

SiD DBD Production (DIRAC)

Christian Grefe, Stephane Poss,

Jan Strube (CERN)

Jeremy McCormick (SLAC)

Overview

Samples and setup

Introduction to ILCDIRAC and SiD presence
on the grid

Lessons Learned

Physics Processes

Samples provided by Common Generators Group (Barklow, Berggren, Miyamoto): all with the correct beamstrahlung spectrum at each energy

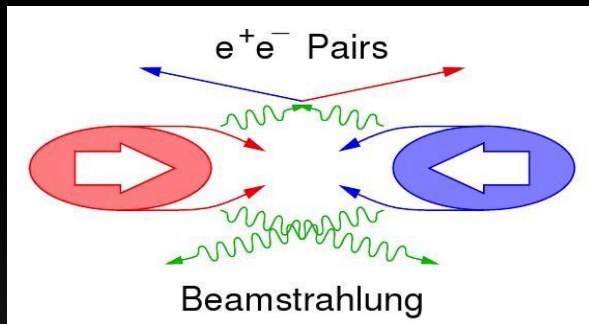
1 TeV (1000 fb⁻¹)

- $\nu\nu H$ signal ($m_H = 125$ GeV)
- $t\bar{t}H$ signal ($m_H = 125$ GeV)
- WW signal
- 1f - 8f SM background ($m_H = 2$ TeV)
- $\gamma\gamma \rightarrow$ hadrons
- Incoherent pairs

500 GeV (250 fb⁻¹)

- Top pairs, $m_{\text{top}} = 174.0$ GeV
- Top pairs, $m_{\text{top}} = 173.5$ GeV
- 6f SM background
- $\gamma\gamma \rightarrow$ hadrons

Beam-Induced Background

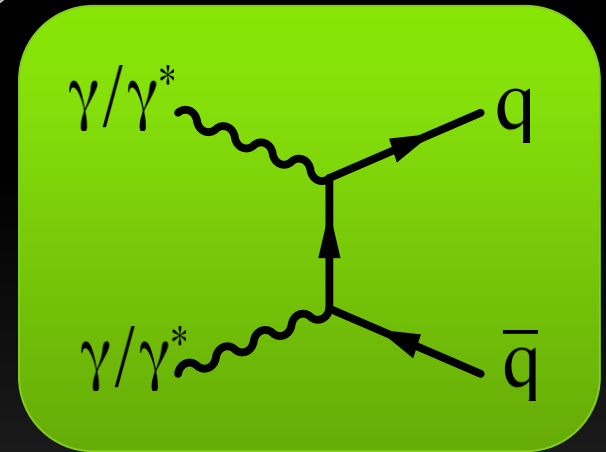


Pair background
1 event per BX
450k particles

Generated by
GuineaPig
ascii → hepevt →
stdhep

Merged with
every
“physics”
event

MCParticles
that don't
make hits will
be dropped



$\gamma\gamma$ interactions

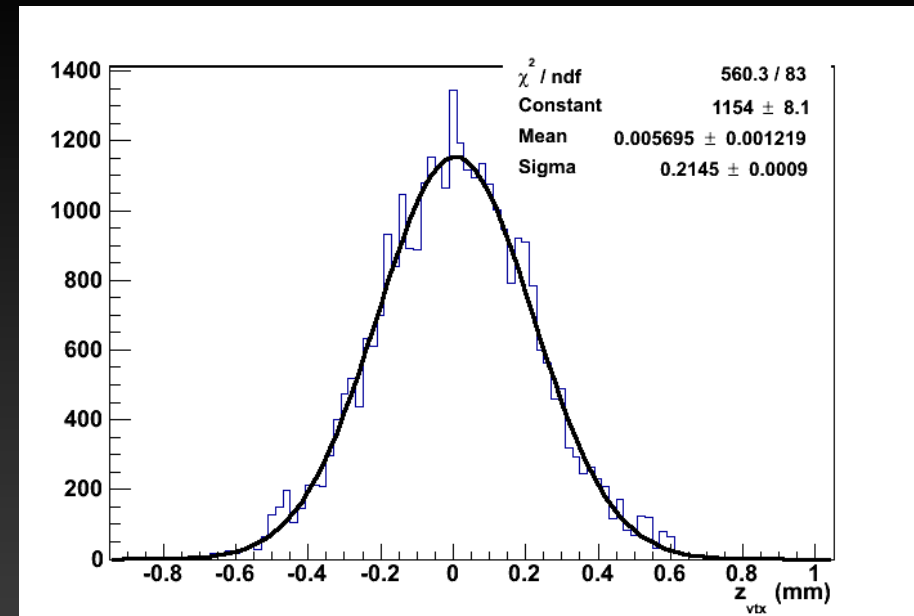
4.1 events per BX @ 1 TeV
1.7 events per BX at 500 GeV

