

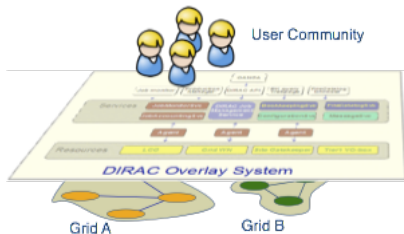
# Status of ILD Mass Productions Using `ilcdircac`

C. Calancha  
KEK  
LCWS15

November 5th, 2015

## DIRAC & ILCDIRAC

- DIRAC(\*) project is a complete Grid solution.
  - DIRAC forms a layer between a particular community and various compute resources.
- ILCDIRAC forms a layer on top of DIRAC providing the computing needs of linear collider experiments.



## ILD Mass productions with ILCDIRAC

- ILD has adopted `ilcdirac` for its mass productions.
  - Reliable, scalable, easy to use, good experts support.
  - Other experiments already using it: CLic, BelleII, Sid.

(\*) Distributed Infrastructure with Remote Agent Control

# What Has Be Done

## ILD productions with `ilcdircac`

- Since the last AWLC15 extensive test and improvement of the ILD modules has being done.
- We have shown that the **ILD modules** are **ready for mass production**.
  - Received several request to produce samples.
  - All of them were completed efficiently.

## Improving existing tools

- set/remove metadata faster/easier.
- Extending tools for monitoring the productions.

# ILD productions with `ilcdirac`

## Set up of ILD Production Module

- Output following ILD name convention.
- Test module for all ILD energy scenario: 250, 350, 500, 1000 GeV .
- Increased speed submission.

```
class ILDProductionJob( ProductionJob ):
    """ILD Production Jobs definition"""
    def __init__( self ):
        super( ILDProductionJob, self ).__init__()
        self.machine = 'ilc'
        self.experiment = 'ILC_ILD'
        self.basepath = self.ops.getValue( '/Production/%s/BasePath' % self.experiment, '/ilc/prod/ilc/mc-dbd/ild/' )
        self.polarization = ""
        self.machineparams = ''
        self.detector = ''
        self.compatmeta = {}
        self.processID = 0
        self.evtclass = ''
        self.evttype = ''
        self.genprocname = ''
        self.usesofttag = False
```

## Improved script setting the production chain.

- Full chain submission: generator split, simulation, reconstruction.
- Improved readability.

```
overlay.setBkgEvtType("aa_lowpt")
if energy==500.: #here you chose the overlay parameters as this determines how many files you need
  #it does NOT affect the content of the marlin steering file whatsoever, you need to make sure the values
  #there are correct. Only the file names are handled properly so that you don't need to care
  overlay.setBXOverlay(BXOverlay)
  overlay.setGGToHadInt(GGToHadInt500)
elif energy == 1000.:
  overlay.setBXOverlay(BXOverlay)
  overlay.setGGToHadInt(GGToHadInt1000)
elif energy == 350.:
  overlay.setBXOverlay(BXOverlay)
  overlay.setGGToHadInt(GGToHadInt350)
elif energy == 250.:
  overlay.setBXOverlay(BXOverlay)
  overlay.setGGToHadInt(GGToHadInt250)
else:
  print "Overlay ILD: No overlay parameters defined for this energy"

##Reconstruction ILD with overlay
mao = Marlin()
mao.setDebug()
mao.setVersion(MarlinVer) ##PUT HERE YOUR MARLIN VERSION
if ild_rec_ov:
  if energy in [250.0, 350.0, 500.0, 1000.0]:
    mao.setSteeringFile("bbudsc_3evt_stdreco.xml")
    mao.setGearFile("GearOutput.xml")
  else:
    print "Marlin: No reconstruction suitable for this energy"
```

## Motivation

- Original DBD 6f\_ttbar samples at 500 GeV show duplicates.
- 6f samples at 500 GeV not enough statistics for several analysts.
- Real test to show ILD `ilcdirac` modules ready for mass production.

## Simulation/Reconstruction Conditions

- Large statistics:  $L = 15 \text{ ab}^{-1}$  per process.
  - Total N events:  $< \approx 31 \text{ Mevts}$ .
- DBD settings: `ildconfig v01-16-p05_500 (*)`
- No REC files: Just Mokka files and DST were produced.

(\*) With no modifications, in particular, same  $\gamma\gamma \rightarrow \text{hadrons}$  .



