ILCDIRAC
A GRID Solution for the LC Community
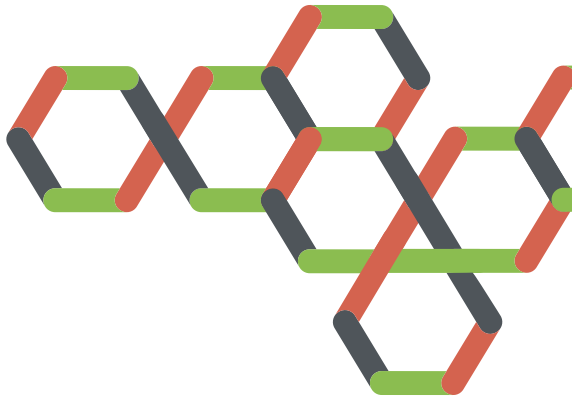
Christian Grefe*, Stéphane Poss*, André Sailer*

*CERN PH-LCD
on behalf of the CALICE collaboration and the CLIC physics and detector study

12. November 2013
Linear Collider Workshop, Tokyo

Outline

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

Overview
Application Framework
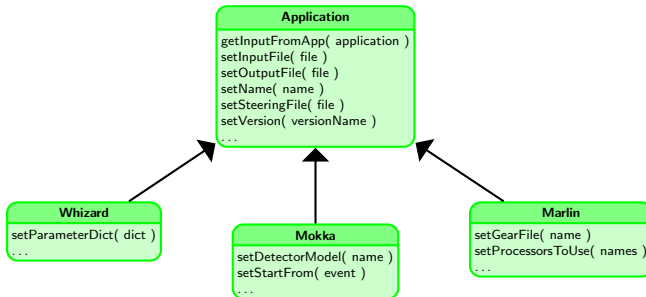Software Management

What is ILCDIRAC?

- Based on DIRAC (LHCb):
  Full Grid solution: Workload management, File catalog with meta data, Production System
- High level interfaces for all linear collider applications to allow easy job definition
- Software management system to ensure availability of all application on all sites
- Overlay system to automatically retrieve files including pile-up
- Multiple VOs supported:
  CALICE and ILC VO share many application and profit from ILCDIRAC
- More than 100 registered users, with $\sim$ 10 very active users
- ILCDIRAC fulfills all requirements $\rightarrow$ currently only occasional bug fixes

ILCDIRAC is stable!

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

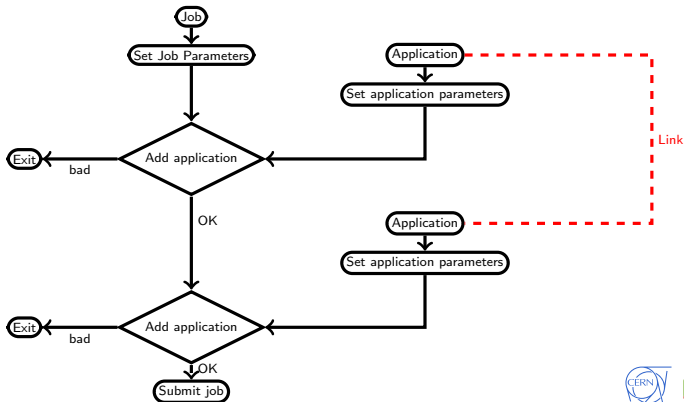Overview
Application Framework
Software Management

The Application Framework

- High level interface for all linear collider applications that streamlines job definition (14 applications)
- Linking of applications: output of one application can be used as input to another application

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

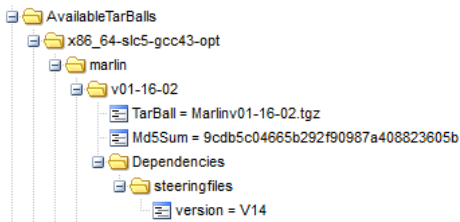Overview
Application Framework
Software Management

The Application Framework

- High level interface for all linear collider applications that streamlines job definition (14 applications)
- Linking of applications: output of one application can be used as input to another application

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

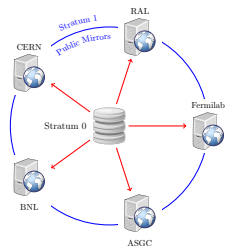Overview
Application Framework
Software Management

