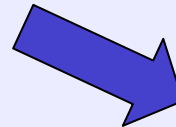
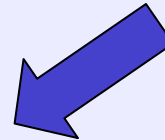


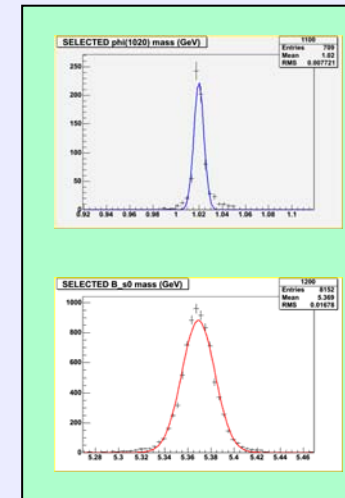
# DIRAC Infrastructure For Distributed Analysis



```
from DIRAC.Client.Dirac import *
dirac = Dirac()
job = Job()
job.setApplication('DaVinci', 'v12r11')
job.setInputSandbox(['Application_DaVinci_v12r11/DV_Pi0Calibr.opts', 'Application_DaVinci_v12r11/lib'])
job.setInputData(['/lhcb/production/DC04/v2/DST/00000746_00002032_9.dst', '/lhcb/production/DC04/v2/DST/00000746_00002038_9.dst'])
job.setOutputSandbox(['pi0calibr.hbook', 'pi0histos.hbook'])
jobid = dirac.submit(job, verbose=1)
print "Job ID = " jobid
```