- Samples are available at KEK site.

**sim** = *DIR\_SIM* (new samples total size 56.9 TB)

- `/ilc/prod/ilc/mc-dbd/ild/sim/500-TDR_ws`

**dst** = *DIR\_DST* (new samples total size 1.7 TB)

- `/ilc/prod/ilc/mc-dbd/ild/dst/500-TDR_ws`

### sub-directories

6f\_eeWW  
6f\_llWW  
6f\_ttbar  
6f\_vvWW  
6f\_yyyyZ  
6f\_xxxxZ  
6f\_xxWW

(\*) *ilcdirac* samples 2 level directory deeper

```
DBD:
DIR_DBD=${DIR_SIM}/6f_llWW/ILD_o1_v05/v01-14-01-p00/
${DIR_DBD}/sv01-14-01-p00.mILD_o1_v05.E500-TDR_ws.I37000.P6f_llxyev.eR.pR-00001.slcio
NEW:
DIR_NEW=${DIR_SIM}/6f_llWW/ILD_o1_v05/v01-14-01-p00/00004863/000
${DIR_NEW}/sv01-14-01-p00.mILD_o1_v05.E500-TDR_ws.I108588.P6f_llxyev.eR.pR_sim_4863_13.slcio
```

ProductionMonitor

Select All  Select None  Start Stop Flush Complete Clean

Selections

Status: Active, Stopped

AgentType: All

Type: All

Group: ILD-DBD\_higgs\_ffh\_250.0

Plugin: All

Date: YYYY-mm-dd

ProductionID:

RequestID:

Submit Reset

ID	Status	Type	Files	Processed (%)	Total Created	Done	Failed
Group: ILD-DBD_higgs_ffh_250.0							
6045	Active	MCRReconstruct...	100	100.0	120	100	20
6044	Active	MCRReconstruct...	100	100.0	112	100	12
6041	Active	MCRReconstruct...	100	100.0	116	100	16
6039	Active	MCRReconstruct...	100	100.0	101	100	1
6037	Active	MCRReconstruct...	100	100.0	117	100	17
6035	Active	MCRReconstruct...	100	100.0	102	100	2
6043	Active	MCSimulation	100	100.0	104	100	4
6042	Active	MCSimulation	100	100.0	134	100	34
6040	Active	MCSimulation	100	100.0	109	100	9
6038	Active	MCSimulation	100	100.0	105	100	5
6036	Active	MCSimulation	100	100.0	114	100	14
6034	Active	MCSimulation	100	100.0	104	100	4
Group: ILD-DBD_higgs_ffh_350.0							
6055	Active	MCRReconstruct...	100	100.0	103	100	3
6054	Active	MCRReconstruct...	100	100.0	101	100	1
6052	Active	MCRReconstruct...	100	100.0	118	100	18
6050	Active	MCRReconstruct...	100	100.0	129	100	29
6047	Active	MCRReconstruct...	100	100.0	123	100	23
6033	Active	MCRReconstruct...	100	100.0	115	100	15

Current Statistics

Global Statistics

Page 1 of 1 Refresh Auto: Disabled Updated: 2015-11-01 03:50 [UTC] Items per page: 1000 Displaying 1 - 36 of 36

Jobs > Production monitor DIRAC: v6r11p25, ILC: v22r1 calanchac@ ilc\_prod (/C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino)

- $e^+e^-H, \mu^+\mu^-H (H \rightarrow \gamma Z), E = 250,350, 500 \text{ GeV}$
- Average error ratio 12 %.
- Main error source: one particular site with low efficiency.

# Improving existing tools

# Improved Tools Monitoring Productions

## Transformation Command Line Interfaz

- Function returning input data query of one production.
- Function returning output data query of one production.
- Filter productions by username/author.
- Get the summary of a list of productions.

## More Improvements

- Get production logs using transformation ID as input.
- Set/delete > 1 metadata field each time in the dirac filecatalog CLI.
- List files ordering by size, using human readable format.
  - dirac filecatalog CLI. and COMDIRAC.
  
- General improvements: not exclusive for `ilcdirac` , applied to `dirac` .

