ILC Computing Model

Frank Gaede, DESY
Birgit Lewendel, DESY
TAB Strategy Meeting for Data Intensive Science
Wuppertal, July 14-15, 2016
Outline

- Introduction ILC
  - Status of ILC in Japan

- Software Framework: iLCSoft

- Computing Strategy
  - Grid usage
  - iLCDirac production system

- Outlook
International Linear Collider - ILC

e+e- collider w/ SCRF
E = 250-500 GeV (1TeV)
length = 31 km (~50 km)

two detectors in push-pull operation: ILD and SiD
  - optimized for PFA
    - high hermiticity
    - highly granular calorimeters
    - excellent tracking and vertexing
ILC physics and detectors

- well known initial conditions and clean events at ILC
- offer a reach precision physics program:
  - Higgs, top, BSM

- detector requirements:
  - very high tracking efficiency
  - extremely good impact parameter and momentum resolution
  - unprecedented jet-energy resolution

- both detectors are optimized for particle flow PFA:
  - high hermeticity
  - highly granular calorimeters

\[ \sigma_{E_{jet}}^2 = \epsilon_{trk}^2 \sum_i E_{trk,i}^4 + \epsilon_{ECal}^2 E_{ECal} + \epsilon_{HCal}^2 E_{HCal} + \sigma_{confusion}^2 \]

\[ \epsilon_{trk} = \delta(1/p) \approx 5 \cdot 10^{-5}, \quad \epsilon_{ECal} = \frac{\delta E}{\sqrt{E}} \approx 0.2, \quad \epsilon_{HCal} \approx 0.5 \]
ILC in Japan

- site proposal by Japanese HEP community
- no official decision by Japanese government
- studies are ongoing on how to implement the ILC at this site
### Status of ILC in Japan

<table>
<thead>
<tr>
<th>Year</th>
<th>Event</th>
</tr>
</thead>
<tbody>
<tr>
<td>10-2012</td>
<td>Japanese association of high energy particle physicists proposes to host the ILC in Japan</td>
</tr>
<tr>
<td>5-2013</td>
<td>MEXT asks Science Council of Japan to review ILC</td>
</tr>
<tr>
<td>9-2013</td>
<td>Report by Science Council on ILC</td>
</tr>
<tr>
<td>5-2014</td>
<td>ILC advisory panel established by MEXT, started deliberations</td>
</tr>
<tr>
<td>6-2015</td>
<td>Interim report by ILC advisory panel, as a result start a group to study the human resources necessary to realize the ILC in Japan</td>
</tr>
<tr>
<td>x-2016</td>
<td>Expect final report from the advisory group on ILC in Japan</td>
</tr>
</tbody>
</table>

- It is clear that no final decision by the Japanese government can be expected before more is known on results from Run II at the LHC.

- Definitely expect a very clear statement (decision?) as input to the next European Strategy (2018/19).
ILC Software

- LC community has a long tradition of collaborating on common software tools
  - for ILC up to 4 different detector concepts in 3 regions and test beams and also CLICdp
  - allow users to share core software tools and focus an algorithm and analysis development
=> do not re-invent the wheel
- basic strategy:
  - use well defined and agreed upon interfaces
  - keep it simple (as simple as possible but no simpler)
  - be as light weight, modular and flexible as possible
- followed this strategy over the years in several projects
  - EUDET, AIDA, AIDA2020, (HSF)
  - recently scope also partly extended to beyond LC:
    - LHC, FCC, CEPC and neutrino
the iLCSoft framework

- well defined interfaces to Event Data and Detector Geometry
- define “Language” through which actual tools are used
- allows for flexible re-use of tools by many clients
  - DESY involved in almost all aspects of the framework
LCIO - common Event Data Model

- common event data model (EDM) and persistency for linear collider community
  - joined DESY and SLAC (and LLR) project
  - first presented @ CHEP 2003 (!)
- non-ROOT Object I/O
  - schema evolution, compressed records, pointers ...
  - EDM decoupled from I/O by interfaces
  - C++, Java (and Fortran)

- used by ILD, SiD, CLICdp and test beams for more than 10 years
- common EDM proven to be crucial for collaborative SW development across detector concepts
DD4hep - detector description for HEP

- generic (experiment independent) detector description toolkit for HEP
- one unique source of geometry for complete data processing
- joined DESY and CERN project in AIDA(2020)
- needed by LC - targeted at all of HEP (now also HSF project)
- based on Geant4 and ROOT
- component based architecture
- use only what you need
- simulation, reconstruction, alignment, conditions data,...

used by:
- ILD, SiD, CLICdp, FCC, CEPC, Calice
- LHCb for upgrade
iLCSoft builds and installations

- have adopted **CMake** build tool early on (2007)

- **ilcinstall**: python scripts for builds and installation of iLCSoft
  - complete installations (official releases)
  - customized installations for users, test beams, ...
  - can build all external tools (ROOT, Geant4, GSL, ...)