Software Management

- Applications and versions defined in configuration system
- Reference tar balls stored on dedicated storage element
- Job takes care of installation if software not available on node
- Install software in shared area if possible
- Locking of installation directory to avoid conflicting installations
- Dependencies are supported ⇒ automatic installation of related packages
- Default steering files (CDR/DBD) are automatically deployed as dependency and don't have to be provided in the input sandbox

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

Overview
Application Framework
Software Management

Moving towards CVMFS

- Distributed read-only file system
- Files are downloaded and cached on demand
- Applications are installed on reference machine at CERN (stratum 0)
- Files replicated to tier 1 sites (stratum 1)
- Worker nodes connect to stratum 1 servers with transparent fallback to other stratum 1

- Modified workflow modules implemented and under test
- Absolutely transparent for the users
- Switch to SLC6 as default installation
- Revisiting steering file distribution $\Rightarrow$ explicit application in job submission instead of "hidden" dependency

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

Overview
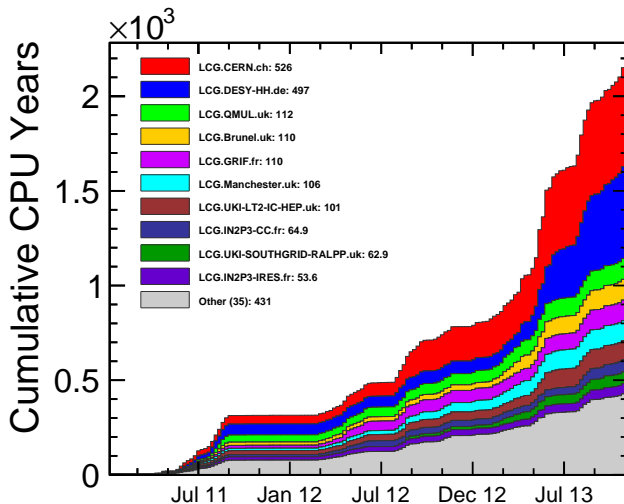Production Jobs
User Jobs
Data Management

- Successfully used in several mass production campaigns:
  CLIC CDR, SiD DBD, ILD DBD (user jobs only), CLIC Higgs Paper
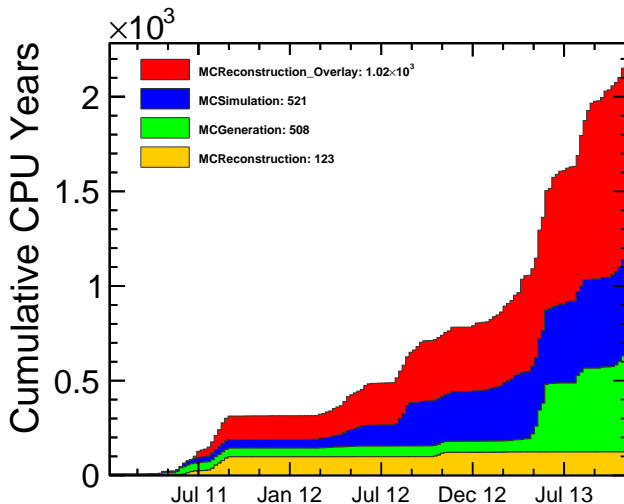- On average several thousand jobs running in parallel (maximum > 13k jobs)
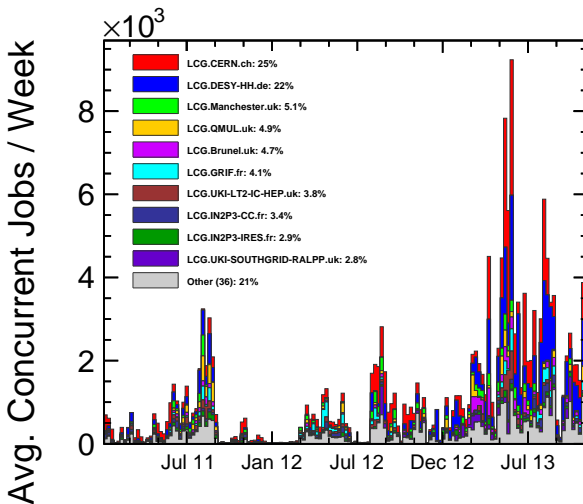
Production Jobs