Generated by Whizard

Luminous Region

- Finite extension: $\sigma_z = 225 \mu\text{m}$
 - conservative compromise
- Events from beam-beam interactions ($\gamma\gamma \rightarrow$ hadrons, incoherent pairs) are distributed randomly over the luminous region
- Physics events always at $z = 0$

Reconstructed primary vertex position for $\gamma\gamma \rightarrow$ hadrons, pairs



Fitted width: 214 μm

Sample Mixing (T. Barklow)

Samples provided 100% polarized

Samples for SiD grouped by processes and luminosity -
weighted

Loss of fraction of files does not cause analysis bias

Mixed to correspond to correct polarization

1 TeV (1 ab^{-1}): $\pm 80\%$ electron, $\mp 20\%$ positron

500 GeV (250 fb^{-1}): $\pm 80\%$ electron, $\mp 30\%$ positron

Sample Summary

2 Machine Energies x 2 Polarization settings

SM processes containing 1-8 fermions

1000 fb⁻¹ @ 1TeV, 250 fb⁻¹ @ 500 GeV

Background from incoherent pairs and hadron events from
photon interactions

66,421,842 events in 60 categories

[Production summary](#) on [SLAC confluence](#)

ILCDIRAC

- Dirac system used in LHCb in Production, consists of
 - File catalog (First used in CLIC CDR)
 - Supports meta data (file ancestry, detector model, ...)
 - Job submission, monitoring and bookkeeping
- ILCDIRAC (S. Poss et al.) developed for CLIC CDR production
 - Support for the plethora of ILC software
- Developed and maintained at CERN

Web-based Job Monitor

The screenshot displays a web-based job monitoring interface. The browser address bar shows the URL: `https://volcd01.cern.ch/DIRAC/ILC-Production/ilc_prod/jobs/ProductionMonitor/display`. The interface includes a sidebar with filters and a main table area.

ProductionMonitor

Select All Select None Start Stop Flush Complete Clean

Status: Active

AgentType: All

Type: All

Group: All

Plugin: All

Date: YYYY-mm-dd

ProductionID:

RequestID:

Request: 0

ID	Status	AgentT...	Type	Name	Files	Processed (%)	Created	Submitted	Waiting	Running	Done
1906	Active	Automatic	MCRReconstructi...	fix_tth-6q-hbb_...	243	0.0	823	823	0	39	0
1905	Active	Automatic	Replication	replicate_1820_...	30105	3.7	7	7	6	0	1
1904	Active	Automatic	Replication	replicate_1818_...	27792	1.6	6	6	6	0	0
1903	Active	Automatic	Replication	replicate_1817_...	49139	1.3	10	10	10	0	0
1902	Active	Automatic	Replication	replicate_1816_...	54390	2.5	11	11	11	0	0
1901	Active	Automatic	Replication	replicate_1814_...	58884	3.1	12	12	12	0	0
1900	Active	Automatic	Replication	replicate_1812_...	17234	0.0	4	4	4	0	0
1899	Active	Automatic	Replication	replicate_1791_...	10466	16.1	3	3	2	0	1
1898	Active	Automatic	Replication	replicate_1791_...	10466	5.8	3	3	3	0	0
1888	Active	Automatic	MCRReconstructi...	hh_nunu_3000_...	49	2.0	50	50	0	0	0
1887	Active	Automatic	MCSimulation	hh_nunu_3000_...	61	80.3	81	81	0	0	48
1886	Active	Automatic	MCGeneration	hh_nunu_3000_...	0	-	61	61	0	0	61
1885	Active	Automatic	MCRReconstructi...	hh_nunu_3000_...	69	4.3	73	73	0	0	3
1884	Active	Automatic	MCSimulation	hh_nunu_3000_...	91	75.8	123	123	0	0	64
1883	Active	Automatic	MCGeneration	hh_nunu_3000_...	0	-	91	91	0	0	91
1881	Active	Automatic	MCSimulation	aa_lowpt_500_...	47378	95.7	47589	47589	0	0	45
1880	Active	Automatic	Split	aa_lowpt_500_...	5357	98.7	5605	5605	0	0	52
1878	Active	Automatic	MCSimulation	aa_lowpt_500_...	45576	97.9	45440	45440	0	0	44
1877	Active	Automatic	Split	aa_lowpt_500_...	5332	99.1	5412	5412	0	0	52
1876	Active	Automatic	MCGeneration	aa qq e1e1 1...	0	-	1	1	0	0	1

