The International Linear Collider

Status & Computing Plans

Marcel Stanitzki, Jan Strube

11/02/2015

Precision for the Terascale

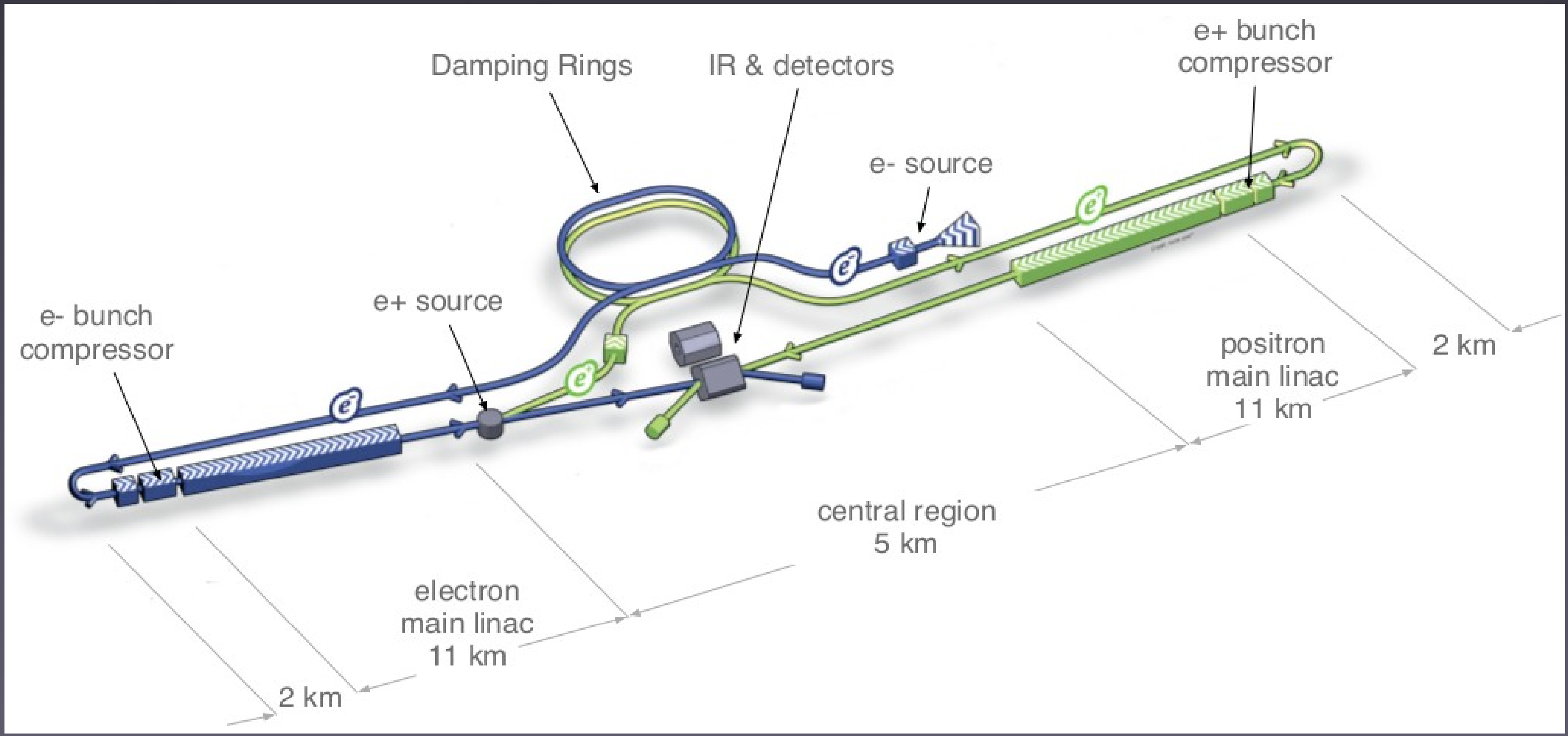
- To complement the LHC/HL-LHC a precision machine is required
  - Only e^+e^- colliders offer the required precision (1%)
- e^+e^- Advantages
  - Well defined initial state and tunable E_{CMS}
  - Clean environment, no QCD backgrounds
  - All background process can be calculated with high precision
  - low radiation environment

Why a Linear Accelerator

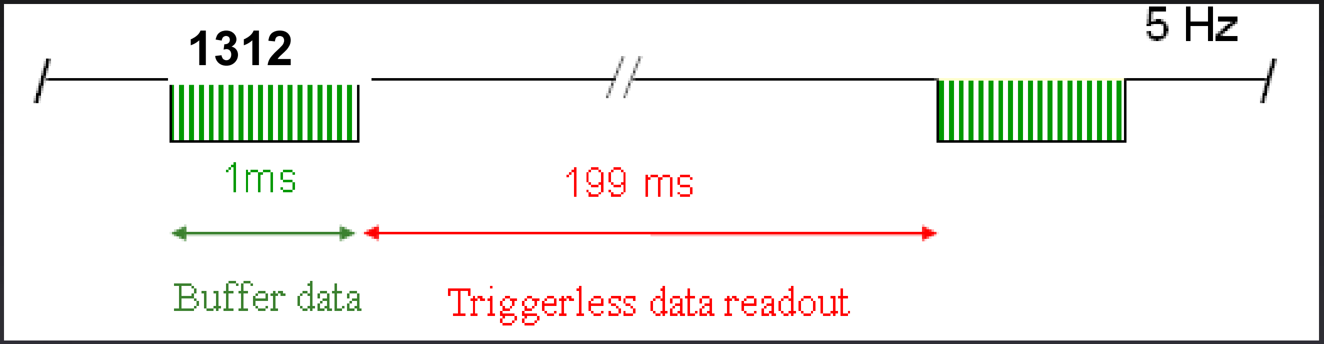
- Basic Limitations e^+e^- synchrotons
  - Synchrotron radiation loss ~ E^4/r
  - Synchrotron cost ~ quadratically with Energy (B. Richter 1980)
    - E_{CMS}=~ 200 GeV as upper limit
- A Linear Accelerator offers a clear way to higher energy
  - Not limited by synchrotron radiation
  - Cost ~ linear with Energy
  - Polarization of both beams
  - "nano beamspot" allows detectors close to the IP → key for c-tagging

