top level engineering/ organization**
  e. **…**

Perspectives of SppC assuming 3/ab data


7x10¹² eV
Beam Energy
10³⁴ cm⁻² s⁻¹
Luminosity
2835
Bunches/Beam
10¹¹
Protons/Bunch
7.5 m (25 ns)
7 TeV Proton Proton
colliding beams
Bunch Crossing 4 10⁷ Hz
Proton Collisions 10⁹ Hz
Parton Collisions
New Particle Production
(Higgs, SUSY, ...)
10⁻⁵ Hz
Selection of 1 event in 10,000,000,000,000

When protons collide…

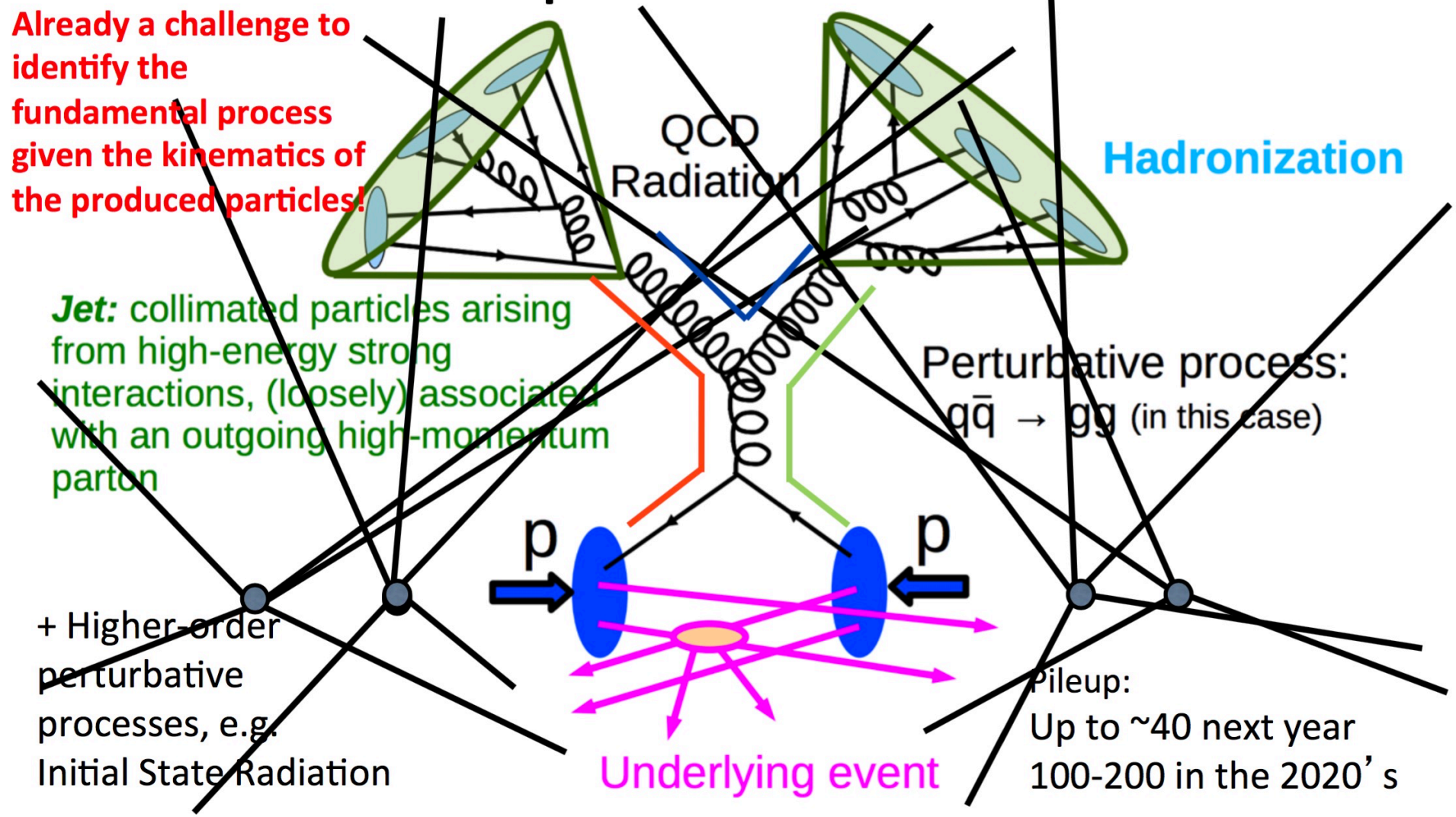


Jet: collimated particles arising from high-energy strong interactions, (loosely) associated with an outgoing high-momentum parton

Perturbative process:
$q\bar{q} \rightarrow gg$ (in this case)

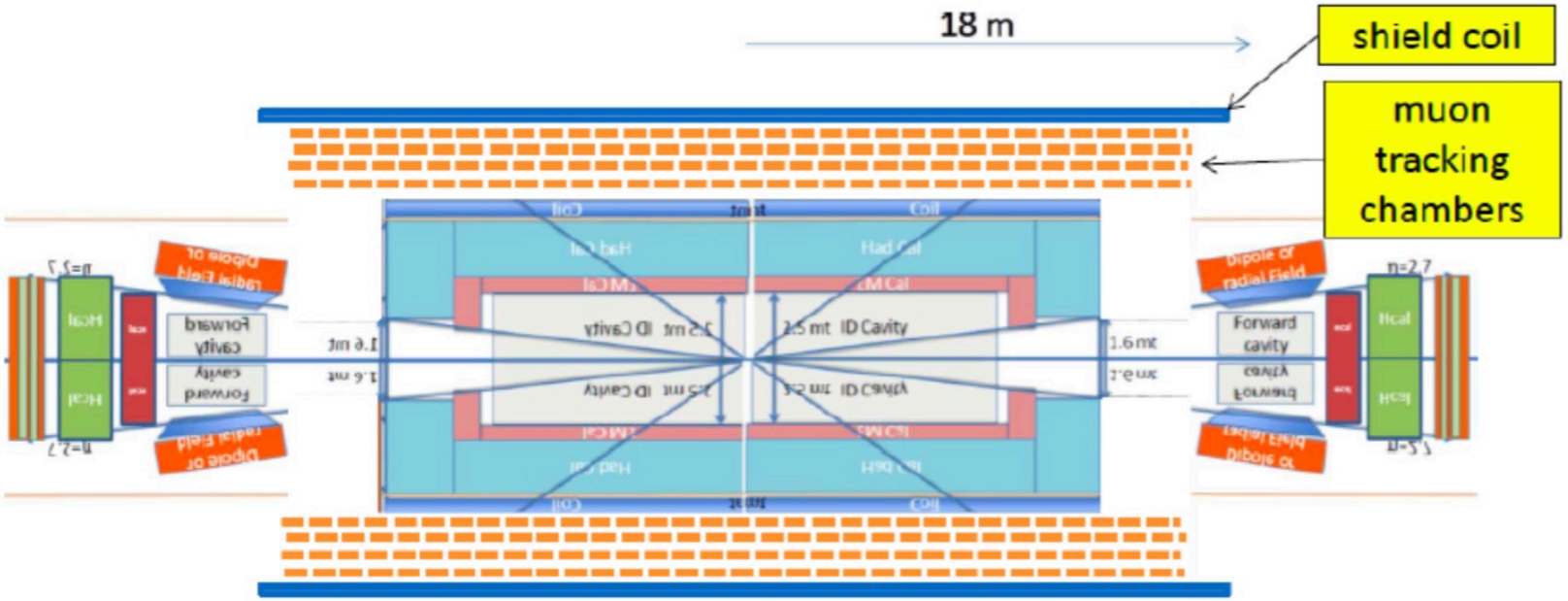
+ Higher-order perturbative processes, e.g. Initial State Radiation

More protons collide...

