



R. Graciani & A. Casajus



Institut de Ciències del Cosmos

# Overview



- **DIRAC**

- what's **DIRAC**?
- DIRAC** in LHCb

- **Recent updates**

- what's new in **DIRAC**?

- **Summary and Outlook**

# What is DIRAC?



**“Distributed Infrastructure**

**with**

**Remote Agent Control”**

A software **Framework**

to **Manage**

**Distributed** computing **Activities**

for a **User Community.**

# DIRAC Design choices



- **VO Centric:**

- Gives to the community, the VO, a central role in the relation of its users with their computing resources.

- **Modularity:**

- To achieve optimal scalability and flexibility, a highly modular design was decided.

- **Pull Scheduling:**

- Implements pull scheduling with late binding of payload to resource to extract optimal performance out of the ever changing underlying resources.

# DIRAC and LHCb



- **DIRAC** is the LHCb solution to manage all its distributed computing activities.
- **DIRAC** started (2002) as tool to manage large scale Monte Carlo simulation with many contributing sites.
- **DIRAC** has evolved into a full, stand-alone grid solution that makes use of different middleware tools and grid services when appropriate.
- **DIRAC** does not define the LHCb computing model, it is a tool to implement it.

# What does DIRAC do for LHCb?



- Detector RAW data upload to CERN and distribution to Tier1's.
- RAW data processing (with data-driven automated job creation and execution):
  - Reconstruction
  - Selection
  - Merging
- Processed DST data replication.
- Monte Carlo simulation and MC data distribution (with parametric automatic job creation and execution).
- Physics analysis of Detector and MC data (with various end-user job definition and submission tool).
- USER data access.
- Supervises, Monitors and Accounts all these activities.

# Recent Updates in DIRAC



- Workload Management
  - Multi-User pilot jobs with gLexec mechanism (beta)
- Data Management
  - New integrated Replica and Metadata File Catalog (beta)
- ☑ Resource Integration
  - New interface to computing clouds (production)