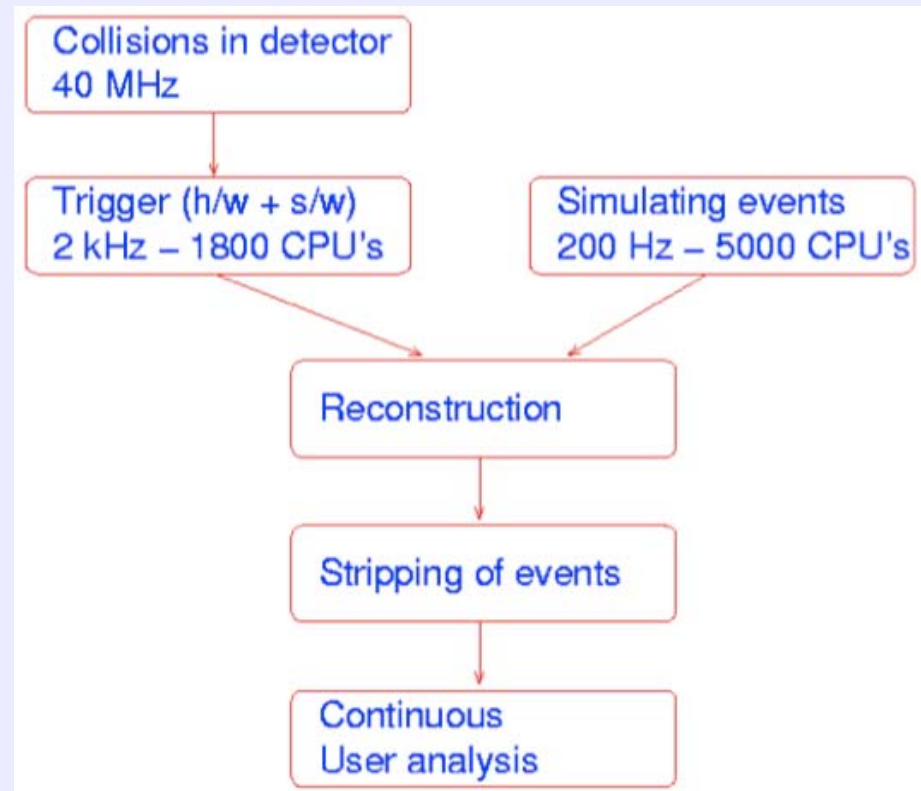


IoP Particle Physics 2006  
Warwick

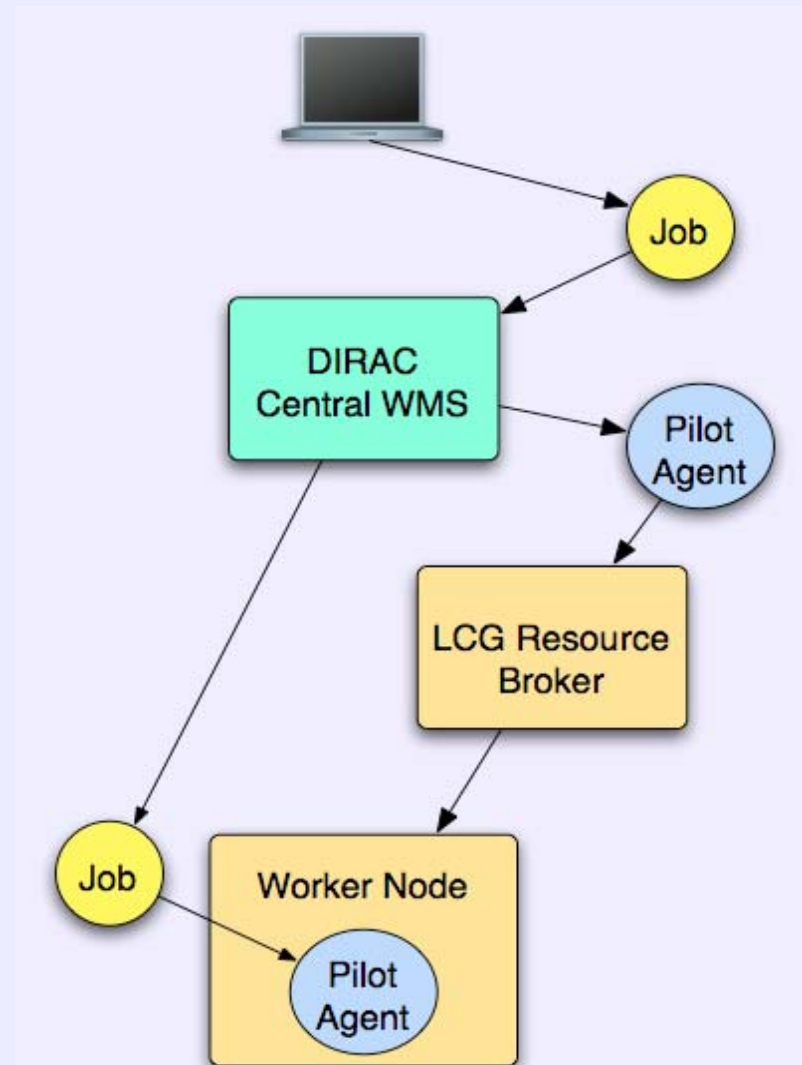


- Distributed Analysis in LHCb
- Introduction to DIRAC
  - DIRAC API
  - Workload Management System
- WMS Strategies & Performance
- Outlook

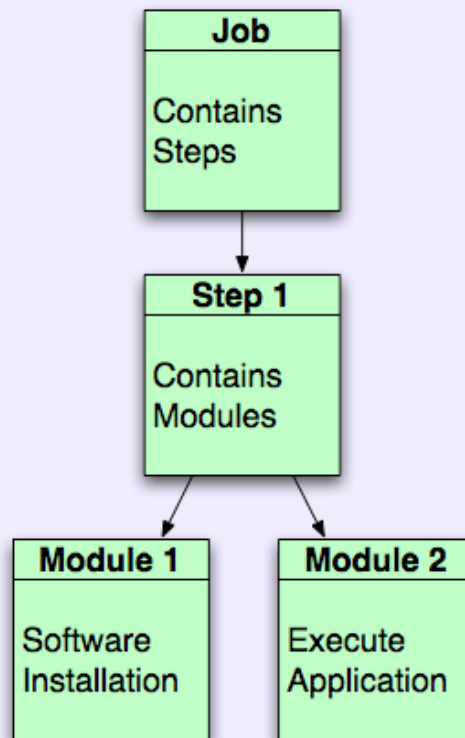
- Stripping is centrally managed analysis run as a production
  - Reduces analysis datasets to a size of  $10^6$  to  $10^7$  events per year
- The user analysis is performed on stripped data throughout the year
  - User analysis will take place at all LHCb Tier 1 centres
  - All stripped data will be disk resident and replicated to all sites
- No direct submission to LCG for LHCb Analysis jobs
  - Instead submit via DIRAC



- The DIRAC Workload & Data Management System (WMS) is made up of **Central Services** and **Distributed Agents**
  - Realizes PULL scheduling paradigm
    - Agents are requesting jobs whenever the corresponding resource is available
    - Execution environment is checked before job is delivered to WN
  - Service Oriented Architecture masks underlying complexity



- The DIRAC API provides a transparent way for users to submit production or analysis jobs to LCG
  - Can be single application or complicated DAGs



```
from DIRAC.Client.Dirac import *
dirac = Dirac()
job = Job()
job.setApplication('DaVinci', 'v12r15')
job.setInputSandbox(['DaVinci.opts', 'lib'])
job.setInputData(['/lhcb/production/DC04/v2/DST/
00000742_00003493_10.dst'])
job.setOutputSandbox(['DVNTuples.root', 'DaVinci_v12r15.log'])
jobid = dirac.submit(job,verbose=1)
print "Job ID = ",jobid
```

- While it may be exploited directly, the DIRAC API also serves as the interface for the GANGA Grid front-end to perform distributed user analysis for LHCb

# LHCb DIRAC Monitoring



DIRAC Monitoring (Analysis) - Opera 8.51

http://lhcb01.pic.es/DIRAC/Monitoring/Analysis/

Accounting Overview

Transfer Accounting

Details

Production ID: All

Site: All  
ANY  
DIRAC.Cambridge.uk  
DIRAC.CERN.ch  
DIRAC.LXGATE34.ch

Job Status: All

App Status: All  
Brunel execution, step 1  
DIRAC job initialization  
DaVinci execution, step 1

Owner: All

Max results: 500

Job Ids:

Jobs updated after: 10/01/2006

Submit Reset

<a href="#">4062</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:00:20	krinnert
<a href="#">4063</a>	failed	Job finished successfully	LCG.CERN.ch	Gauss_1	2006-01-11 14:00:22	jdickens
<a href="#">4064</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:00:22	pkoppenb
<a href="#">4065</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:02:34	krinnert
<a href="#">4066</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 14:45:03	lessnoff
<a href="#">4067</a>	failed	Job finished successfully	LCG.Lyon.fr	Gauss_1	2006-01-11 14:57:13	scan
<a href="#">4068</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:54:12	powell
<a href="#">4069</a>	failed	Job finished successfully	LCG.CERN.ch	Gauss_1	2006-01-11 15:02:43	scan
<a href="#">4071</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:14:30	paterson
<a href="#">4072</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:08:37	paterson
<a href="#">4073</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:07:35	paterson
<a href="#">4074</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:07:58	paterson
<a href="#">4075</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:08:42	paterson
<a href="#">4076</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:02:19	pkoppenb
<a href="#">4077</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:07:30	nstyles
<a href="#">4078</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:18:38	soler
<a href="#">4079</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:14:59	soler
<a href="#">4080</a>	outputready	Job finished successfully	LCG.PIC.es	DaVinci_1	2006-01-11 15:33:38	nraja
<a href="#">4081</a>	outputready	Job finished successfully	LCG.PIC.es	DaVinci_1	2006-01-11 15:31:43	nraja
<a href="#">4082</a>	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 15:12:28	nraja
<a href="#">4083</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:07:54	pkoppenb
<a href="#">4085</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:22:02	pkoppenb
<a href="#">4086</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:41:45	uegede
<a href="#">4087</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:53:16	powell
<a href="#">4088</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 15:55:49	uegede
<a href="#">4089</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 18:08:33	soler
<a href="#">4090</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 18:03:43	soler
<a href="#">4091</a>	outputready	Job finished successfully	LCG.RAL.uk	DaVinci_1	2006-01-11 17:48:03	soler
<a href="#">4092</a>	outputready	Job finished successfully	LCG.RAL.uk	DaVinci_1	2006-01-11 17:46:25	soler
<a href="#">4093</a>	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 17:43:18	uegede
<a href="#">4094</a>	outputready	Job finished successfully	LCG.RAL.uk	DaVinci_1	2006-01-11 17:31:36	paterson

# LHCb DIRAC Monitoring



DIRAC Monitoring (Analysis) - Opera 8.51

http://lhcb01.pic.es/DIRAC/Monitoring/Analysis/

Accounting Overview

Transfer Accounting

Details

Production ID: All

Site: All, ANY, DIRAC.Cambridge.uk, DIRAC.CERN.ch, DIRAC.LXGATE34.ch

Job Status: All

App Status: All, Brunel execution, step 1, DIRAC job initialization, DaVinci execution, step 1

Owner: All

Max results: 500

Job Ids: [input field]

Jobs updated after: 10/01/2006

Submit Reset

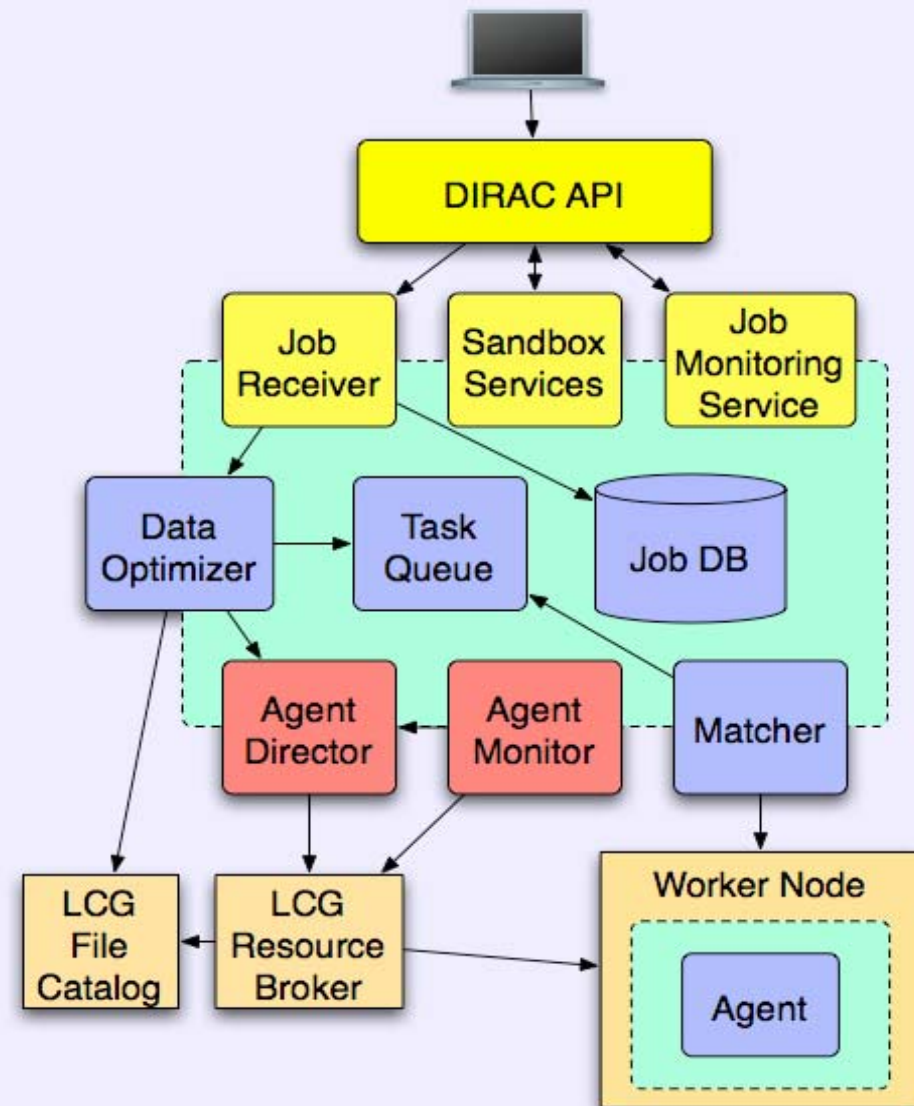
Job ID	Status	Message	Site	Job Name	Start Time	Owner
4062	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:00:20	krinnert
4063	failed	Job finished successfully	LCG.CERN.ch	Gauss_1	2006-01-11 14:00:22	jdickens
4064	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:00:22	pkoppenb
4065	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 14:02:34	krinnert
4066	outputready	Job finished successfully	LCG.Lyon.fr	DaVinci_1	2006-01-11 14:45:03	lessnoff
4067	failed				2006-01-11 14:57:13	scan
4068	outputready				2006-01-11 14:54:12	powell
4069	failed				2006-01-11 15:02:43	scan
4071	outputready				2006-01-11 15:14:30	paterson
4072	outputready				2006-01-11 15:08:37	paterson
4073	outputready				2006-01-11 15:07:35	paterson
4074	outputready				2006-01-11 15:07:58	paterson
4075	outputready				2006-01-11 15:08:42	paterson
4076	outputready				2006-01-11 15:02:19	pkoppenb
4077	outputready				2006-01-11 15:07:30	nstyles
4078	outputready				2006-01-11 15:18:38	soler
4079	outputready				2006-01-11 15:14:59	soler
4080	outputready				2006-01-11 15:33:38	nraja
4081	outputready				2006-01-11 15:31:43	nraja
4082	outputready				2006-01-11 15:12:28	nraja
4083	outputready				2006-01-11 15:07:54	pkoppenb
4085	outputready				2006-01-11 15:22:02	pkoppenb
4086	outputready				2006-01-11 15:41:45	uegede
4087	outputready				2006-01-11 15:53:16	powell
4088	outputready				2006-01-11 15:55:49	uegede
4089	outputready				2006-01-11 18:08:33	soler
4090	outputready				2006-01-11 18:03:43	soler
4091	outputready				2006-01-11 17:48:03	soler
4092	outputready				2006-01-11 17:46:25	soler
4093	outputready	Job finished successfully	LCG.CERN.ch	DaVinci_1	2006-01-11 17:43:18	uegede
4094	outputready	Job finished successfully	LCG.RAL.uk	DaVinci_1	2006-01-11 17:31:36	paterson