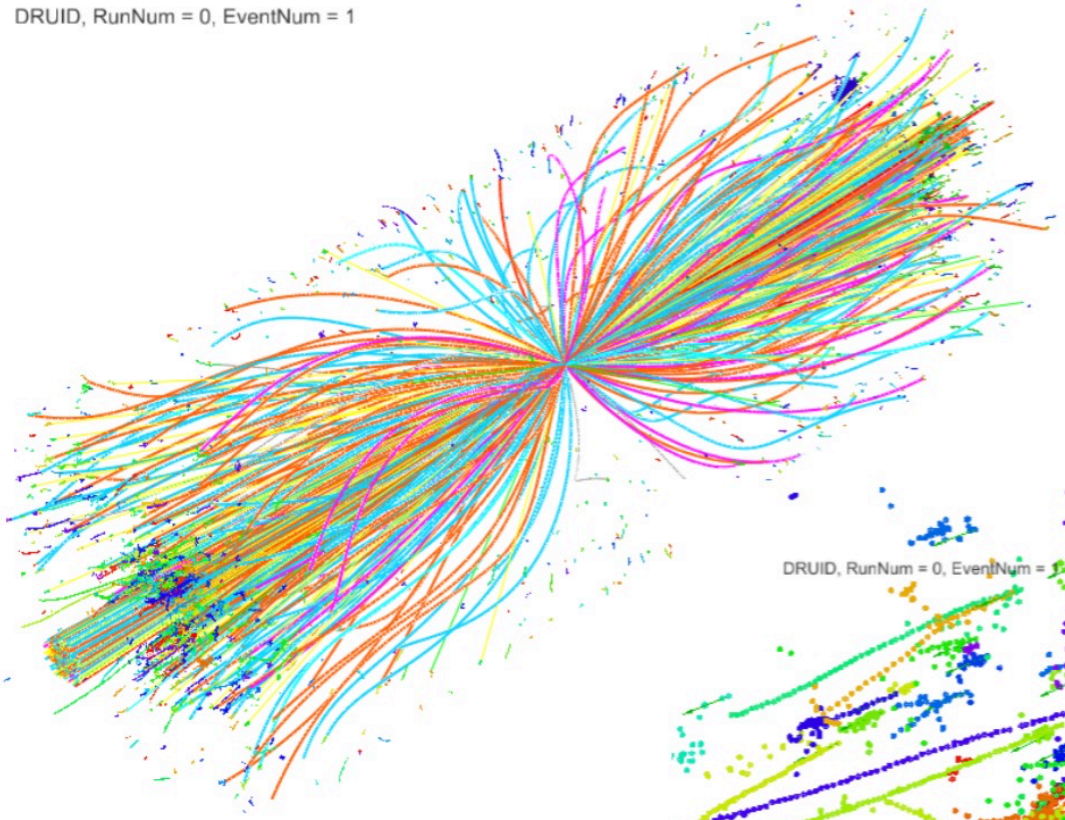


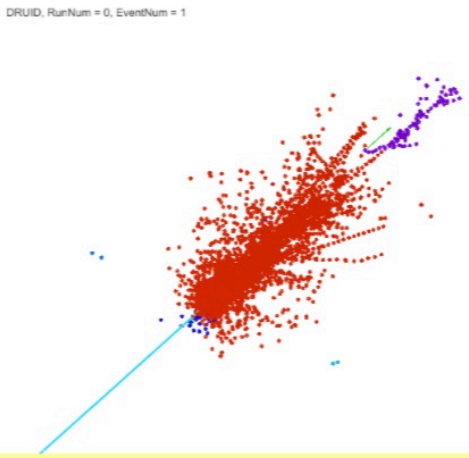

If you know how to track each single fragment of a firework from a B&W movie of only 12 frames contact us: we are interested in your solution!