# DIRAC-Cloud Interface

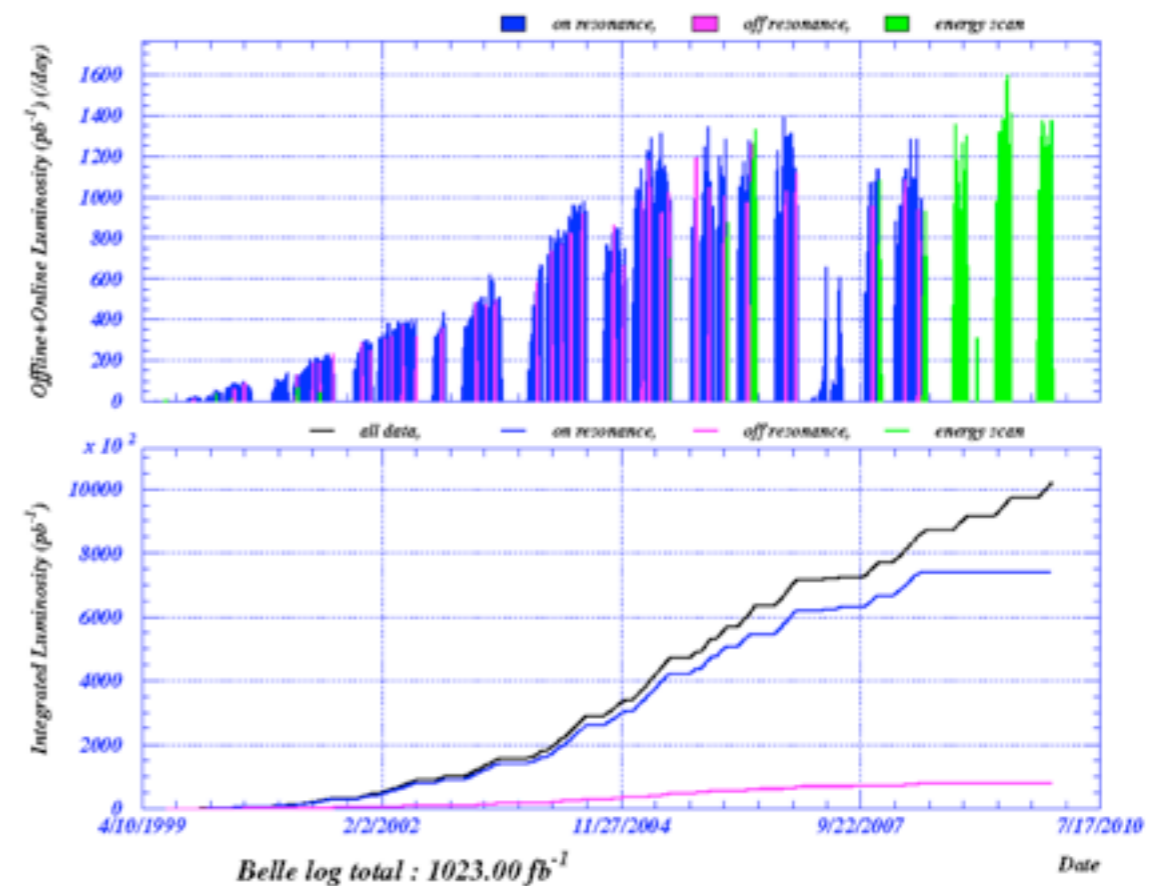


- Use case:
  - Belle experiment (KEK, Japan) has large peaks of CPU needs related to Monte Carlo simulations.
  - Belle II (KEK, Japan) will require 50 times more resources.