Details of job  
lhcb01.pic.es

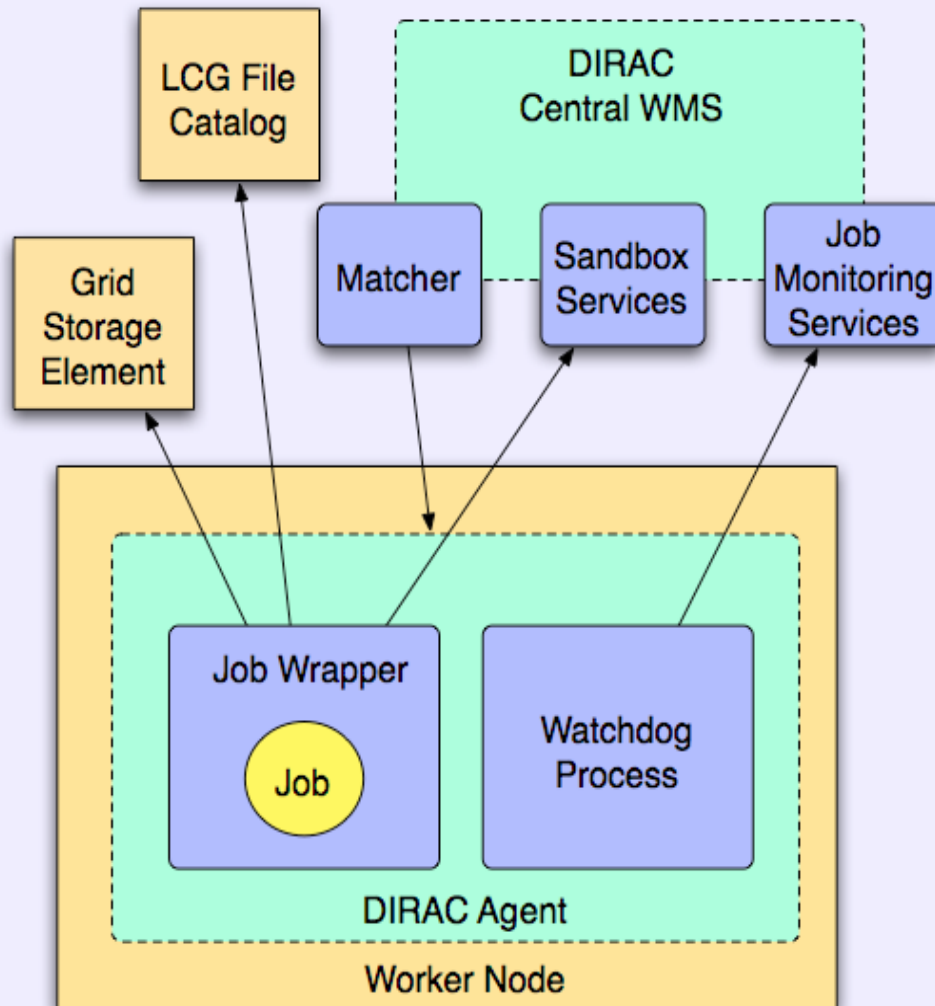
Parameter Name	Value
ApplicationStatus	Job finished successfully
JDL	<pre>[   Requirements = { member("CNAF_Castor", other.Location);   Arguments = "jobDescription.xml";   JobName = "DaVinci_1";   parameters =   [     STEPS = "1";     STEP_1_NAME = "0_0_1"   ];   SoftwarePackages =   {     "DaVinci.v12r14"   };   JobType = "user";   Executable = "\$LHCBPRODROOT/DIRAC/scripts/jobexe";   CPUtime = "500";   StdOutput = "std.out";   Owner = "paterson";   OutputSandbox =   {     "pool_xml_catalog.xml",     "std.out",     "std.err",     "DVNTuples.hbook",     "DaVinciRun.csh",   } ];</pre>