Page 1 of 7 Refresh Auto: Disabled Updated: 2012-10-23 04:33 [UTC] Items per page: 25 Displaying 1 - 25 of 164

jobs > Production monitor jstrube@ ilc_prod (/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=jstrube/CN=678050/CN=Jan Fridolf Strube)

Meta Data Interface

The screenshot displays the Meta Data Interface web application. The browser address bar shows the URL: `https://volcd04.cern.ch/DIRAC/ILC-Development/ilc_prod/data/MetaCatalogue/display`. The interface is divided into three main sections:

- MetadataCatalog:** A list of metadata tags on the left, including EvType, NumberOfEvents, BXOverlaid, Polarisation, Datatype, Luminosity, Energy, MachineParams, DetectorType, Machine, ProdID (selected), Owner, SoftwareTag, DetectorModel, StartDate, and JobType.
- Metadata Query:** A search area in the middle with a "Path to start from:" field set to "/" and a "ProdID:" field containing the value "1880".
- File List:** A table on the right showing a list of files. The "File Name" column contains paths such as `/ilc/prod/ilc/sid/500_m80p20/aa_lowpt/GEN/00001880/003/aa_lowpt_m80p20_gen_1880_3000_001.stdhep`.

At the bottom of the interface, there are buttons for "Refresh", "Submit", and "Reset". The status bar at the very bottom shows the user information: `jstrube@ ilc_prod (/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=jstrube/CN=678050/CN=Jan Fridolf Strube)`.

Supported Software

Software is modularized in Dirac

Sets the context of the program (env vars, dependencies)

Allows to chain different modules together

Currently supported Physics applications:

Whizard, Pythia, Mokka, Marlin, PandoraPFA, SLIC, slicPandora, lcsim, etc.

Mix and match, supply your own steering files

ILC VO

Before Summer 2012:

US colleagues:

Open Science Grid

ILC VO managed at Fermilab



Grid authorization
prevents data
exchange

European / Asian colleagues:

Worldwide LHC Computing Grid

ILC VO managed at DESY

Now:

Virtual Organizations have been merged

Actively exchanging computing and storage resources

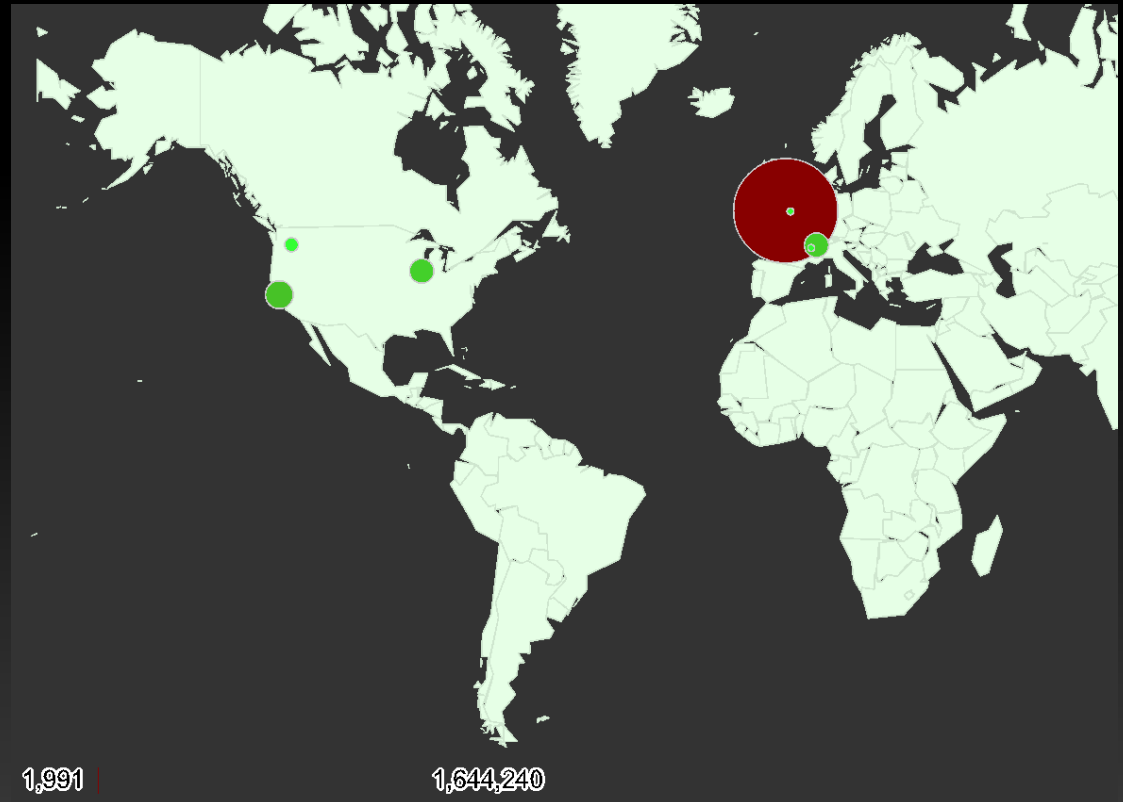
Both, OSG and WLCG sites supported in DIRAC through gLite

