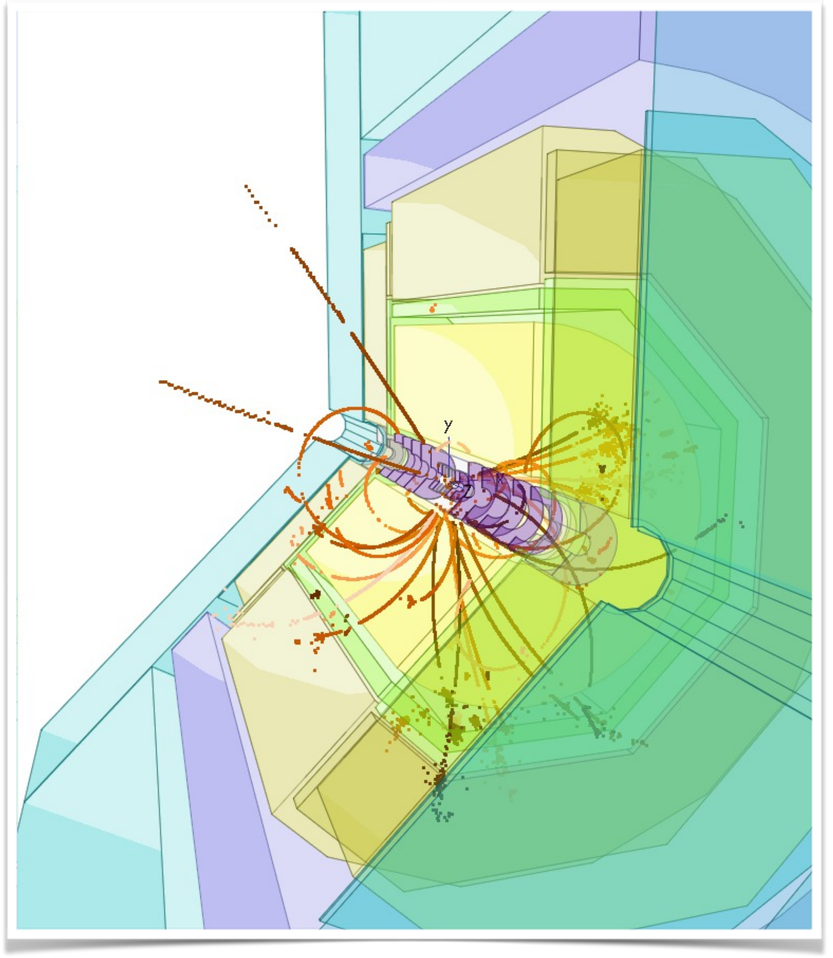
# Software Coordinator Report

## ILD Software and Analysis Meeting



**Frank Gaede, DESY, 04.11.20**

# Generator

M.Berggren, J.Tian

- just finishing the generation of 2f
  - only one job running
  - resubmitted twice, should finish today

- gamma-gamma:
  - real-real -> ll
  - huge cross section due to missing Q^2-cut for muons
  - to be fixed
  - need to decide on splitting the samples into different processes/directories
  - 2f samples with virtual gammas have smaller cross sections

  - also smaller 3f, 5f, … samples need to be done

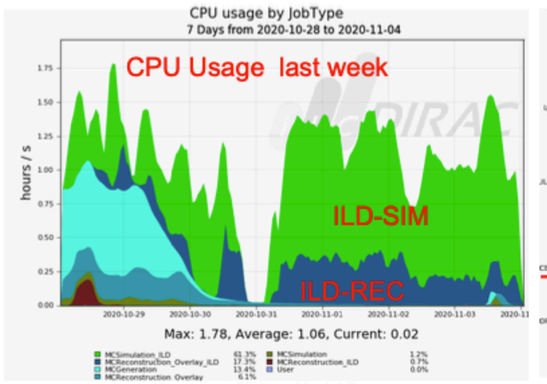
- last set of higgs samples done
  - exclusive higgs channels
  - all channels with 9 decays and 4 polarizations
  - O(100k) events per sample