# DIRAC Workload Management System

- DIRAC API interfaces with the client securely
  - Uses DISET Framework
- Data Optimizer queries the LCG File Catalogue (LFC) to determine suitable sites for job execution
- Agent Director and Agent Monitor services handle submission to LCG
  - Pilot Agents go through the LCG Resource Broker as normal jobs







- Dynamically creates Job Wrapper
  - Handles Input/Output Sandboxes
  - Transfers and registers Output Data as required, according to LHCb conventions
- Sends accounting and site specific information back to DIRAC WMS
- Runs a Watchdog process in parallel to the application
  - Provides 'heartbeat' for DIRAC Monitoring
- Provides access to any requested Input Data

- A study was performed to compare the possible WMS strategies for analysis jobs
  - We know DIRAC can deal with long production jobs, what about the other extreme?
- Consider shortest possible 'useful' job
  - Analysis of 500 Events
- These short jobs are of high priority since results must be back as soon as possible

- There are several ways to use the DIRAC infrastructure but the end goal is to **minimise the start time** of user analysis jobs
  - The following explores some of the possibilities
- DIRAC Modes of submission
  - 'Resubmission'
    - Pilot Agent submission to LCG with monitoring
    - Multiple Pilot Agents may be sent in case of LCG failures
  - 'Filling Mode'
    - Pilot Agents may request several jobs from the same user, one after the other
  - 'Multi-Threaded'
    - Same as 'Filling' Mode above except two jobs can be run in parallel on the Worker Node