Storage Elements

RAL has been largest grid site for SiD since the LOI production

We could not have done the production without being permitted to use 300% of our allocation

PNNL has started making large storage and CPU resources available, able to host one copy of whole production



Total Storage: 127 TB -- 2,681,083 Files

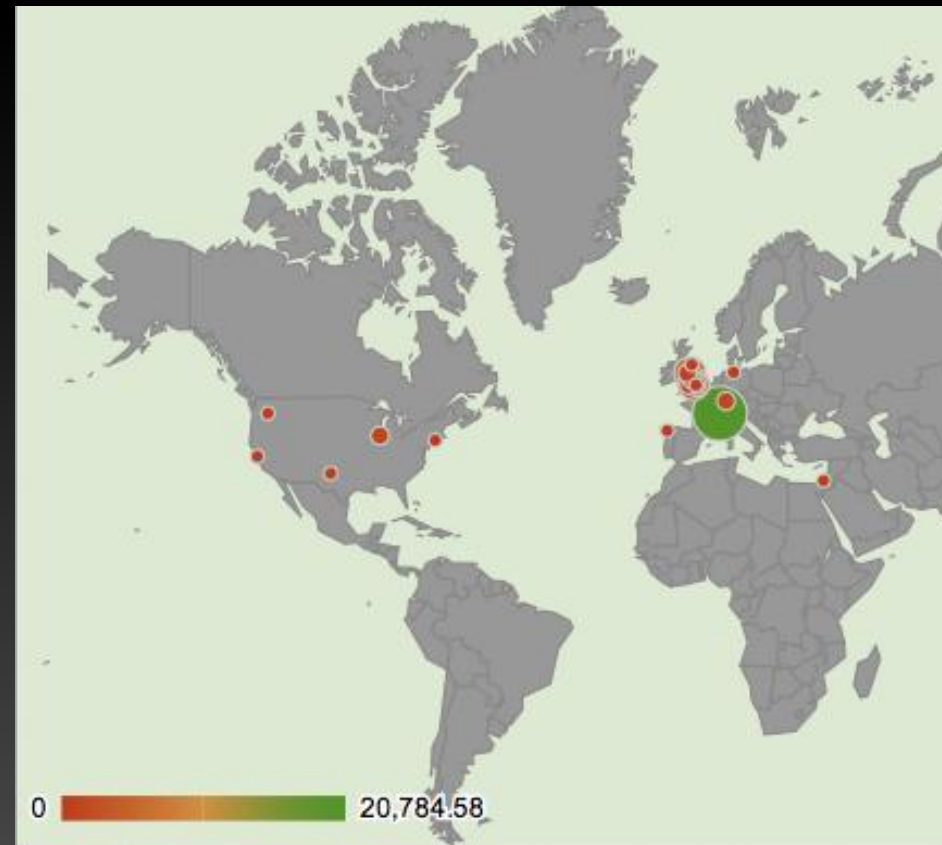
Computing Elements

During LOI: 70% of production at SLAC farm, other 30% used to gain experience at various grid sites across Europe