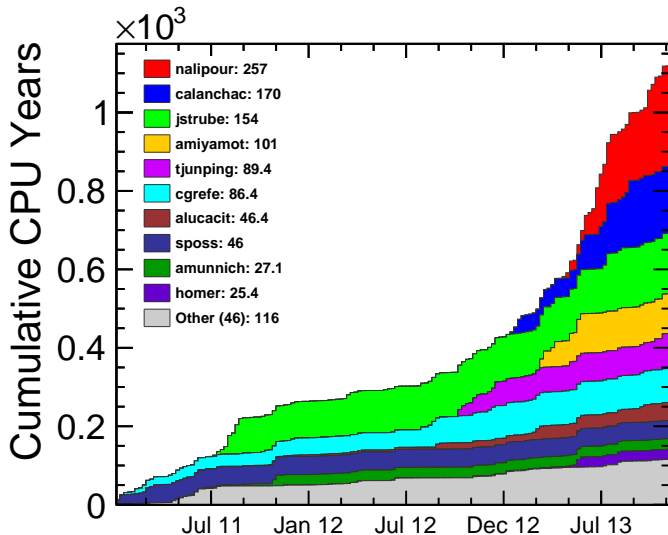- More than 2k CPU years over past 3 years used in mass production

Production Jobs

- Most time used for reconstruction (including beam-beam interactions)

User Jobs

- DIRAC file catalog (DFC) knows about all files produced using ILCDIRAC
- Most ILD DBD files migrated into DFC
- Production files have meta data and ancestry set:
  $\Rightarrow$ searchable from the web interface and CLI
- Failover mechanism prevents data loss
- Automatic replication between sites with at least two replicas stored by default

Storage	Total [TB]	CLIC [TB]	ILC [TB]	User [TB]	# Files
CERN	1196	962	7	227	5.7×10^6
DESY	196	-	170	16	8.8×10^5
KEK	155	-	100	56	7.0×10^5
RAL	168	4	89	76	1.5×10^6
IN2P3	57	-	33	24	4.2×10^5
PNNL	26	-	25	-	7.4×10^5
Total	1792	973	424	401	1.0×10^7

What is ILCDIRAC?
Status of the System
Contact and Support
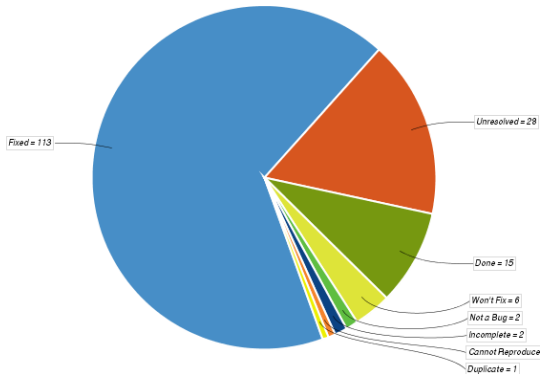Summary and Outlook

Getting Started
Bug Tracking

Getting Started

- (For now) ILCDIRAC requires dedicated registration in addition to registration with the ILC VO
- Mailing lists:
    - Registration: ilcdirac-register@cern.ch
    - Questions: ilcdirac-support@cern.ch
- Forum:
    http://forum.linearcollider.org/index.php?t=index&cat=22
- Documentation:
    http://twiki.cern.ch/twiki/bin/view/CLIC/DiracUsage

Any new user is welcome!

What is ILCDIRAC?
Status of the System
Contact and Support
Summary and Outlook

Getting Started
Bug Tracking

Bug Tracking

- JIRA is used for bug tracking:

  `http://its.cern.ch/jira/browse/ILCDIRAC`

- JIRA tickets can be created directly from the web interface
  ⇒no account required

Summary and Outlook

- ILCDIRAC is a complete Grid solution providing dedicated interfaces for all linear collider applications
- It fulfills all requirements and is stable
- Software distribution will be done using CVMFS soon
- Interface for Whizard2 will be added
- ILCDIRAC was extremely useful in the mass productions for CLIC and SiD
- ILD is now adopting it as their production system
- Active support from the CERN LCD group

Thanks

Our thanks go to

- the GRID site administrators for fixing issues quickly
- the DIRAC developers for fast replies and discussions
- the users for useful and positive feedback