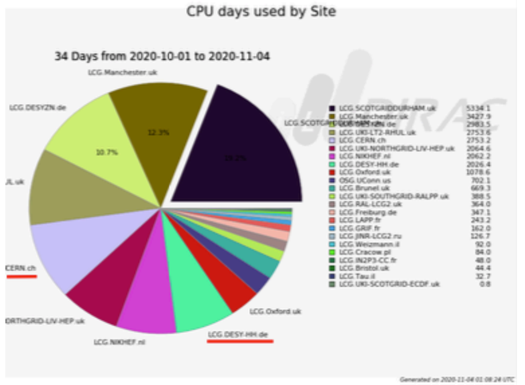
  - next step will be 6f samples
  - somewhat lower priority

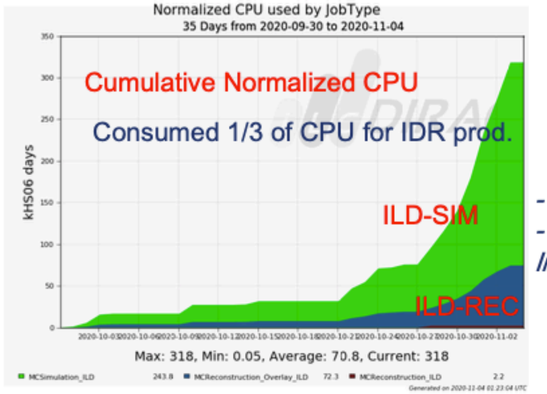
# Monte Carlo Production

A.Miyamoto, H.Ono

- Produced so far ( in progress, modifying production scripts.)
  - Latest information is in https://ild.ngt.ndu.ac.jp/elog/dbd-prod/
  - Flavortag samples,
    - 250 GeV 4q and ZH→vvqq : done
    - 91 GeV qq, nobg and overlay : done
  - 250 GeV SM samples,
    - Higgs samples : done
    - Low Xsect 3f,4f,5f, aa_4f samples : done
    - 4f-Mid Xsect samples : mostly done
    - 4f-High Xsect samples : next
    - 2f, additional 3f and 5f : next-to-next
- DST merged files
  - Replicas at DESY-SRM and KEK-DISK
  - /ilc/prod/ilc/mc-2020/ild/dst-merged/<Energy-Machinpara>/<EvtType>/<DetectorModel>/<ILDConfig>
  - Newer samples with large Xsect. :
    - Sub directories below <ILDConfig>, similar to REC, SIM files.
- REC files
  - Only 10% (2.5% flavortag) are saved initially. Now limited to 500 rec. files max
- SIM files
  - Initially DESY-SRM. Now KEK-SRM. Not able to keep them long.
- DST and gensplit files : will be removed soon