  - Explore the cost of using commercial computing resource providers as complement of grid and proprietary resources.



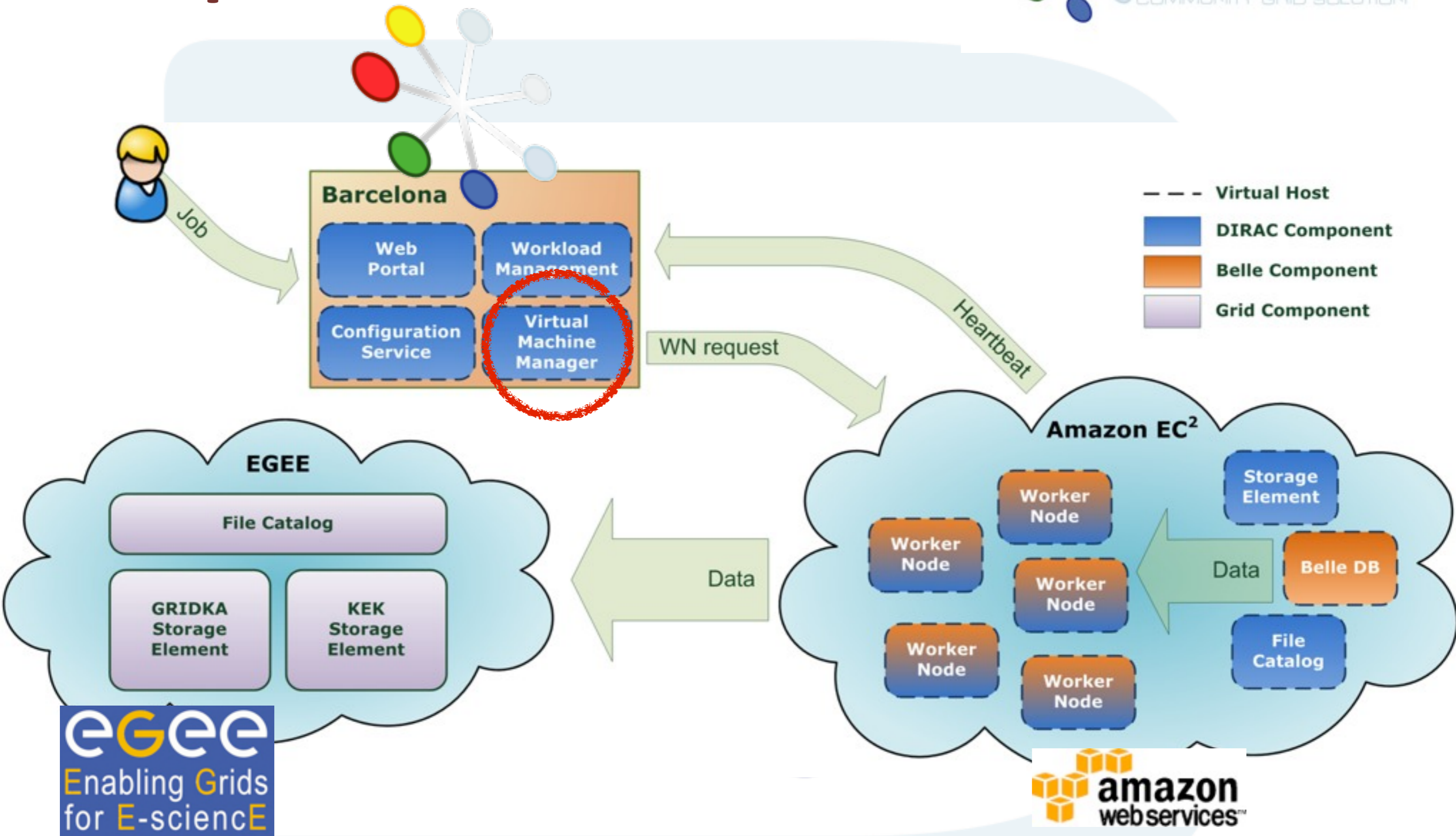
Offline+Online Luminosity ( $\text{pb}^{-1}$ ) (/day)