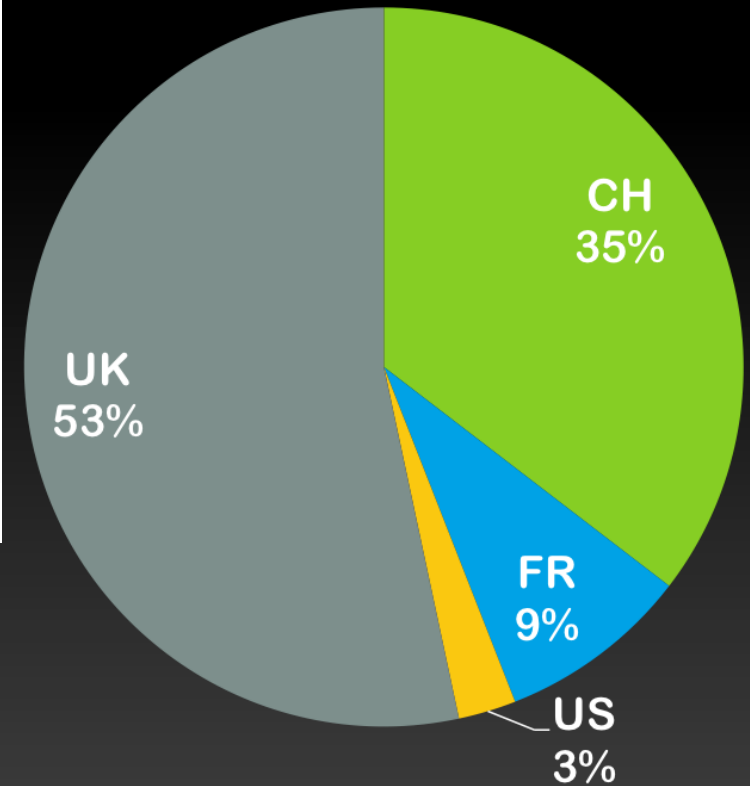
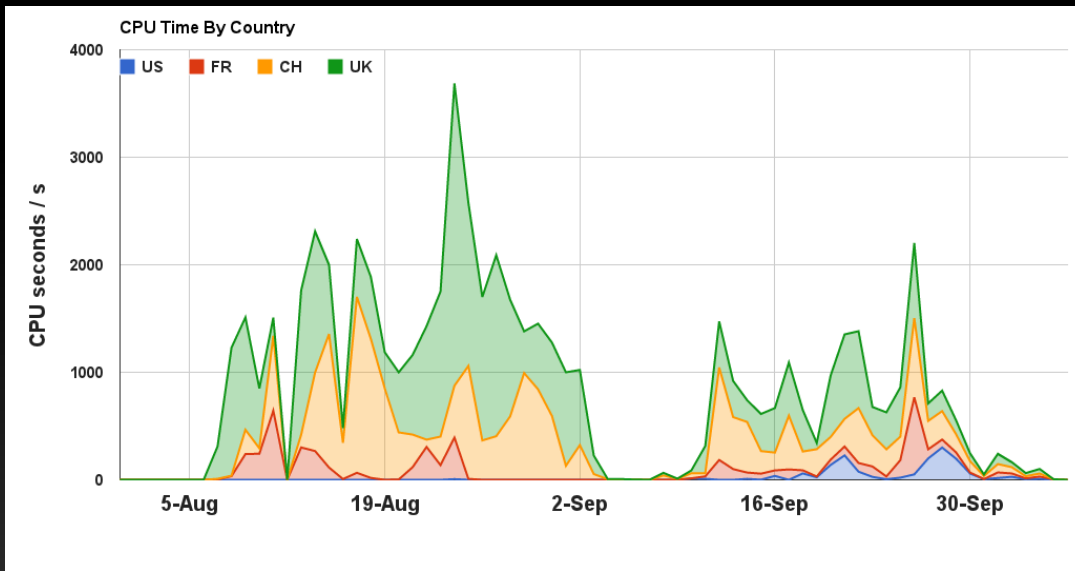
Tremendous benefit from LOI and CDR experience