- Set/Remove unlimited metadata fields each time.

```
FC:/> meta get /ilc/user/c/calanchac/tests/setRemoveMeta
*Owner : calanchac
FC:/> meta set /ilc/user/c/calanchac/tests/setRemoveMeta Energy 500 Machine ilc
/ilc/user/c/calanchac/tests/setRemoveMeta {'Machine': 'ilc', 'Energy': '500'}
FC:/> meta get /ilc/user/c/calanchac/tests/setRemoveMeta
!Machine : ilc
*Owner : calanchac
!Energy : 500
FC:/> meta remove /ilc/user/c/calanchac/tests/setRemoveMeta Energy Machine
FC:/> meta get /ilc/user/c/calanchac/tests/setRemoveMeta
*Owner : calanchac
FC:/>
```

- Improved parsing of input options.
- Supported short and long options.
- Supported order by size, reverse order, human readable file size.

```

FC:/> ls -lSRH /ilc/user/c/calanchac/tests/

drwxrwxr-x 0 calanchac ilc_prod      0.0B 2015-09-15 01:00:10 setRemoveMeta
drwxrwxr-x 0 calanchac ilc_user      0.0B 2014-11-12 05:31:21 set_randomSeed
-rwxrwxr-x 2 calanchac ilc_user    23.4KiB 2014-10-29 11:34:45 Dirac-Dst-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user    33.1KiB 2014-10-29 11:29:40 Dirac-Dst-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user    50.2KiB 2014-10-31 01:16:15 Dirac-Dst-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user    67.9KiB 2014-10-29 11:24:57 Dirac-Dst-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user    69.8KiB 2014-10-29 11:34:13 Dirac-Dst-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user   199.7KiB 2014-10-29 11:25:15 Dirac-Sim-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user   484.1KiB 2014-10-29 11:29:58 Dirac-Sim-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user   531.5KiB 2014-10-31 01:15:57 Dirac-Sim-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user   551.3KiB 2014-10-29 11:34:31 Dirac-Sim-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
-rwxrwxr-x 2 calanchac ilc_user   569.4KiB 2014-10-29 11:35:04 Dirac-Sim-E1000-B1b_ws.P4f_size_l.Gwhizard-1_95.eL.pR.I200034.005
FC:/>

```

- Summary productions inside the transformation CLI.

```
(Cmd) summaryTransformations 6035 6036 6037
TransID      Type              Status      F_Total    F_Proc.(%)  F_Proc.    F_Unused
  J_Run      J_Done      J_Fail      J_Stalled
6035         MCREconstruction_Overlay  Stopped    100        100.0       100         0
  0          100         2           0
6036         MCSimulation      Stopped    100        100.0       100         0
  0          100         14          0
6037         MCREconstruction_Overlay  Stopped    100        100.0       100         0
  0          100         17          0
None
(Cmd)
```

- Production log files can be retrieved by production ID (-I, -ProdID below).

```
lirac-ilm-get-prod-log -F /ilc/prod/.../LOG/.../somefile
```

#### General options:

```
-o --option <value>      : Option=value to add
-s --section <value>     : Set base section for relative parsed options
-c --cert <value>        : Use server certificate to connect to Core Services
-d --debug                : Set debug mode (-dd is extra debug)
-h --help                 : Shows this help
```

#### Options:

```
-D --LogFileDir <value>  : Production log dir to download
-F --LogFile <value>     : Production log to download
-O --OutputDir <value>   : Output directory (default ./)
-I --ProdID <value>     : Production ID (default )
```



- Input/Output can be listed from the transformation CLI.

```
(Cmd) getInputDataQuery 6036
```

```
{'Datatype': 'gen', 'Energy': '250', 'MachineParams': 'TDR_ws', 'GenProcessID': '108716', 'Ma  
az', 'SoftwareTag': 'v01-14-01-p00', 'ProdID': 5973}
```

```
(Cmd)
```

```
(Cmd) getOutputFiles 6035
```

```
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_1.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_10.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_100.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_101.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_102.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
250-TDR_ws.I108714.Pe2e2h_az.eR.pL_rec_6035_11.slcio  
/ilc/prod/ilc/mc-dbd/ild/rec/250-TDR_ws/higgs_ffh/ILD_o1_v05/v01-16-p10_250/00006035/000/rv01-16-p10_250.s  
(Cmd)
```