The ILC Project

- The ILC (International Linear Collider)
  - A 500 GeV (baseline) GeV e+e- Linear Collider
  - Upgrade Path to 1 TeV
  - Polarization of both beams
- Interaction Region with two detectors

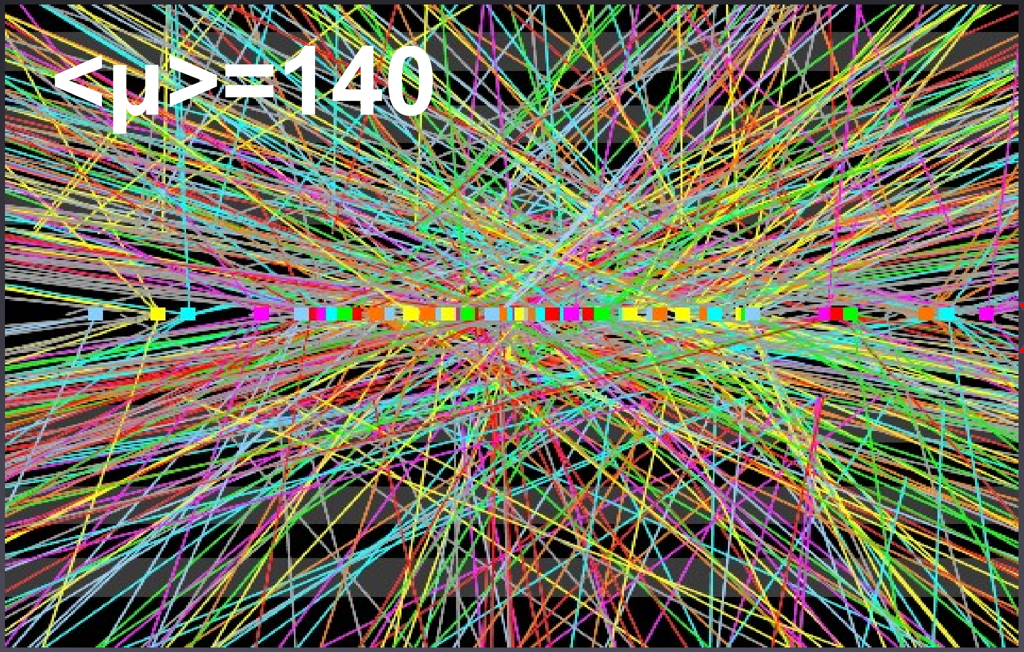


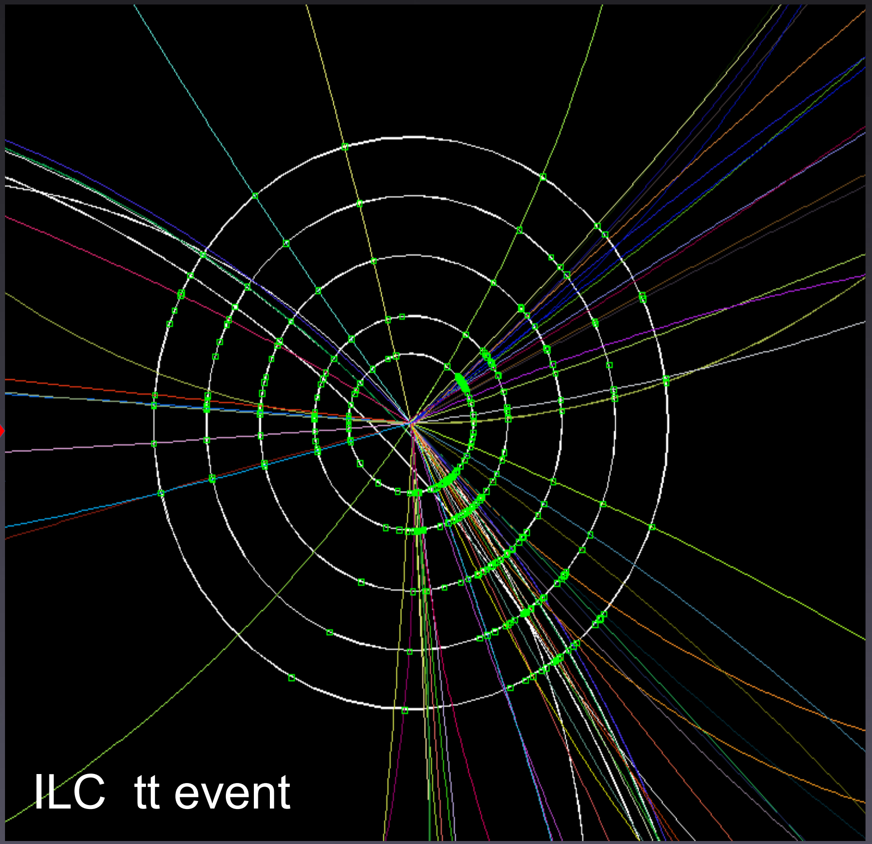
ILC Environment



- ILC environment is very different compared to LHC
  - Bunch spacing of ~ 554 ns (baseline)
  - 1312 bunches in 1 ms
  - 199 ms quiet time
- Occupancy dominated by beam background & noise
  - ~ 1 hadronic Z ($e^+e^- \rightarrow Z \rightarrow q\bar{q}$) per train ...
- Readout during quiet time possible