Production now simply scales with additional sites

Several limitations of the ILC DIRAC CDR setup have been lifted

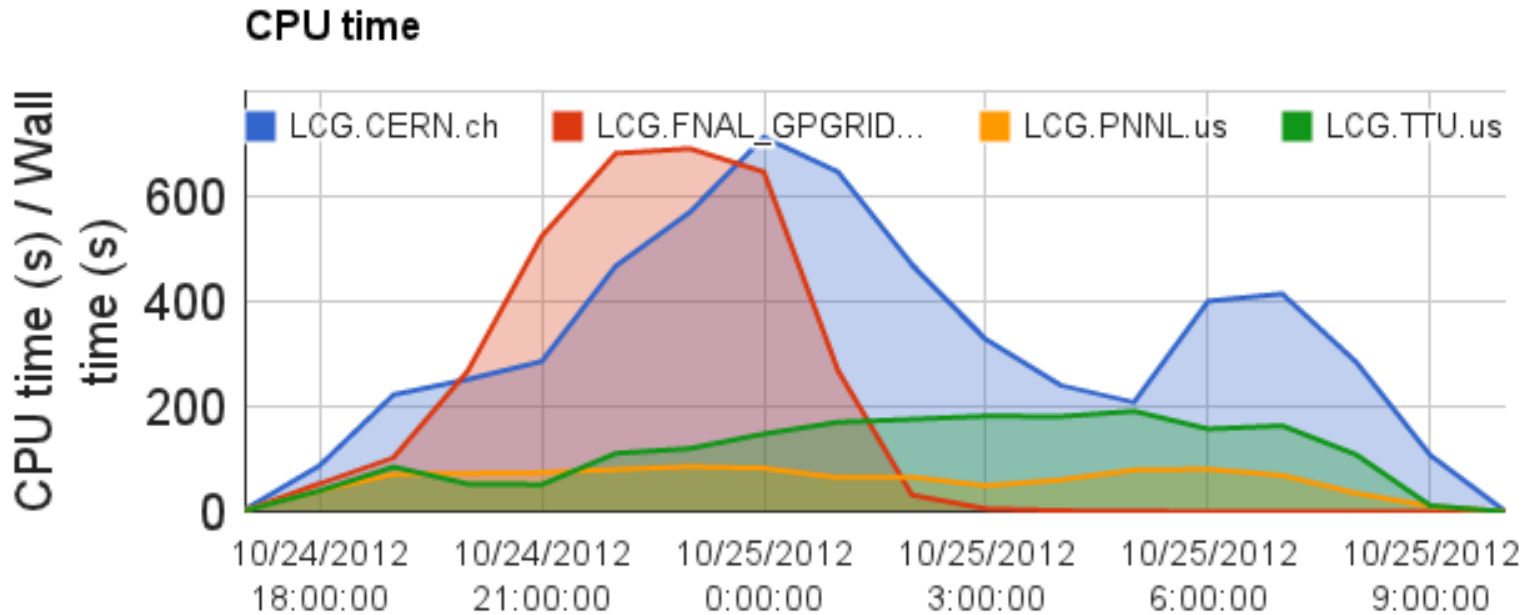


CPU Time by Country



Country	CPU Time (years)
UK	85.8
CH	56.9
FR	13.9
US	4.2

Production update



US sites are now configured to take jobs

Increasing efficiency very time-intensive, PNNL only US site with manpower to liaise