- We can list productions by username/author.

```
(Cmd) getALLByUser calanchac Stopped
TransformationID      Status      AgentType      TransformationName
ionDate              AuthorDN
5893                  Stopped     Manual         37588_1000.0_split_test-1000gev_20151007_1
10-07 12:46:54      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
5894                  Stopped     Manual         37588_1000.0_ild_sim_test-1000gev_20151007_
10-07 12:47:44      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
5895                  Stopped     Manual         37588_1000.0_ild_rec_overlay_test-1000gev_2
10-07 12:48:32      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6032                  Stopped     Manual         108709_350.0_ild_sim_calancha_20151021_2
10-21 00:37:38      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6033                  Stopped     Manual         108709_350.0_ild_rec_overlay_calancha_20151
10-21 00:38:28      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6034                  Stopped     Manual         108714_250.0_ild_sim_calancha_" + genproces
10-21 02:14:25      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6035                  Stopped     Manual         108714_250.0_ild_rec_overlay_calancha_" + g
10-21 02:15:14      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6036                  Stopped     Manual         108716_250.0_ild_sim_calancha_" + genproces
10-21 02:19:31      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
6037                  Stopped     Manual         108716_250.0_ild_rec_overlay_calancha_" + g
10-21 02:20:19      /C=JP/O=KEK/OU=CRC/CN=CALANCHA Constantino
(Cmd)
```

## Summary

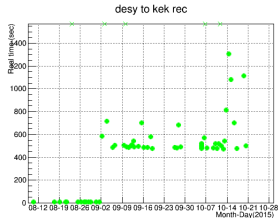
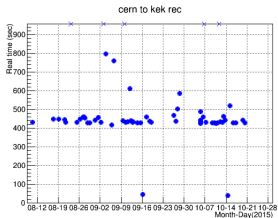
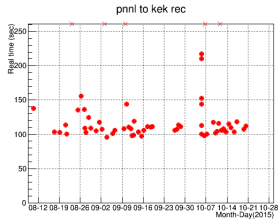
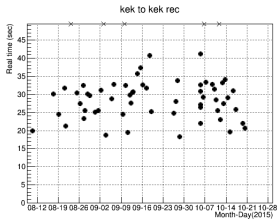
- ILD has adopted `ilcdircac` for its mass productions.
  - Three official request completed successfully using this new tool.
  - Ready to accept more request.
- Improving existing tools to monitoring the productions more effectively.

## Plan

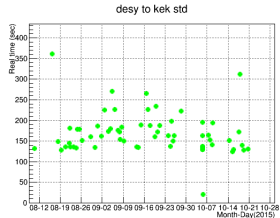
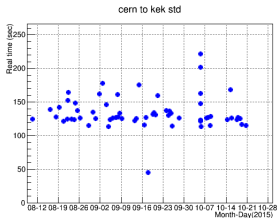
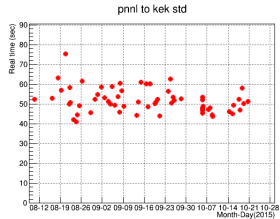
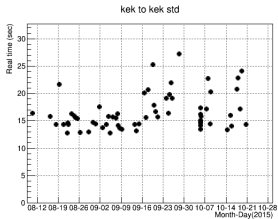
- Keep producing samples requested by ILD members.
- Keep improving/extending the tools monitoring the productions.
- Command line interfaz code.
  - Factorize to avoid duplication of code.
  - Extend/improve functionality.
  - Improve speed on listing LFNs.

# Back Up

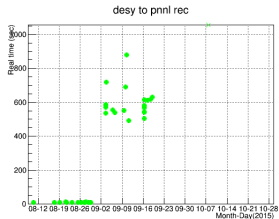
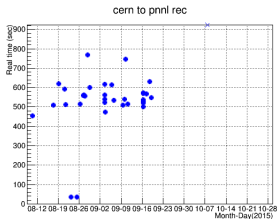
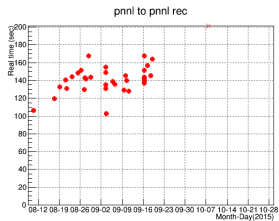
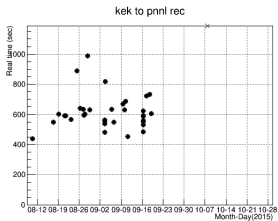
# Transfer Between Sites (Preliminary)