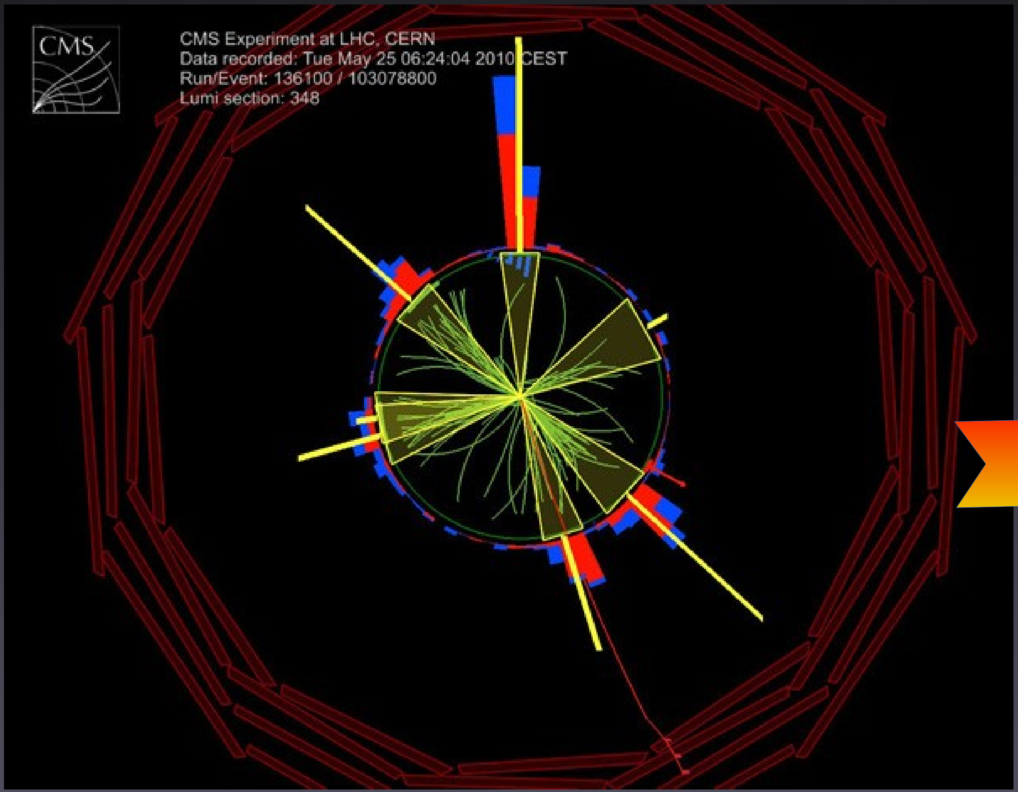
Tracking from HL-LHC to ILC

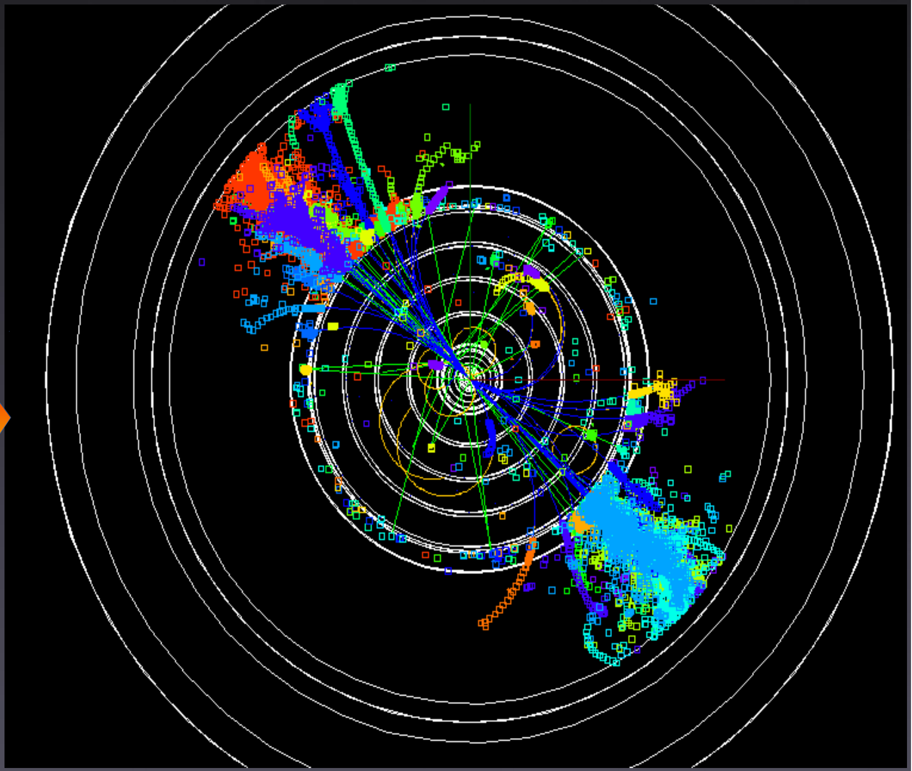




Moving from 140 interactions per crossing to ~1 event/train

Calorimetry from LHC to ILC

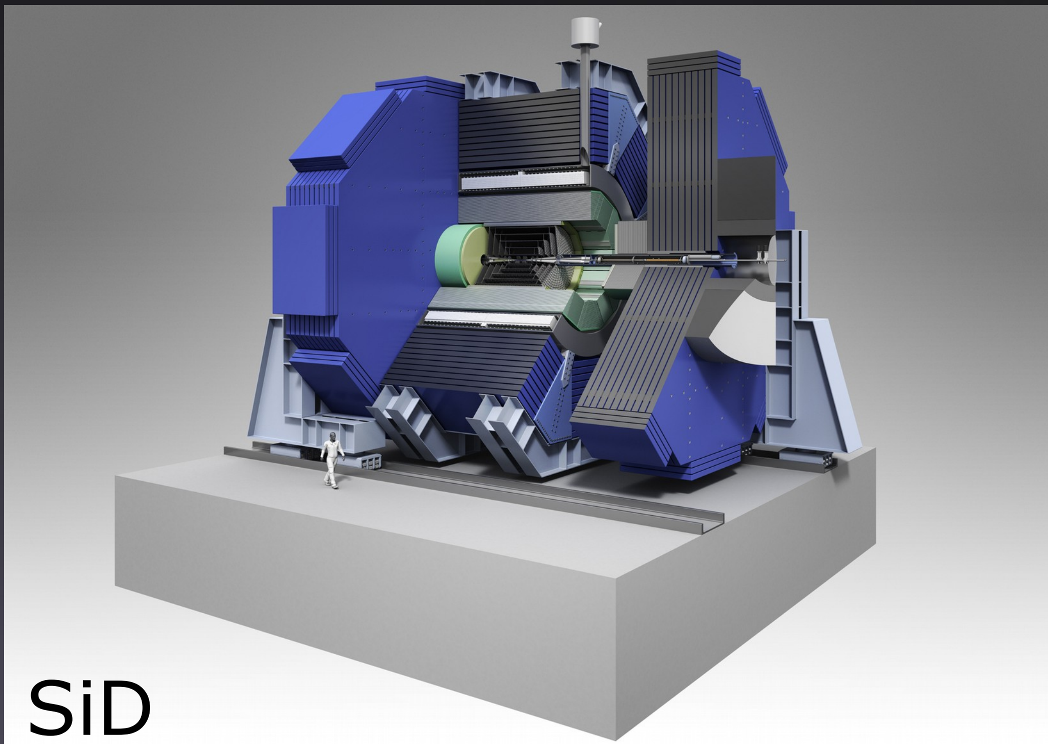


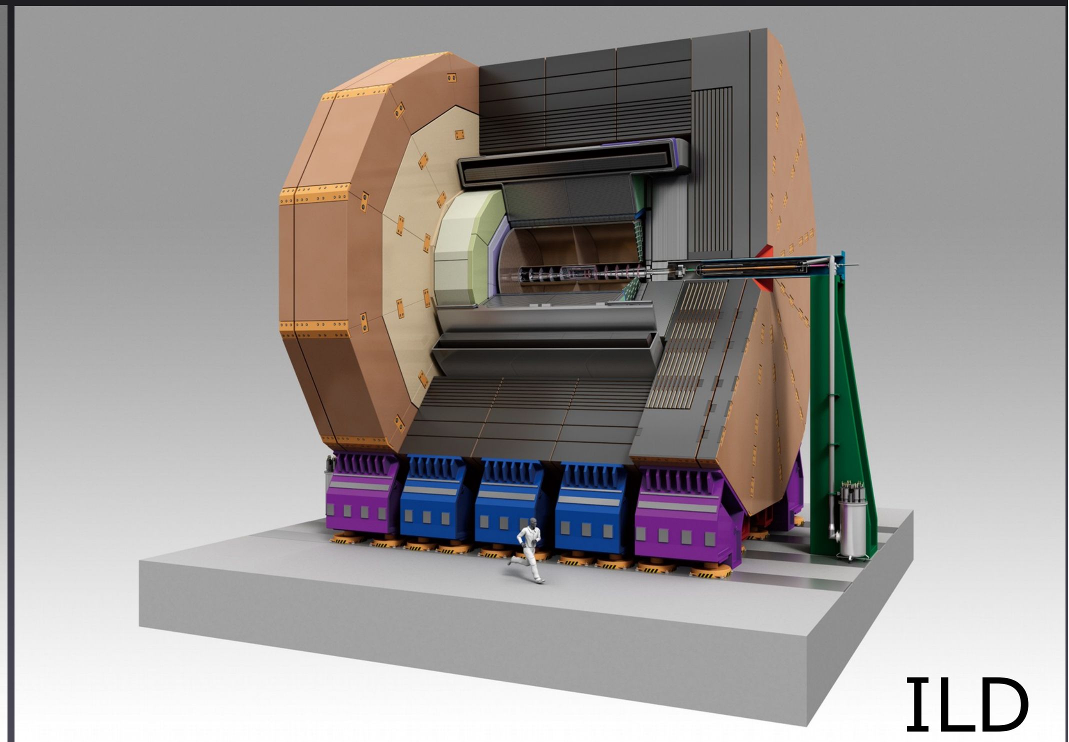


LHC Today

ILC

SiD & ILD

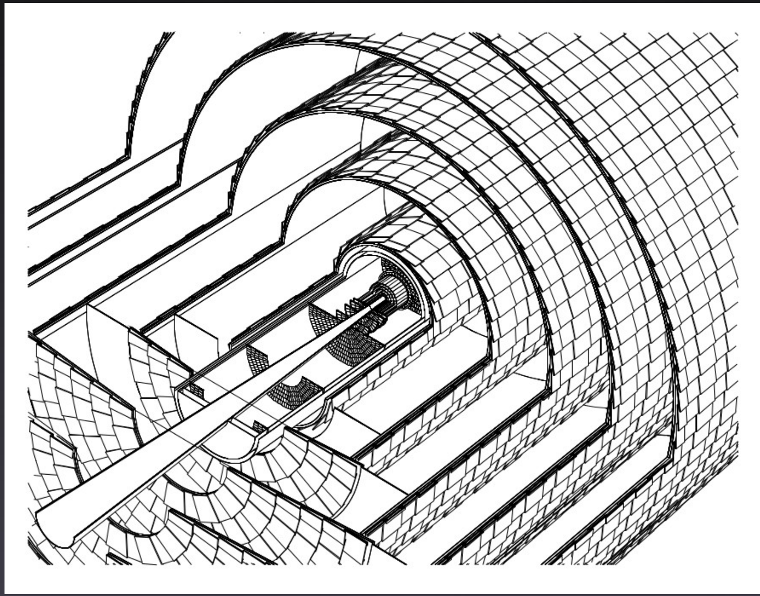


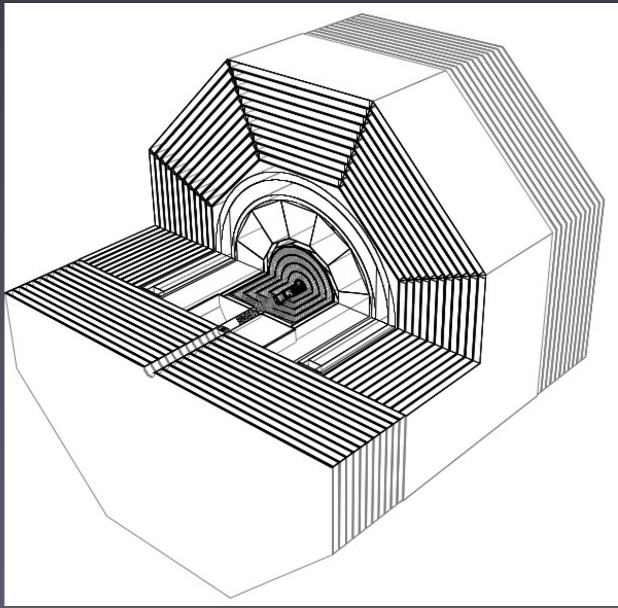


- SiD
  - $r_{Tracker}$=1.25 m
  - B = 5 T
  - All-silicon tracking
- ILD
  - $r_{Tracker}$=1.8 m
  - B = 3.5 T
  - Time Projection Chamber

Simulation & Reconstruction

- Full Simulation& Reco