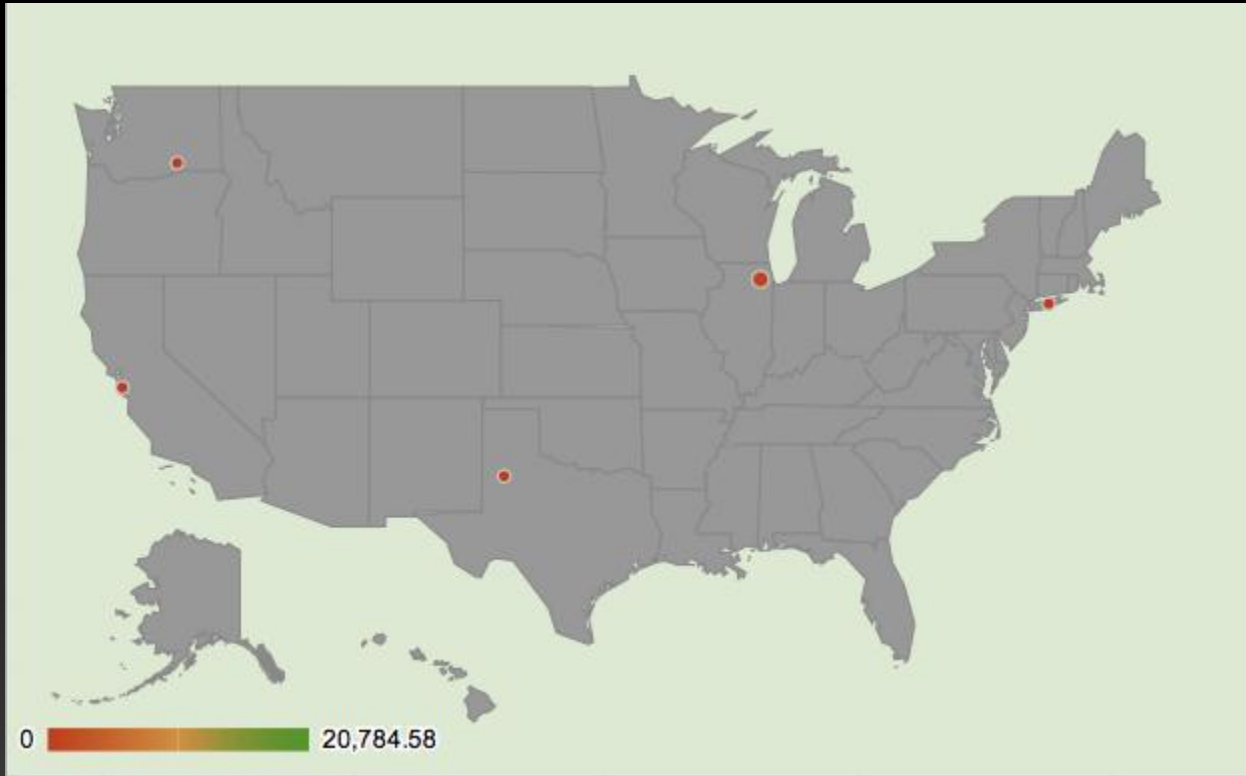
Conclusions / Lessons Learned

- ILCDIRAC production system was further improved in terms of efficiency and error handling
- Bookkeeping is most time consuming part of production
 - Automate as much as possible
- Site configuration problems (wrong defaults) can lead to deterioration of the overall performance
 - Optimal performance during production requires continuous monitoring
 - Every site is different
- Even with the best tools, running a production is a manpower-intensive task