# Experiments Performed

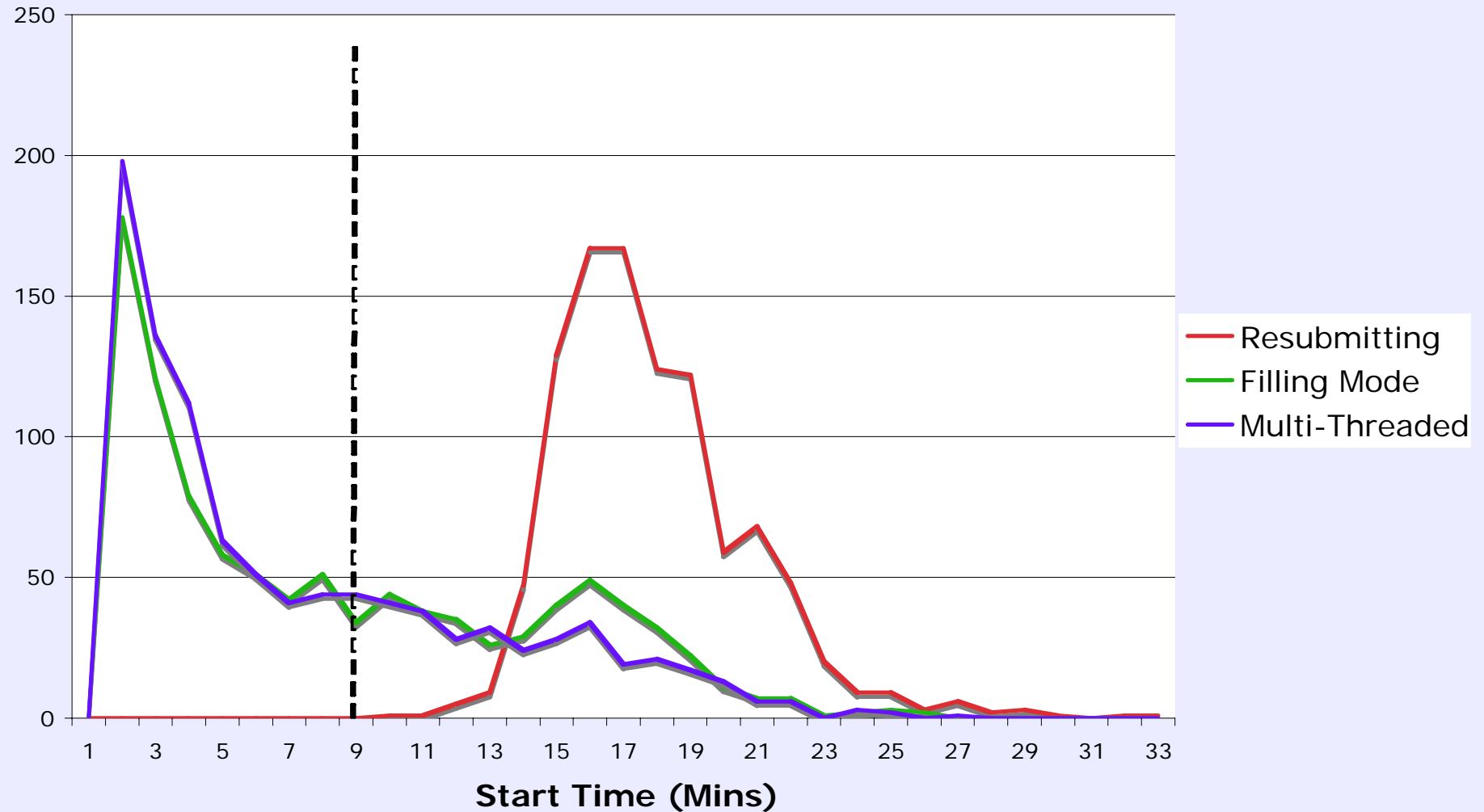


- Short (5 Mins) analysis jobs running at Tier 1 Centres
- Jobs for different modes were intermixed in order to ensure the same Grid 'weather'
- Submission of jobs was at the pace of the Resource Broker
- Job start time is defined as the time between submission and the application starting on the WN
- For each experiment 3 Users submitted 100 jobs
  - Each user submitted jobs for one of the three different modes described above
- These experiments were repeated several times to see if results are reproducible

# Start Times for 10 Experiments, 30 Users



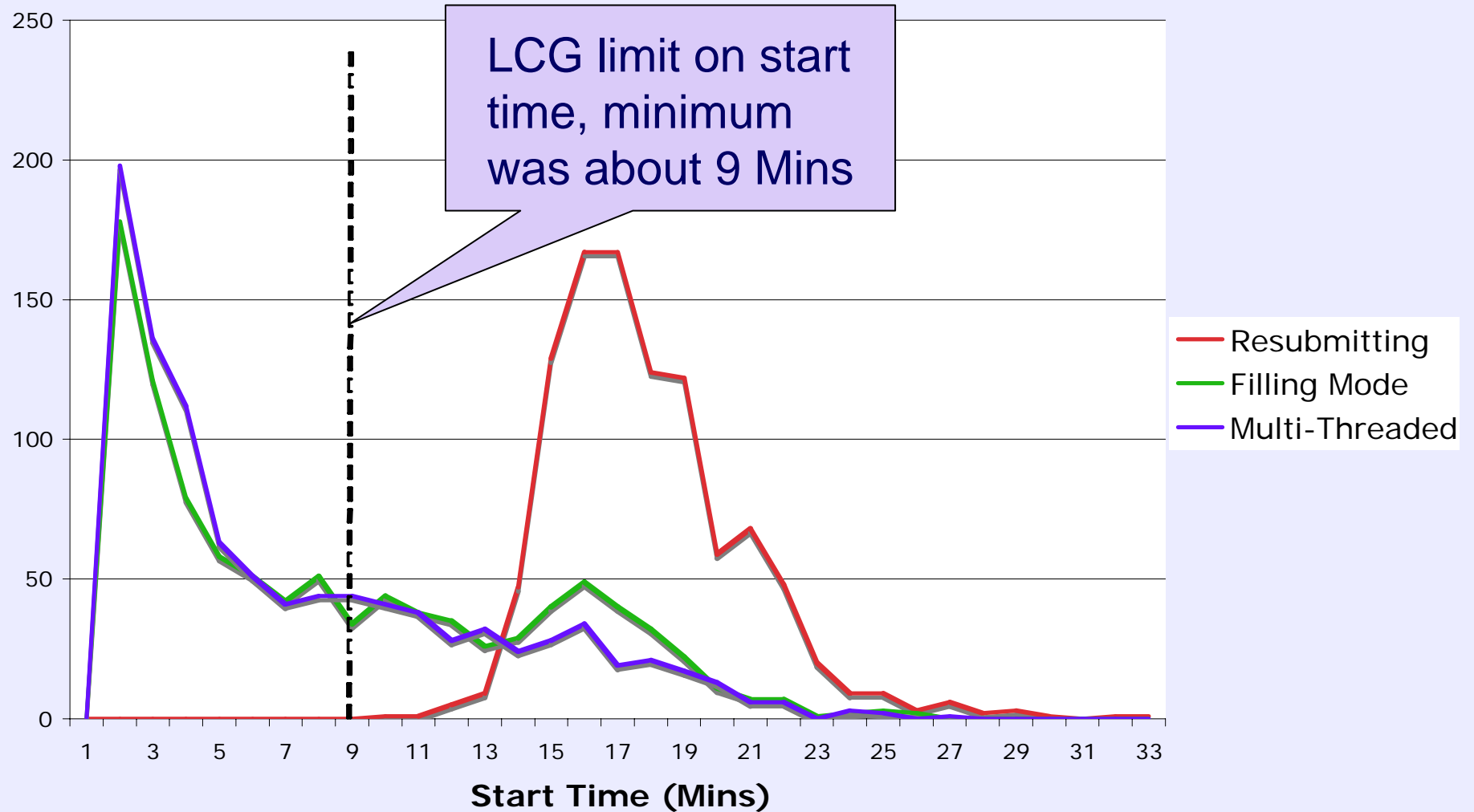
Start Times for 30 Users  
3000 Jobs, 1.5 Million Events