  - Including full beam backgrounds
- Simulation
  - Detailed GEANT4 detector simulation
  - Including "dead areas"
- Full Reconstruction
  - Digitization, Tracking, Particle Flow, Flavor Tagging
  - No cheating at all





The ILC TDR

"As compared to other projects of similar scale (ITER, LHC, ATLAS, CMS, ALMA, XFEL, FAIR, ESS, SSC) the quality of the documentation presented by the GDE team is equal or superior to that utilized to launch into a similar process."

The ILC is good to go!

- ILC TDR (Technical Design Report)
  - Five Volumes covering
  - Physics, Accelerator & Detectors
  - Culmination of eight years of effort
- Very favorable review
- Wide Community support
  - 48 countries
  - 392 Institutes
  - 2400 signatories

SiD on the Grid

- Resources on the grid are shared within the ILC virtual organization (VO)
- VO members are (main sites)
  - ILD (DESY and KEK)
  - Clicdp (CERN / IN2P3)
  - SiD (PNNL, SLAC, everything else opportunistically)
- All three are currently running detector optimization campaigns
  - No major production effort on the horizon
- Storage
  - 150 TB at RAL (full, some used by Clicdp), 150 TB PNNL, 20 TB at SLAC
- CPU
  - ILC VO Total: ~10k CPU + PNNL