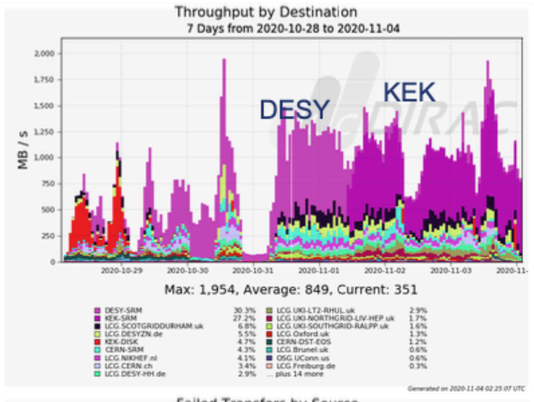



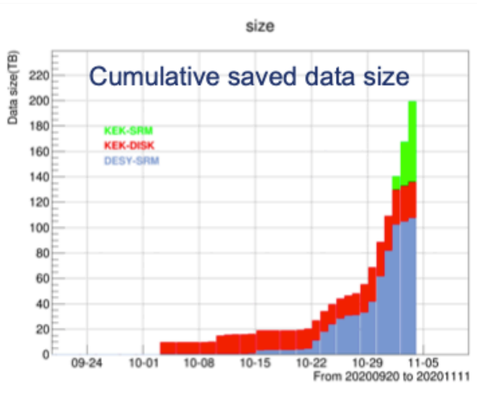



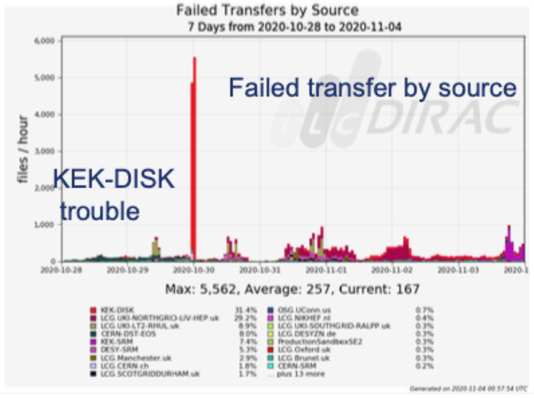

*- KEK CE is not yet used, because of ARC.*
*- DESY-HH and CERN were main resource In ILDIDR*

# Monte Carlo Production

A.Miyamoto, H.Ono




Data written since mid. October

| | Data Size (TB) | | | |
|---|---|---|---|---|
| | Total | DESY-SRM | KEK-DISK | KEK-SRM |
| mc-2020/SIM | 165.1 | 79.1 | 22.4 | 63.3 |
| mc-2020/REC | 17.0 | 16.0 | 1.1 | --- |
| mc-2020/Gen | 6.7 | 5.2 | 1.5 | --- |
| mc-2020 total | 199.7 | 107.4 | 28.6 | 63.3 |

## Resource needs for MC2020 (update 04-Nov)

Using samples available, estimated from 50 events of every processes.

- CPU=Sim + Rec ( Sim/Rec ~ 1/3 for higgs )
- DST = DST-Merged + DST.
  - DST could be removed )
  - Replica of DST-Merged at KEK-DISK and DESY-SRM
- REC = Keep only 10% (2.5%) of total ( 500 files max not take into account )
- SIM = Need to store in SE temporary. Originally planned to remove later.

| Process_type | CPU (k-days) | DST(TB) | REC(TB) | SIM(TB) | | SE |
|---|---|---|---|---|---|---|
| flavortag250(Done) | 1.37 | 0.82 | 0.38 | 12.92 | | KEK-DISK |
| SM(low Xsec)(Done) | 0.26 | 0.18 | 0.37 | 2.49 | | DESY-SRM |
| higgs(On-going) | 2.20 | 1.29 | 2.67 | 22.50 | | DESY-SRM |
| flavortag91 | 1.02 | 0.61 | 0.27 | 9.10 | | KEK-DISK |
| SM(med_Xsec) | 22.54 | 13.53 | 29.16 | 236.68 | | DESY-SMR → /KEK-SRM |
| SM(4f_high_Xsec) | 88.23 | 55.06 | 119.26 | 1034.12 | | KEK-SRM |
| Total | 118.62 | 71.47 | 152.10 | 1317.81 | | |

Available space in DESY and KEK

- Quota : 300TB (DESY-SRM, disk), 300TB(KEK-DISK, disk), ~500TB(KEK-SRM, tape)
- Used since Sep. 2020 : 11TB(DESY-SRM), 16TB(KEK-DISK), including generated, logs, etc.

New plan for SIM data

- SIM files have to be removed soon, if no more storage is provided.
  - Reconstruction with o2, o3 options requires significant CPU time for sim.
  - Do we remove all SIM files ? Any needs to keep some of them ?

5

need decision on what to do with SIM files -> see next slide

# Monte Carlo Production

A.Miyamoto, H.Ono

- impossible to keep all SIM files of the large 250 GeV production
  - unless significantly more storage resource would become available

- SIM files would be needed to
  - re-reconstruct the o1 technology option
  - reconstruct any of the alternative technology options o2 or o3

- agreement in software and physics working groups:
  - we do not need the full statistics with other technologies
  - also highly impractical to reconstruct the complete set several times

- proposal:
- keep the same amount of SIM files as for REC:  10% or 500 files/process maximum
- -> could later reconstruct some samples with alternative technologies
- -> need agreement from ILD