# Start Times for 10 Experiments, 30 Users



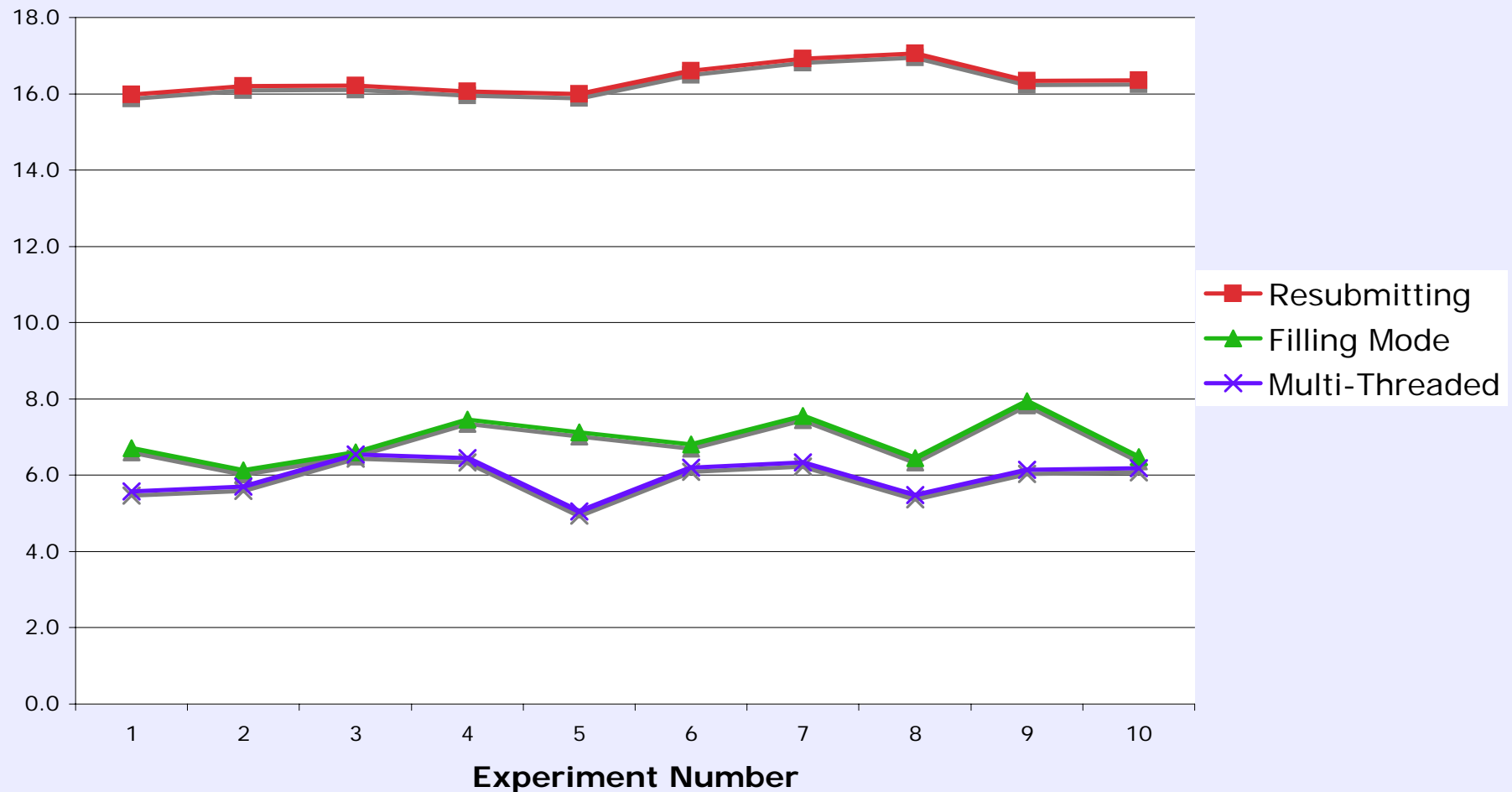
Start Times for 30 Users  
3000 Jobs, 1.5 Million Events



# Result of 10 Experiments from 30 Users



Mean Start Times for 10 Experiments with 30 Users  
3000 Jobs, 1.5 Million Events



- All 3000 jobs successfully completed within ~40 Mins of starting
- Can see clear improvement on start times for Multi-Threaded and Filling modes
  - Even when LCG performance is optimal, see significant reduction
  - A factor of 3 to 4 fewer LCG jobs need to be sent
    - These modes therefore reduce the load on LCG
- Optimization of workload is performed for each user independently
  - Here each user submitted 100 jobs
    - This can be further improved by optimizing the workload on the level of the Virtual Organisation (VO)



- DIRAC masks the inefficiencies of LCG
  - In these tests LCG performed well, this is not always the case...
  - Can execute LHCb VO policy in one central place for the benefit of all users
    - e.g. troublesome LCG sites can be banned for all users at once
- DIRAC WMS currently demonstrated ~2Hz throughput for analysis jobs
  - Actual throughput limited by available LCG resources and capacity of Resource Broker