2009/12/28 07:33





# Proposed DIRAC solution



# Replacing Pilots by Virtual Machines



- **VirtualMachine Scheduler:**
  - Executes in a VM, either local or remote.
  - Monitor DIRAC TaskQueues and request new VM from resource provider if appropriated.
- **VirtualMachine Monitor:**
  - Executes on the VM, reports back the activity and halts the VM if no longer needed.
  - Asynchronous upload of output data.
- **VirtualMachine Manager:**
  - Executes in a VM with other WMS components.
  - Collect information about requested, running and halted VMs. Allows monitoring of their resource usage.

# And there we go !



- Tasks taken from Belle official MC requests.
- Preparations:
  - “Execution VM” image with preinstalled Belle SW and **DIRAC**.
  - “SE VM” image with **DIRAC** FC/SE and postgresql DB.
- Input at KEK:
  - collection of bash scripts
  - event description and backgrounds (1 GB/ 1 M evt)
- Output sent back to grid SE’s (GRIDKA & KEK).
  - 22 GB / 1 M evt

# Starting up (April 14th)



The screenshot shows the DIRAC Job Monitoring interface. The main table displays the following data:

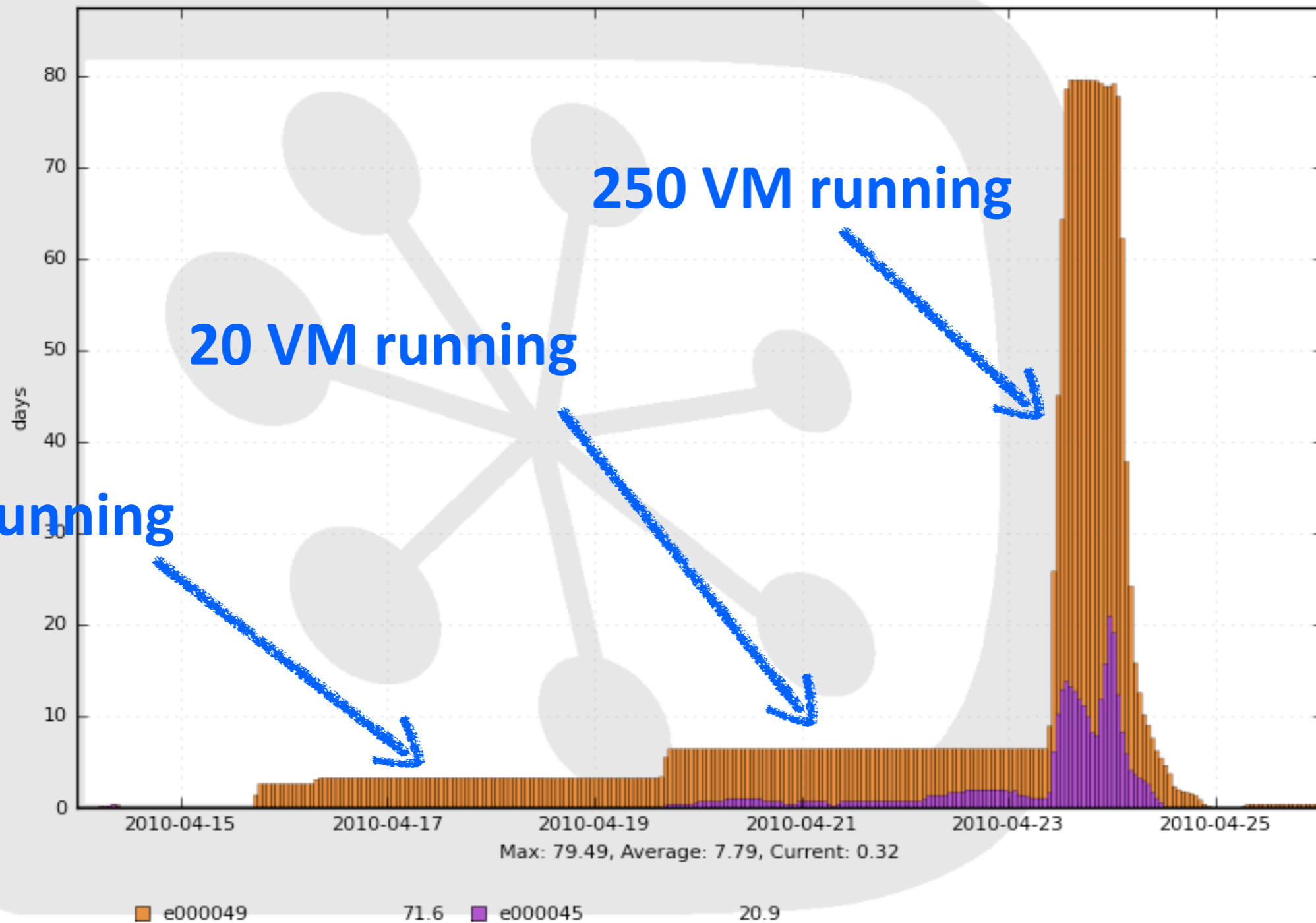
JobId	Status	MinorStatus	ApplicationStatus	Site	JobName	LastUpdate [UTC]	LastSignOfLife [UTC]
670	Running	Job Initialization	Unknown	DIRAC.Amazon.us	e000049r000702	2010-04-14 17:27	2010-04-14 17:27
385	Running	Job Initialization	Unknown	DIRAC.Amazon.us	e000049r000120	2010-04-14 17:23	2010-04-14 17:23
1030	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000448	2010-04-14 14:42	2010-04-14 14:42
1031	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000449	2010-04-14 14:42	2010-04-14 14:42
1032	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000450	2010-04-14 14:42	2010-04-14 14:42
1022	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000435	2010-04-14 14:42	2010-04-14 14:42
1023	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000436	2010-04-14 14:42	2010-04-14 14:42
1021	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000429	2010-04-14 14:42	2010-04-14 14:42
1019	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000372	2010-04-14 14:42	2010-04-14 14:42
1020	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000428	2010-04-14 14:42	2010-04-14 14:42
1017	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000369	2010-04-14 14:42	2010-04-14 14:42
1018	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000371	2010-04-14 14:42	2010-04-14 14:42
1015	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000364	2010-04-14 14:42	2010-04-14 14:42
1016	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000367	2010-04-14 14:42	2010-04-14 14:42
1014	Waiting	Pilot Agent Submis:	Unknown	DIRAC.Amazon.us	e000045r000363	2010-04-14 14:42	2010-04-14 14:42