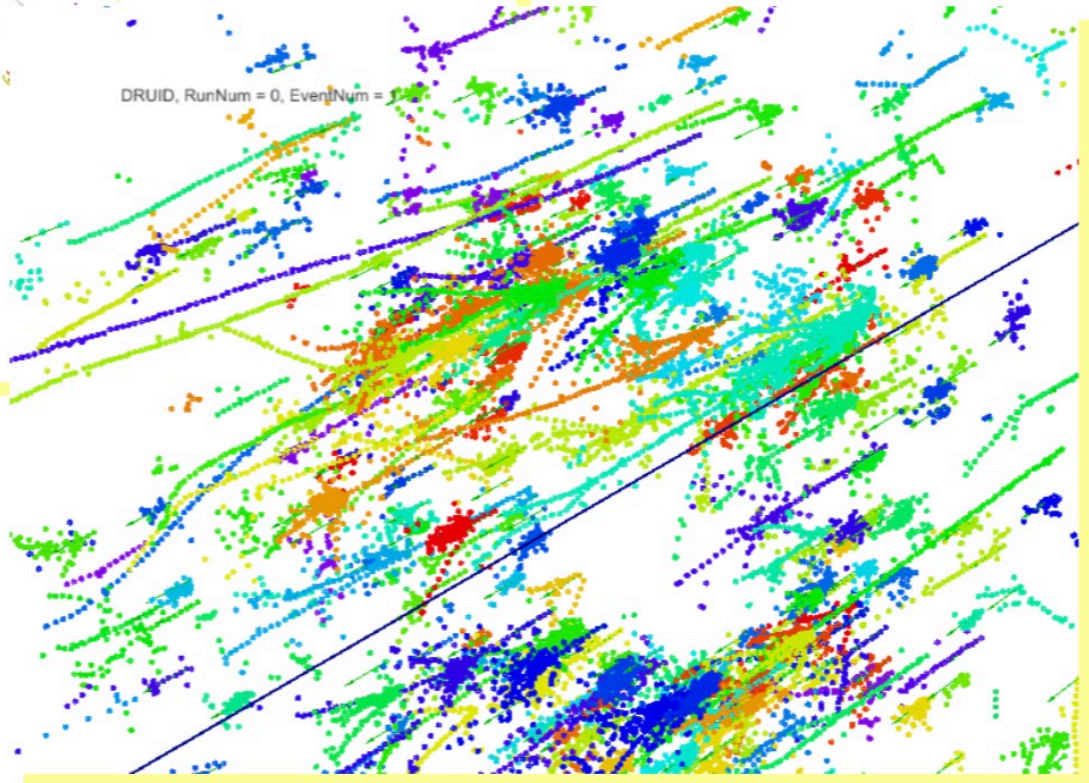
SPPC detector: exploration



C. Young: Solenoid + Dipole pairs for 100 TeV pp collider







Messy Environments

Overlap 20
Min-bias events

需要向工业界学习和寻求帮助

- Signal/image processing
  - Digital-Analog Conversions (including calibrations)
  - Pattern recognition, “clustering”
- Topological problems
  - Closest neighbor, minimum path, space partitioning
- Gaming *(our main source of inspiration!)*
  - “walk-through” complex 3D geometries
  - Detection of “collisions” (particles with surfaces!)
- Navigation/Avionics (Kalman filtering)
  - Tracking in a force field in presence of “noise”
  - Trajectory identification and prediction
- • Regression, classification, statistical analysis
  - Determination of physical parameters
  - Assign probabilities at various level of the data hierarchy
  - Statistic analysis with full data sample

Summary

- CEPC has a small scale storage and limited computing cluster
- A prototype software tools can meet the demand of detector optimization
- But not for a real experiment

- Real challenge comes from SppC stage
  - Huge amount data
  - Complicated event type
  - Computing model & algorithms
- We need help and corporation from computing experts, as well as from industry

Extras