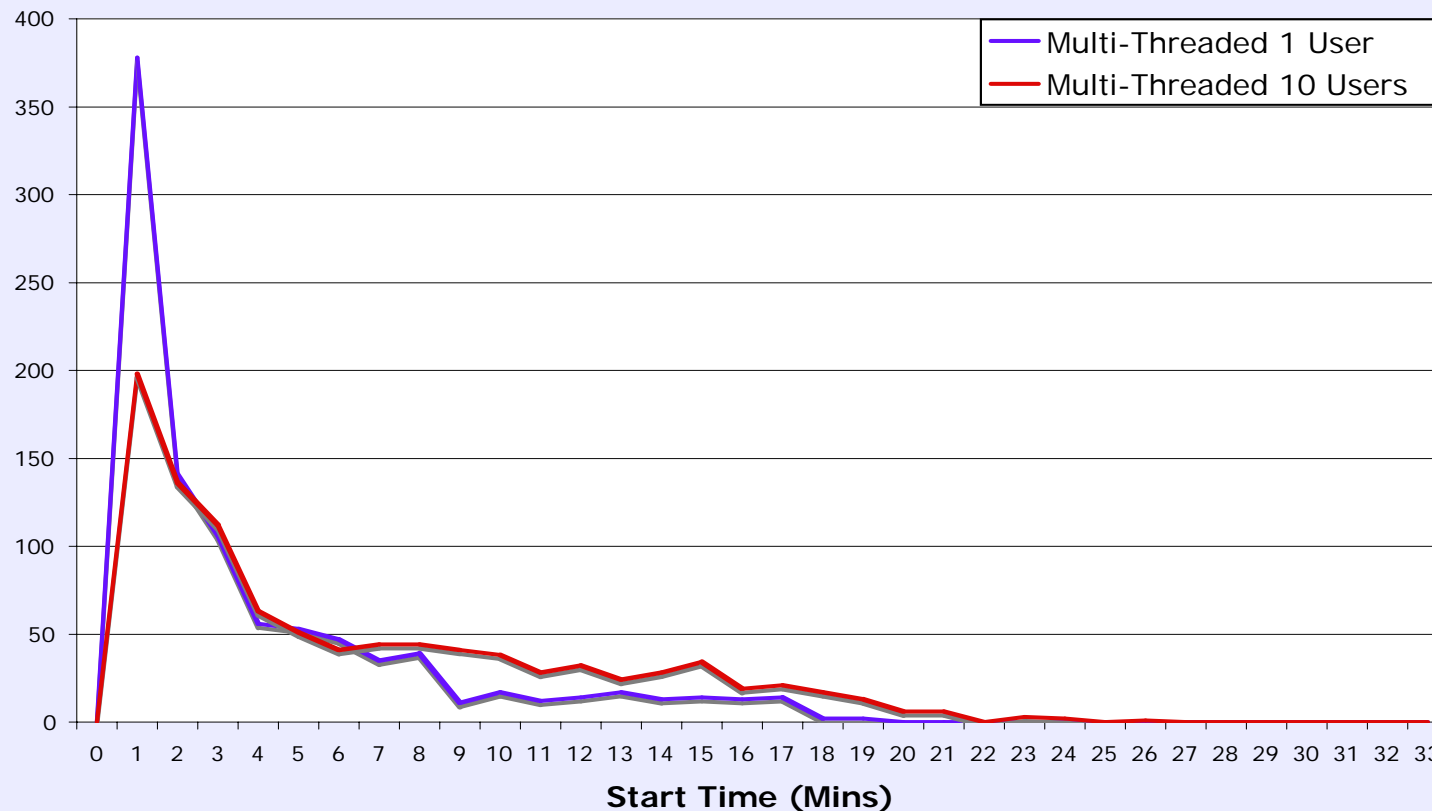
- The DIRAC API provides a simple yet powerful tool for users
  - Access to LCG resources is provided in a simple and transparent way
- DIRAC Multi-Threaded and Filling modes show significant reductions on the job start times
  - Also reduce the load on LCG
- Workload management on the level of the user is effective
  - Can be more powerful on the level of the VO
- DIRAC infrastructure for distributed analysis is in place
  - Now have real users

- DIRAC Infrastructure for Distributed Analysis, Paterson, S., Tsaregorodtsev, A., CHEP06, Mumbai, India, 13-17 February, 2006.
- DIRAC, the LHCb Data Production and Distributed Analysis system, Andrei Tsaregorodtsev et al., CHEP06, Mumbai, India, 13-17 February, 2006.
- Experience with Distributed Analysis in LHCb, Ulrik Egede et al., CHEP06, Mumbai, India, 13-17 February, 2006.
- DIRAC Security Infrastructure, Casajus Ramo, A., Graciani Diaz, R., CHEP06, Mumbai, India, 13-17 February, 2006.
- DIRAC Distributed Analysis on the Computing Grid, Paterson, S., Tsaregorodtsev, A., LHCb Note 2005-072, December, 2005.

# Backup Slide - Workload Optimization



Comparison of 1 and 10 Users for Multi-Threaded Mode



- Two cases,
  - 1 user submitting 1000 jobs
  - 10 users submitting 100 jobs

- There are clear gains in efficiency when the number of users is reduced
- Thus, workload optimization **per user** is effective but less effective than optimization **per Virtual Organisation** could be

# Backup Slide - Total Times For 30 Users



Total times for 30 users  
3000 Jobs, 1.5 Million Events