ILCDirac

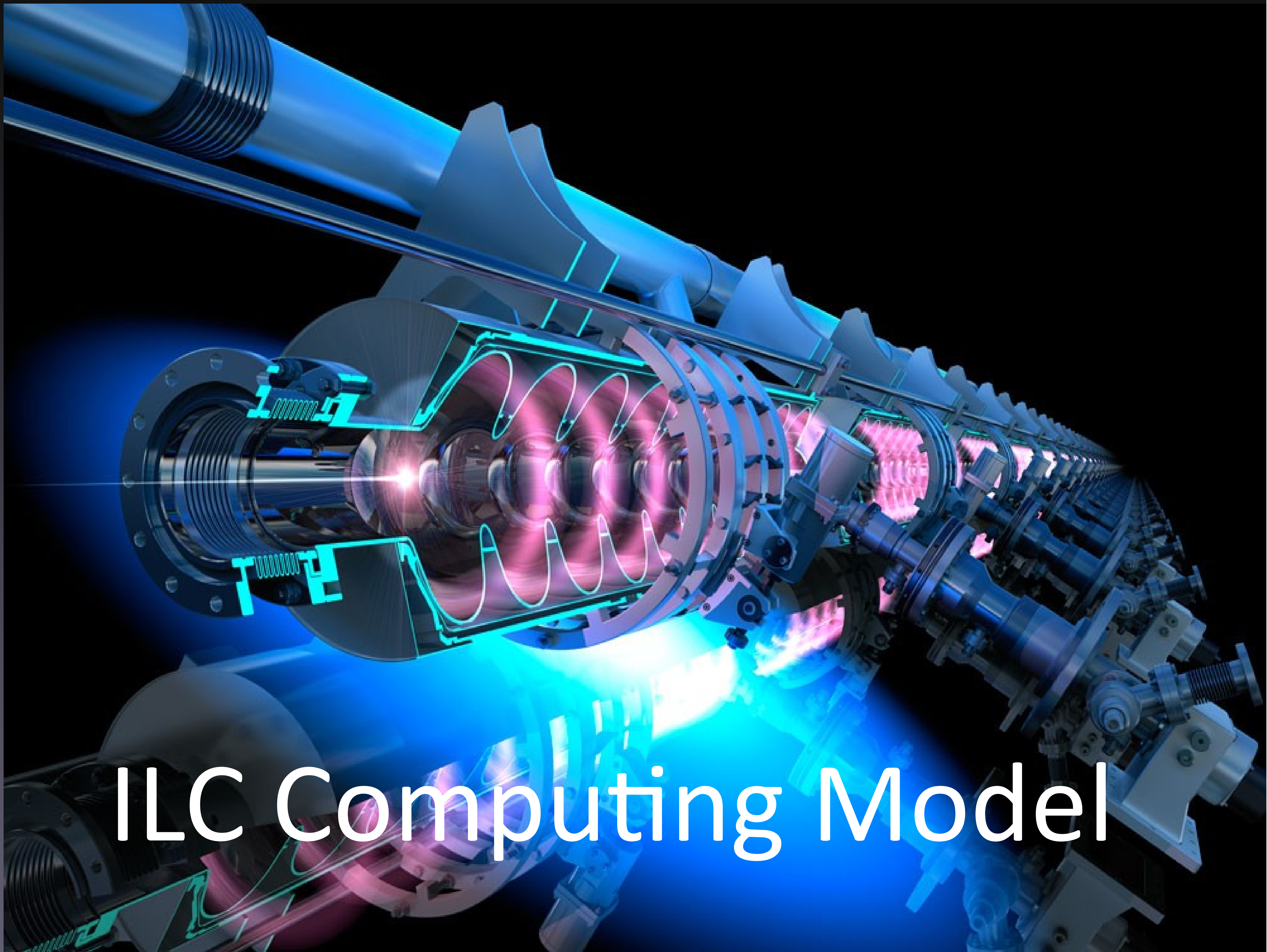
- An offspring of LHCb DIRAC package
- Submit scripts written in Python
- Web interface http://ilcdirac.cern.ch/DIRAC/
  - bookkeeping, restart failed jobs
- File Catalog
- Good Experience during TDR production
  - Reliable
  - "easy" to use

SiD on the Grid

dedicated resources: PNNL, SLAC, CERN
temporary quota increase: FNAL, RAL Tier 1
opportunistic use: all others

Worldwide LHC Computing Grid (WLCG) resources have been established during LOI and CLIC CDR efforts

SiD takes advantage of the international computing grid infrastructure


ILC Computing Model

Potential ILC Timeline

- 2016: Japan announces decision to host the ILC
- 2019: Completion of the Site-specific TDR
- 2020: Begin of Construction
- 2029: Commissioning
- 2030: Physics

A technology-driven Timescale…

Possible ILC Site

Planning the Computing Needs

- Cost of computing needs for the detectors was not included in the TDR.
  - Explicit request to specify computing costs
- For the LHC experiments, this was a very large item and (still is) a major source of headaches.
  - The Grid has been considered "The 5th experiment" at LHC.
- For ILC, we consider computing more like another sub-detector, not least because of the PFA paradigm.
- Understanding the requirements for CPU / storage / networking will be important for planning the computing center
  - Online vs. offline computing
  - Prompt reconstruction vs. distributed computing
  - Manpower / user support

SiD Data Rate

Table II-9.2 Overview of read-out details for the various subdetectors. Occupancies and data volumes are for a full bunch train at 1 TeV and include beam-induced background as well as charge sharing between pixels/strips. Safety factors of five and two have been applied to the rates of incoherent pairs and $\gamma\gamma \rightarrow$ hadrons respectively. BeamCal and Lumical are expected to be using the Bean chip with a buffer depth of 2820.

