

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

# ILC DIRAC, a grid solution for the LC community

S. Poss<sup>1</sup> and P. Majewski<sup>1,2</sup>

<sup>1</sup>CERN, Switzerland

<sup>2</sup>Gdansk University of Technology, Poland

March 2010 / LCWS10, Beijing

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances
  - Achievements
  - Number of sites
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
    - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances
  - Achievements
  - Number of sites
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions

# Aim of this work

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

In 2011, CLIC community releases the Conceptual Design Report. Volume 3 describes the physics and detector studies.

This needs:

- Generation of MC events for the benchmark channels and background events
- Simulation of detector
- Reconstruction and analysis
- For both ILD and SiD geometries

Need to heavily use the GRID, not much time to start from scratch

⇒ DIRAC

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

## 1 Introduction

- Aim

- **DIRAC**

## 2 ILCDIRAC

- Our developments

## 3 Performances

- Achievements

- Number of sites

- Storage of output data

## 4 Issues

- MySQL

- File catalog

- Others

## 5 Outlook and conclusions

# DIRAC:

## Distributed Infrastructure with Remote Agent Control

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

Developed as a full GRID solution for the LHCb experiment.

- System designed to manage large amount of data
- Comply with VO specific problems: heterogeneous resources, applications, etc.
- Overcome deficiencies of standard GRID middleware
- Alleviate the excessive burden from sites in supporting multiple VO

Not the only solution to solve those problems, other LHC experiments also developed their tools

# More on DIRAC

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

- Pilot jobs: higher job efficiency than with standard jobs
  - Jobs are pulled from the central task queue
  - Multiple jobs can run in the same CPU slot (Filling mode)
- Apply VO policy directly in DIRAC, not by the site
- Security follows GRID standards

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

## 1 Introduction

- Aim
- DIRAC

## 2 ILCDIRAC

- Our developments

## 3 Performances

- Achievements
- Number of sites
- Storage of output data

## 4 Issues

- MySQL
- File catalog
- Others

## 5 Outlook and conclusions



# Interfacing with DIRAC core

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

DIRAC project already made such that adding a new “client” is easy

- LHCbDIRAC is an extension

Our work up to now:

- Use the principles of LHCbDIRAC to build ILCDIRAC
- Wrap around Mokka and Marlin applications to run them safely on the GRID sites

## Installation procedure:

- Download tar ball from server and untar it
- Set up MySQL server

## User input needed:

- Steering file
- Generator file
- Number of events to run
- Optional: events number to start from, DB slice, output file name (when running Marlin after)

One does not need to make changes to the steering file previously used for interactive tests ⇒ **fewer errors introduced**

## Installation procedure:

- As Mokka: download and untar tar ball
- Setup the environment variables

## Input needed:

- XML file for steering
- Optional if Mokka was ran before: Gear file, slcio file list, number of events to process

Again, no modification needed in input XML, all are done automatically, **more user friendly**

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

**Achievements**

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

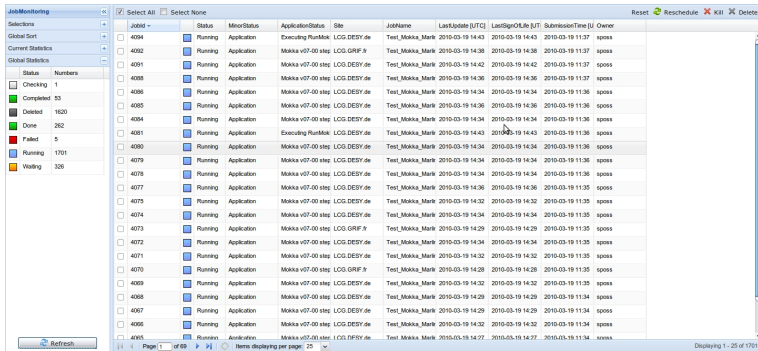
Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances**
  - Achievements**
  - Number of sites
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions

# Number of simultaneously running jobs

2000 jobs each using 1000 3TeV  $e^+e^- \rightarrow \mu\mu\nu\nu$  events



The screenshot displays the JobMonitoring interface with a table of job details. The table includes columns for JobID, Status, MinorStatus, Application, ApplicationStatus, Site, JobName, LastUpdate [UTC], LastSignOff [UTC], SubmissionTime [UTC], and Owner. The 'Status' column shows that 1701 jobs are in a 'Running' state, while others are in 'Checking', 'Completed', 'Deleted', 'Done', 'Failed', or 'Waiting' states. The 'Application' column indicates that most jobs are 'Application' type, with some being 'Executing RunMok' or 'Executing RunMok'.