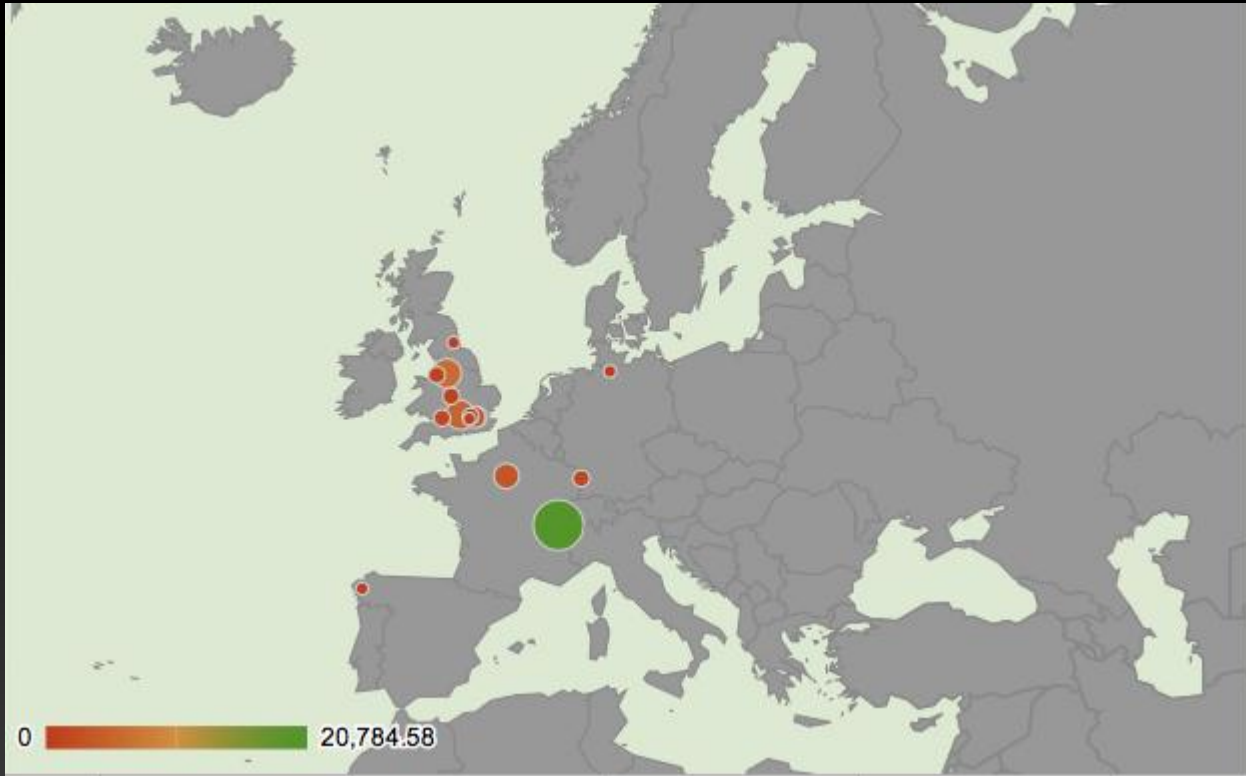
Summary

- More than 66 million events have been processed in about 6 weeks
 - Re-processing stage currently on-going, estimated 3 weeks to finish
- The Fermilab and DESY ILC VOs have been merged
 - ILCDIRAC supports sites in OSG and in WLCG
 - **Talk to us** if you have idle resources, default setup is likely to cause problems
- **All data** produced in the context of the SiD DBD and **all tools are freely available** to anybody
 - Theorists, or Students can perform analysis on **realistic** simulation of the detector concept at a 1 TeV ILC
 - **Talk to us** if you would like to participate

BACKUP

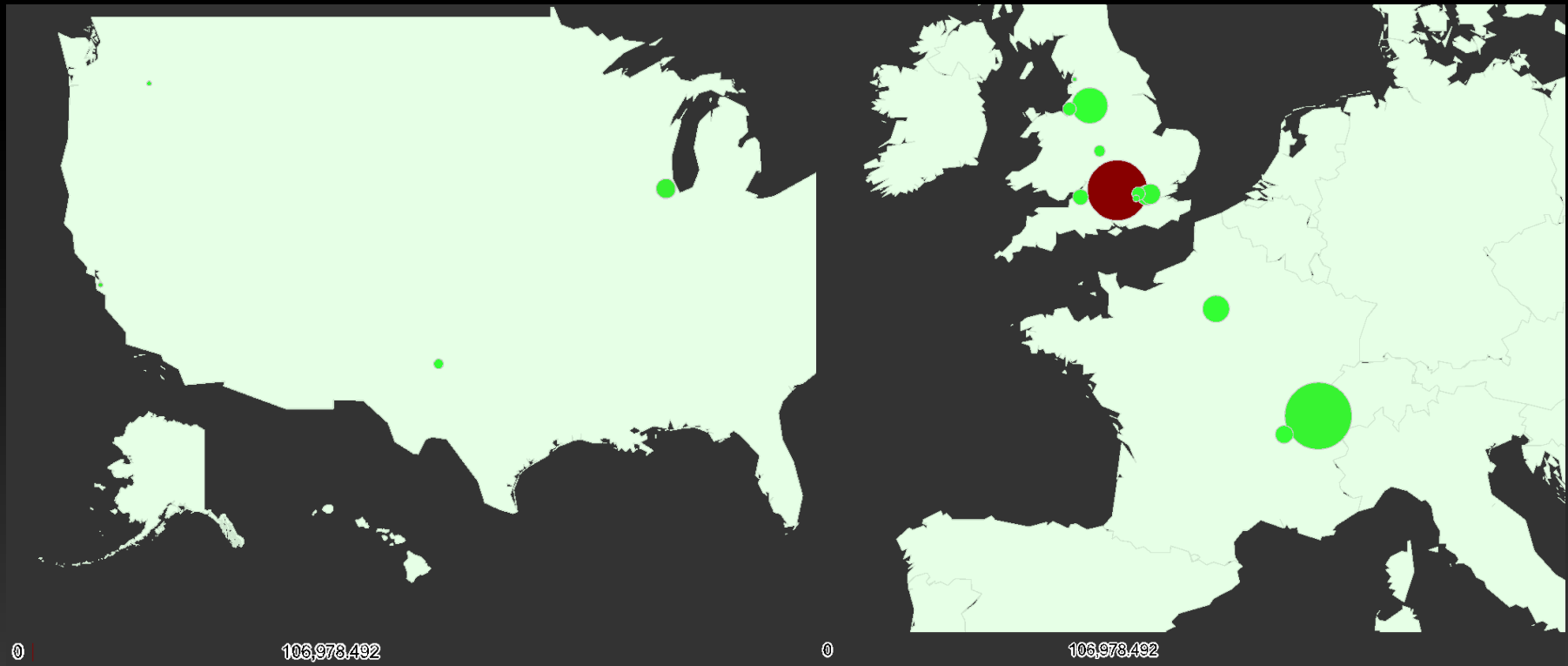


SiD Production Status LCWS 2012



SiD Production Status LCWS 2012

Global Distribution of Resources



Size of circles proportional to CPU time

Color: Amount of storage used during production