- reference installations in afs/cvmfs
  - SL6, gcc4.4 and gcc4.8 (C++11), ubuntu
  - can be used from everywhere (w/ compatible OS)
    - also for building user packages against releases
Grid services for ILC

- VOs ILC and Calice
  - VOMS hosted at DESY
  - also used by CLICdp

- software & configurations distributed via cvmfs
  - stratum0 operated ad DESY: `/cvmfs/ilc.desy.de/sw/`
  - stratum1 @ DESY, CERN and RAL

- VO ILC uses EGI resources in Europe and some OSG resources in US provided by LCG Grid sites

- using DESY-SE as “Tier0”
  - tape backed dCache for large MC campaigns and test beam data
  - disk based for user data and more volatile files
Grid production system

- Dirac: production interface
  - originally developed for LHCb
  - uses pilot jobs on the Grid
  - provides data catalogue

- iLCDirac:
  - provides interface for running iLCSoft specific applications
  - handling of ILC specific meta data

- operated by CERN CLICdp group for the VO ILC
  - used by CLIC
    - mostly CERN-SE
  - and ILC
    - mostly DESY-SE
resource usage VO ILC

- 4 Million jobs, 2500 CPU years
  - Peak of about 15 to 20 thousand jobs
  - Mostly CLICdp productions
- No very long lasting activity
  - Not using up all resources yet
**VO ILC - resource usage - ILC**

- used Grid resources for several large Monte Carlo and test beam campaigns
- \(O(10^{e8})\) ILD/Calice events stored in DESY Grid-SE
- used as T0 for ILD and Calice:
  - >200 TByte disk, 150 TByte on tape
- last 12 months
  - \(\sim 780\) kSi2K*y

- **small** (1-2%) compared to LHC - but **crucial** for ILC
near term plans

- ILD has recently adopted DD4hep as new geometry model and simulation tool
  - also used for SiD and CLICdp
- allows for more flexible scaling of detector geometries
- adopted the track reconstruction to use an automatically generated reconstruction geometry from the detailed simulation model
  - using surfaces attached to volumes
  - averaging material properties
- plan to have “large” Monte Carlo production for ILD in first half of 2017
  - comparing several detector variants
  - $O(10^7)$ events
  - exact resource estimate to be done
  - expect to be small compared to LHC
Computing resource estimate for ILC

- first version of ILC computing model developed in 2014
- assuming a raw data rate of 1-1.5 GByte/s
- very preliminary estimate shows that computing needs for ILC (first seven years) are going to be rather modest compared to LHC-Run 2 or Belle-II

<table>
<thead>
<tr>
<th>Item</th>
<th>unit</th>
<th>Belle-II</th>
<th>ATLAS</th>
<th>CMS</th>
<th>ILC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Date</td>
<td>year</td>
<td>2016-2022</td>
<td>2015-17</td>
<td>2015-17</td>
<td>2029-2036</td>
</tr>
<tr>
<td>Storage</td>
<td>PB</td>
<td>485</td>
<td>337</td>
<td>295</td>
<td>250</td>
</tr>
<tr>
<td>CPU</td>
<td>kHEPSpec</td>
<td>1650</td>
<td>1695</td>
<td>1690</td>
<td>650</td>
</tr>
<tr>
<td>Raw data rate from detector to the main campus</td>
<td>GB/s</td>
<td>1.8</td>
<td></td>
<td></td>
<td>1.5</td>
</tr>
</tbody>
</table>
Summary & Outlook

- iLCSoft: simulation and reconstruction software for linear collider studies (ILC and CLIC)
- develop as much as possible common software tools
  - also shared beyond LC, e.g. CEPS, FCC, LHC, neutrino, ...

- VO ILC uses mostly EGI (and OSG) resources
  - VOMS, cvmfs and ILC-T0 provided by DESY
  - iL CDDLrac as production system (CERN/CLICdp)

- current resource usage small compared to LHC (1-2%)
- will grow considerably once a positive signal from Japan
- expect overall ILC computing resources to be rather modest compared to LHC Run-2 and Belle-II