JobID	Status	MinorStatus	Application	ApplicationStatus	Site	JobName	LastUpdate [UTC]	LastSignOff [UTC]	SubmissionTime [UTC]	Owner
4094	Running	Application	Executing RunMok	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:43	2010-03-19 14:43	2010-03-19 11:37	spos
4092	Running	Application	Mokka v07-00 step	LOG.GRF.fr		Test_Mikka_Marke	2010-03-19 14:38	2010-03-19 14:38	2010-03-19 11:37	spos
4091	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:42	2010-03-19 14:42	2010-03-19 11:37	spos
4098	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:36	2010-03-19 14:36	2010-03-19 11:37	spos
4096	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:36	spos
4095	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:36	2010-03-19 14:36	2010-03-19 11:36	spos
4094	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:36	spos
4081	Running	Application	Executing RunMok	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:43	2010-03-19 14:43	2010-03-19 11:36	spos
4090	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:36	spos
4079	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:36	spos
4078	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:36	spos
4077	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:36	2010-03-19 14:36	2010-03-19 11:35	spos
4075	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:32	2010-03-19 14:32	2010-03-19 11:35	spos
4074	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:35	spos
4073	Running	Application	Mokka v07-00 step	LOG.GRF.fr		Test_Mikka_Marke	2010-03-19 14:29	2010-03-19 14:29	2010-03-19 11:35	spos
4072	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:34	2010-03-19 14:34	2010-03-19 11:35	spos
4071	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:32	2010-03-19 14:32	2010-03-19 11:35	spos
4070	Running	Application	Mokka v07-00 step	LOG.GRF.fr		Test_Mikka_Marke	2010-03-19 14:28	2010-03-19 14:28	2010-03-19 11:35	spos
4069	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:32	2010-03-19 14:32	2010-03-19 11:35	spos
4068	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:29	2010-03-19 14:29	2010-03-19 11:34	spos
4067	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:29	2010-03-19 14:29	2010-03-19 11:34	spos
4066	Running	Application	Mokka v07-00 step	LOG.DESY.de		Test_Mikka_Marke	2010-03-19 14:32	2010-03-19 14:32	2010-03-19 11:34	spos
4065	Running	Application	Mokka v07-00 step	LOG.GRF.fr		Test_Mikka_Marke	2010-03-19 14:27	2010-03-19 14:27	2010-03-19 11:34	spos

1700 running in parallel

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

**Number of sites**

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances**
  - Achievements
  - Number of sites**
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions

# Sites used

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
**Number of sites**  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

Using only SL5, 64 bit software, only 6 sites usable:

- DESY (DE)
- KEK (JP)
- GRIF (FR)
- CEA (FR)
- POLGRID (FR)
- IRES (FR)

25 are available for the VO

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances**
  - Achievements
  - Number of sites
  - Storage of output data**
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions



Currently using only IN2P3 storage, but CERN resources should be available soon

- Output files automatically copied to Storage Element (SE) if larger than 10Mb
- Output data is systematically copied to SE
- Files are added to the catalog

**Users don't need to worry about data management**, they only need to specify their output

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances
  - Achievements
  - Number of sites
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - Others
- 5 Outlook and conclusions

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

## Problems encountered:

- Threading: MySQLd server must run in background
- Have to wait for socket creation, depends on the site
- Socket path length: limited to 108 characters, while standard GRID directories have  $> 200$
- Using /tmp/ on worker node: sites don't really like that

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances
  - Achievements
  - Number of sites
  - Storage of output data
- 4 Issues**
  - MySQL
  - File catalog**
  - Others
- 5 Outlook and conclusions

# FileCatalog and Bookkeeping

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
**File catalog**  
Others

Outlook and  
conclusions

Backup

Need to:

- Determine common data path structure
- Populate the catalog(s)

In DIRAC, using multiple catalogs in parallel is possible

- DIRAC provides one, in development
- Possibility to implement an interface to the DESY catalog (Jan Engels)

# Outline

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

- 1 Introduction
  - Aim
  - DIRAC
- 2 ILCDIRAC
  - Our developments
- 3 Performances
  - Achievements
  - Number of sites
  - Storage of output data
- 4 Issues
  - MySQL
  - File catalog
  - **Others**
- 5 Outlook and conclusions

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

## Software specific:

- Software compatibility with different OS and architecture to run on more sites
- Some channels ( $t\bar{t}$ ) take a long time to process in Mokka, need to split input generator file to avoid hitting the CPU time limit.

## VO specific:

- Another VO with same name exists! The resource discovery agent got confused.

# Outlook and conclusions

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

## Plans for future:

- File catalog & bookkeeping
- File splitter
- Production system
- Add SiD software (SLIC and LCSIM)

## Conclusions:

- Simple to use for any user
- Reliable
- Efficient



# Acknowledgments

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

Many thanks to the DIRAC developers: A. Smith, S. Paterson,  
A. Tsaregorodstev, J. Closier, V. Roma

Thanks to A. Gellrich for sorting out the VO problems

Thanks to the CERN IT for the support

Thanks to André Sailer for all the help understanding Mokka/Marlin

# Backup slides

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim

DIRAC

ILCDIRAC

Our developments

Performances

Achievements

Number of sites

Storage of output data

Issues

MySQL

File catalog

Others

Outlook and  
conclusions

Backup

# VO Box requirements

ILC DIRAC, a  
grid solution for  
the LC  
community

S. Poss and  
P. Majewski

Introduction

Aim  
DIRAC

ILCDIRAC

Our developments

Performances

Achievements  
Number of sites  
Storage of output data

Issues

MySQL  
File catalog  
Others

Outlook and  
conclusions

Backup

- Local account dirac
- /opt/dirac partition owned by dirac user. Must be large as contains DB and sandbox
- Ports 80, 443 opened for web server
- Ports range 9130:9200 opened for DIRAC services
- Host certificate

## Installation procedure:

- MySQL inside Mokka tar ball
- Setup the environment variables

## Input needed:

- Paths to Local socket, data directory and binary directory:  
determined at run time
- Database dump file: either specified by user or default from  
tar ball used

## After Mokka run:

- Shutdown MySQLd server through socket
- Remove MySQL from disk

⇒ **No user intervention needed.**