# The execution: Phase I



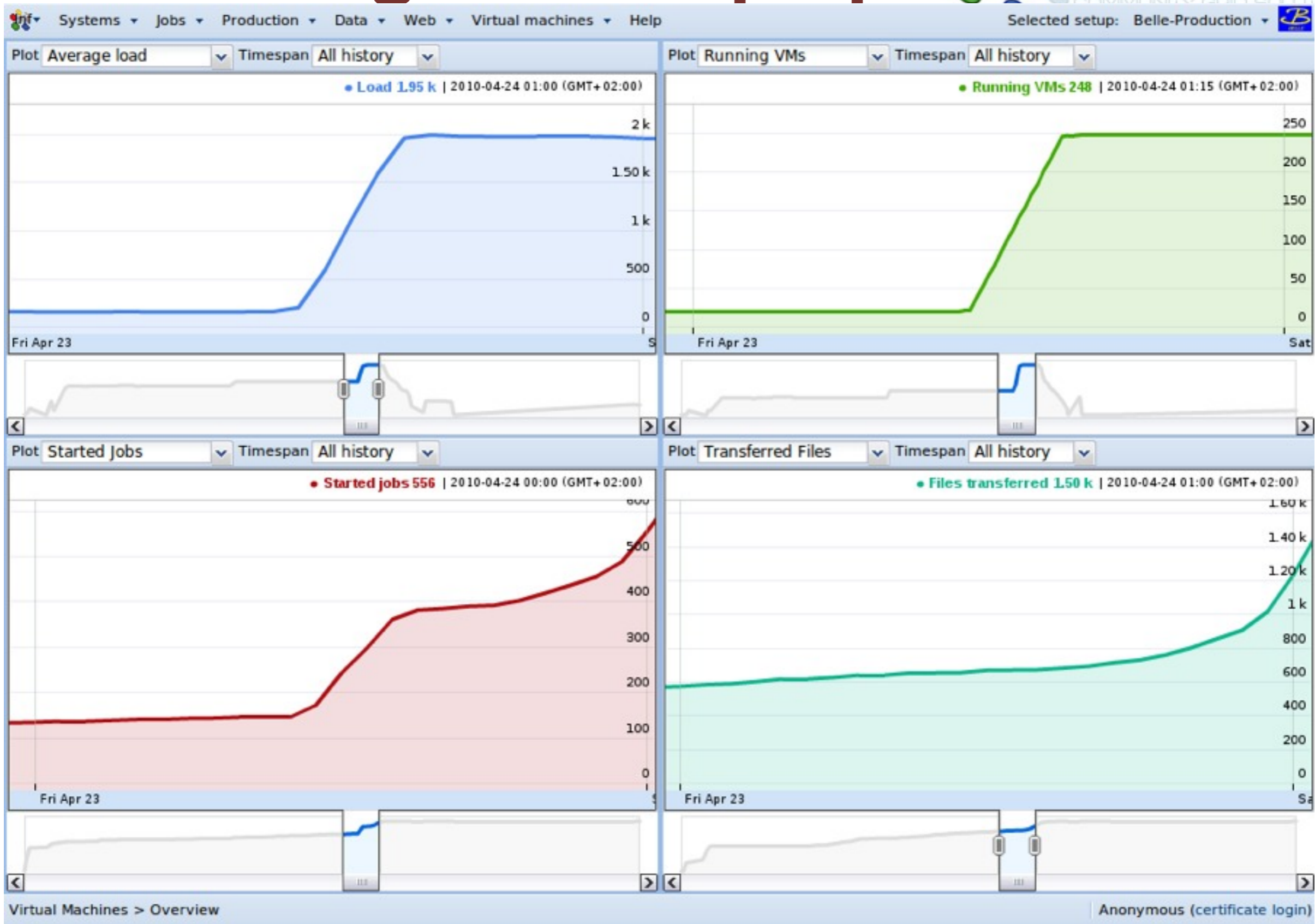
CPU days consumed by simulation Experiment / hour