	cell size (mm^2)	number of channels (10^6)	av. to max. occ. (%)	approx. # bits per hit (bit)	data volume (Mbyte)
VXD barrel	0.02×0.02	408	8 - 60	32	130
VXD disks inner	0.02×0.02	295	4 - 70	32	50
VXD disks outer	0.05×0.05	980	0.5 - 20	32	20
Main tracker barrel	0.05×100	16	33 - 300	32	20
Main tracker disks	0.05×100	11	4 - 500	32	2
ECAL barrel	3.5×3.5	72	2 - 45	40	7
ECAL endcap	3.5×3.5	22	33 - 2300	40	36
HCAL barrel	10×10	30	0.07 - 200	40	0.1
HCAL endcap	10×10	5	96 - 3600	40	24
LumiCal	2.5×var.	0.061	≫100	16	340
BeamCal	2.5(5.0)×var.	0.076	≫100	16	430

1060

Assuming 2450 BX/train @ 5 Hz

→ 5.3 GB/s

Data Rates

- From the Detectors:
  - ~1.1 GB raw data @ 500 GB, both for SiD/ILD
- Assuming half of a 1.6 ⨉ 107 s year:
  - ~9 PB raw data / year / detector
- Slightly less at lower energies:
  - 350 GeV SiD: ~4.9 PB, ILD ~6.3 PB
  - 250 GeV SiD: ~3 PB, ILD ~5.5 PB
- Large uncertainties from estimation of beam backgrounds

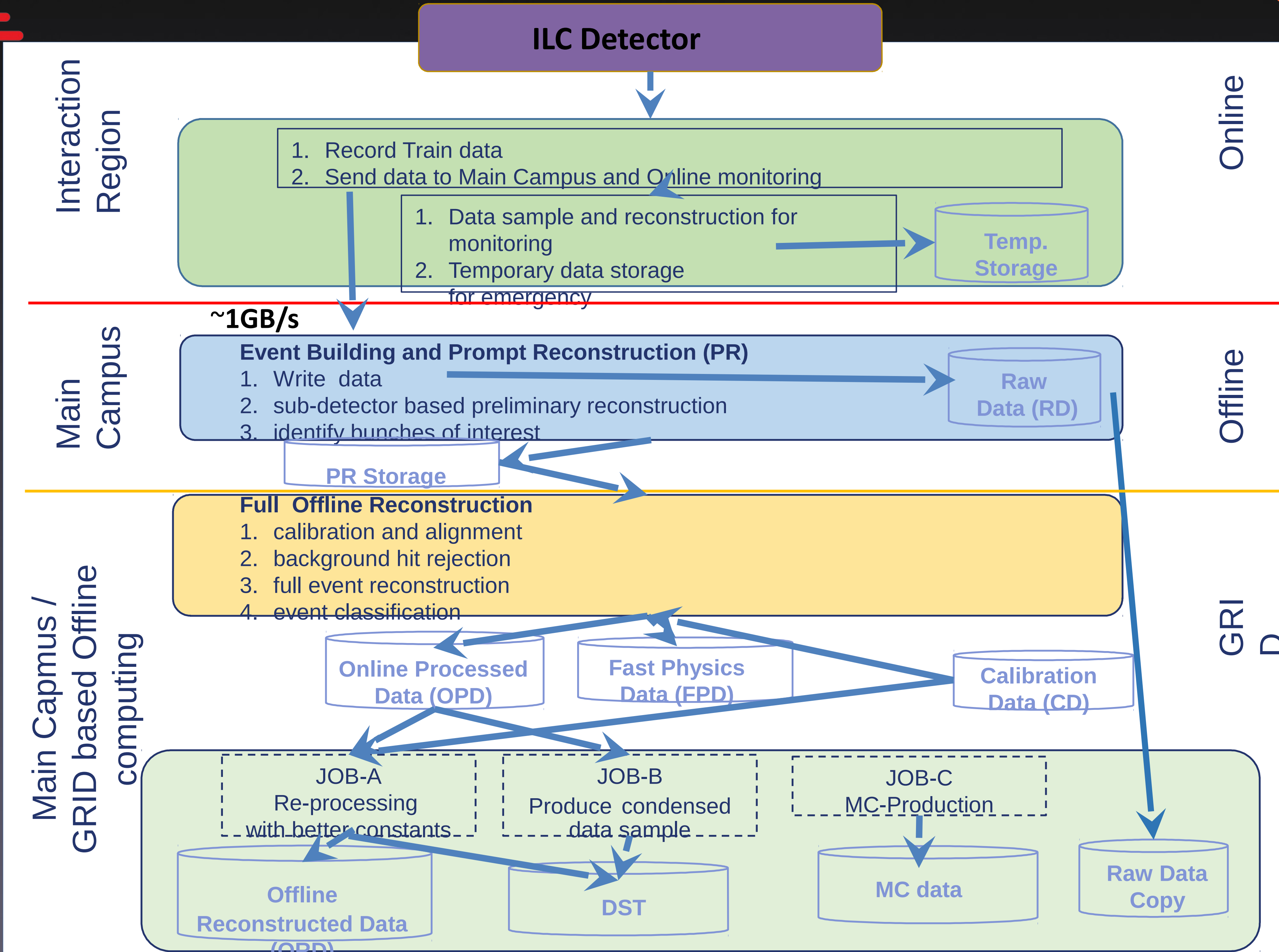
Current Strategy

- Move the data out of the mountain as quickly as possible
  - Need few days buffer for network outage
- Store a raw copy on campus
- Process (filter) events for analysis and ship to grid.
  - CPU needs for this being investigated. Filtering could happen on the grid. Under study.
  - A complete copy of the raw data needs to be shipped for redundancy anyway.
  - Plan to use existing “Tier-1” centers.
  - Probably want certain “streams” with fast feedback.