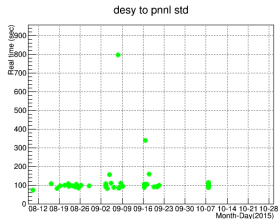
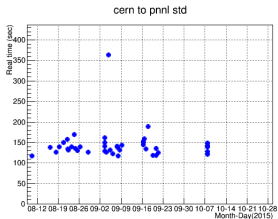
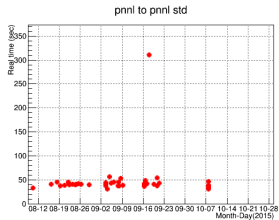
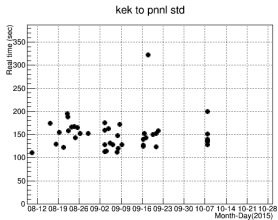
# Transfer Between Sites (Preliminary)



# Transfer Between Sites (Preliminary)

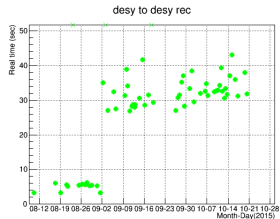
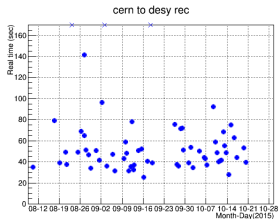
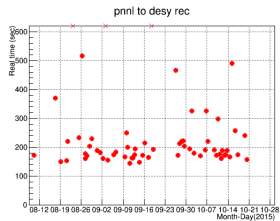
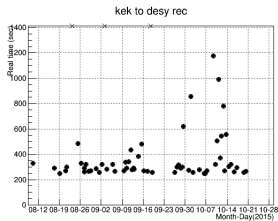


# Transfer Between Sites (Preliminary)

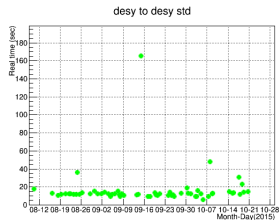
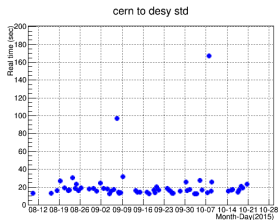
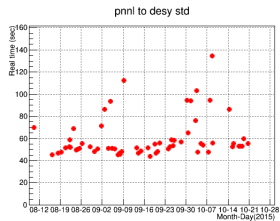
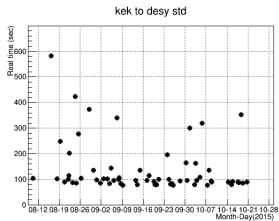




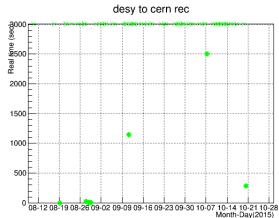
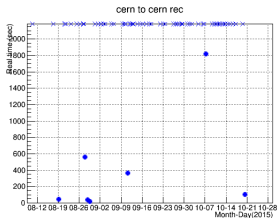
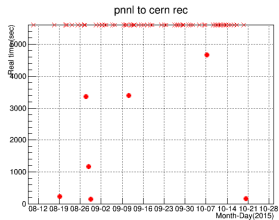
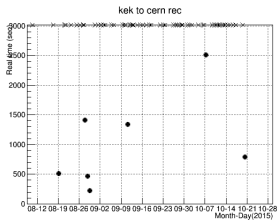
# Transfer Between Sites (Preliminary)



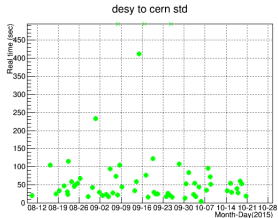
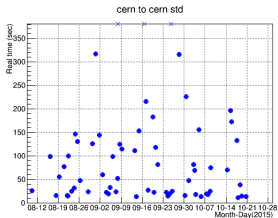
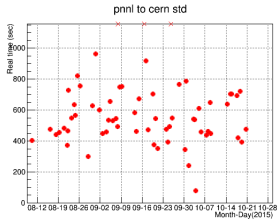
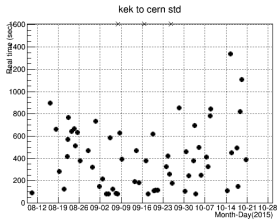
# Transfer Between Sites (Preliminary)



# Transfer Between Sites (Preliminary)



# Transfer Between Sites (Preliminary)



# Desy to KEK Parallel Transfer

- x axis spreads from July 24 00:00 JST to July 30 17:00 JST.