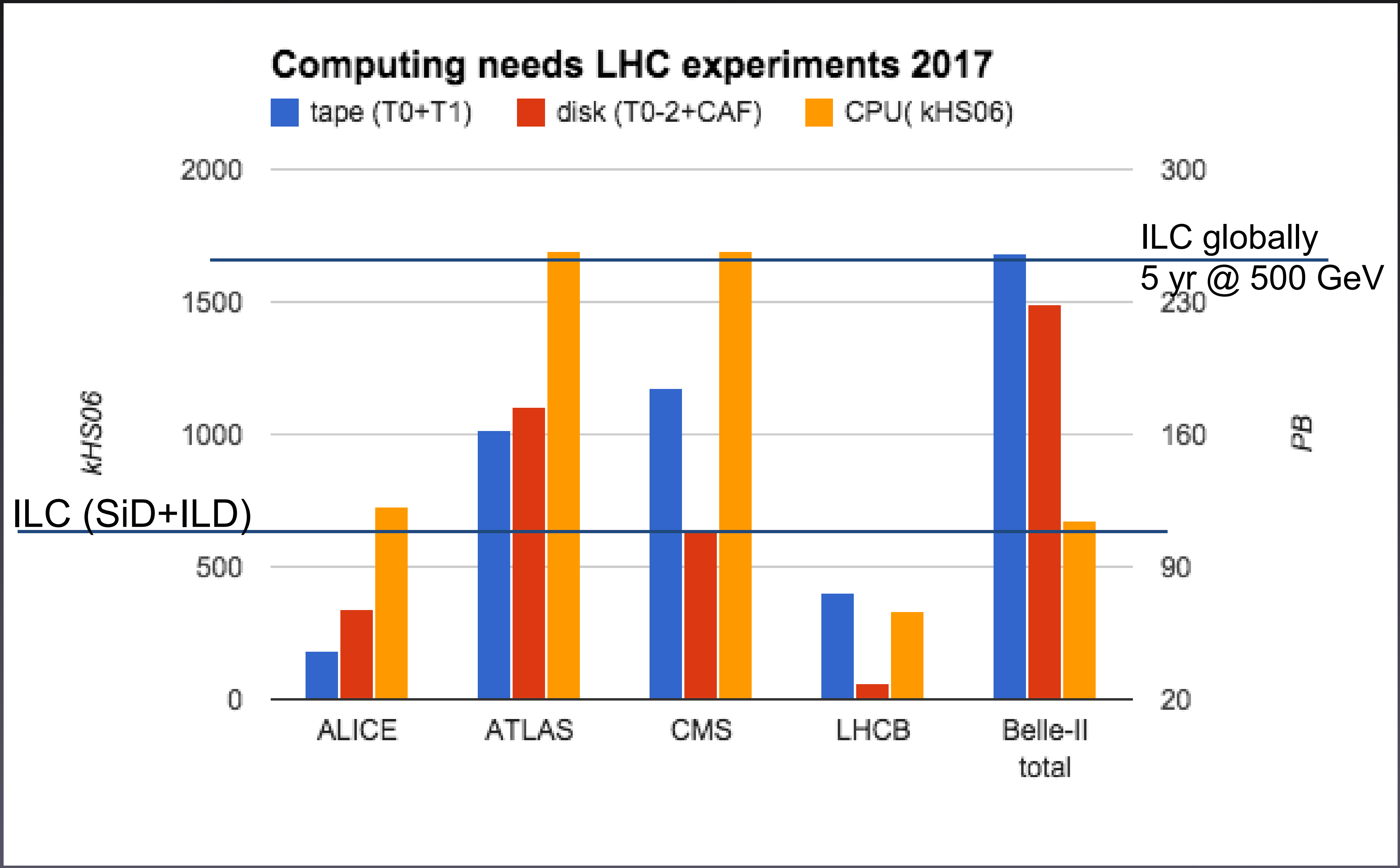
Caveats

Data rate looks scary at first sight: > LHC! But:

- This will be in ~2030. HL-LHC will have set a new bar by then.
- There will be many experiments with data rates in the same ball park by then.
- As mentioned, large safety factors on background. Using 5, while 2 is advertized by the machine.
- Data reduction techniques being investigated. (Size is mostly from pair background)


ILC Detector
Interaction Region
Online
1. Record Train data
2. Send data to Main Campus and Online monitoring
1. Data sample and reconstruction for monitoring
2. Temporary data storage for emergency
Temp. Storage
~1GB/s
Main Campus
Offline
Event Building and Prompt Reconstruction (PR)
1. Write data
2. sub-detector based preliminary reconstruction
3. identify bunches of interest
Raw Data (RD)
PR Storage
Main Capmus / GRID based Offline computing
GRI D
Full Offline Reconstruction
1. calibration and alignment
2. background hit rejection
3. full event reconstruction
4. event classification
Online Processed Data (OPD)
Fast Physics Data (FPD)
Calibration Data (CD)
JOB-A Re-processing with better constants
JOB-B Produce condensed data sample
JOB-C MC-Production
Offline Reconstructed Data (ORD)
DST
MC data
Raw Data Copy

ILC Computing Needs



Summary

- Input to Japanese Funding Agency is being prepared
- Large uncertainties on data rates and CPU
  - Data rates conservative, CPU unknown
- Data rates from detectors would compete with LHC today
  - > 10 years from now, our data rates will not require large additional investments
- Expected continued development of networks lets us take advantage of a distributed infrastructure
  - Allows granular contributions from Funding Agencies