12 Days from 2010-04-13 to 2010-04-25



Generated on 2010-05-01 15:23:48 UTC

# Monitoring the ramp up

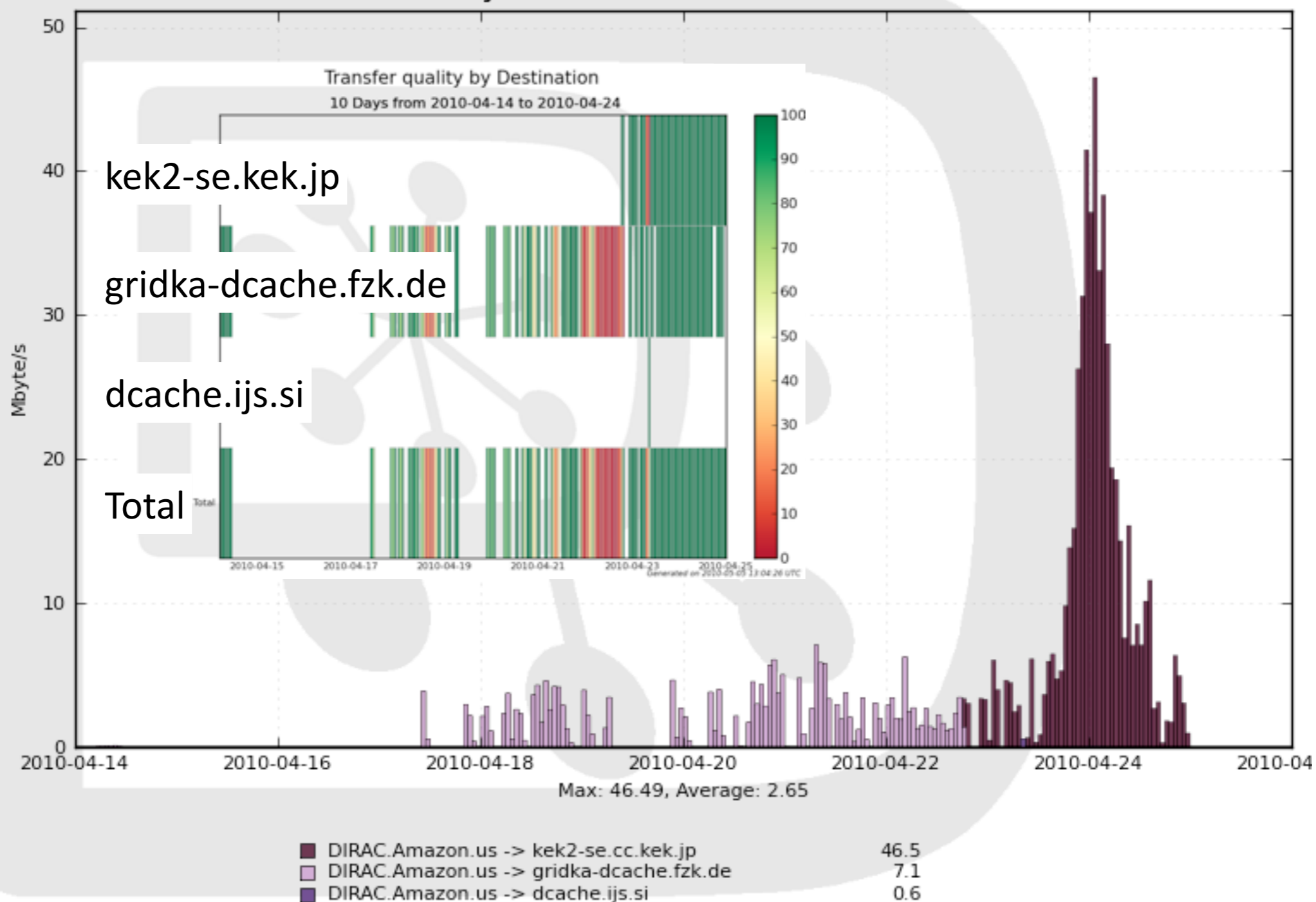


# And the data back to grid



## Transferred data by Channel

11 Days from 2010-04-13 to 2010-04-25



Generated on 2010-05-01 15:05:41 UTC

# Results (I)

- Phase I (cloud test):

- production ready:

- 5% of Belle production in 10 days
- 120 M evt (~2.7 TB)
- 2250 CPU days used

- proven stability and scalability:

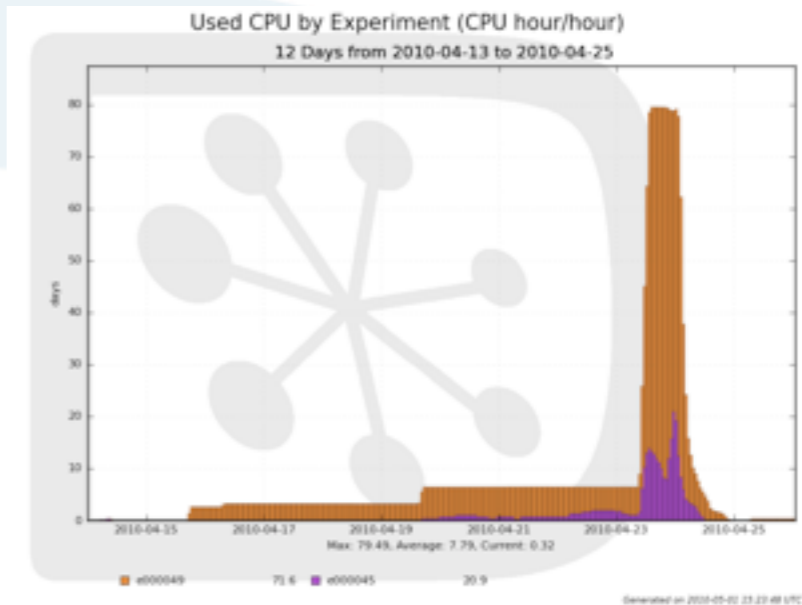
- 2000 CPUs peak achieved in < 4 hours
- > 90 % efficiency in CPU usage

- first cost estimation:

- 0.46 USD/10k evt

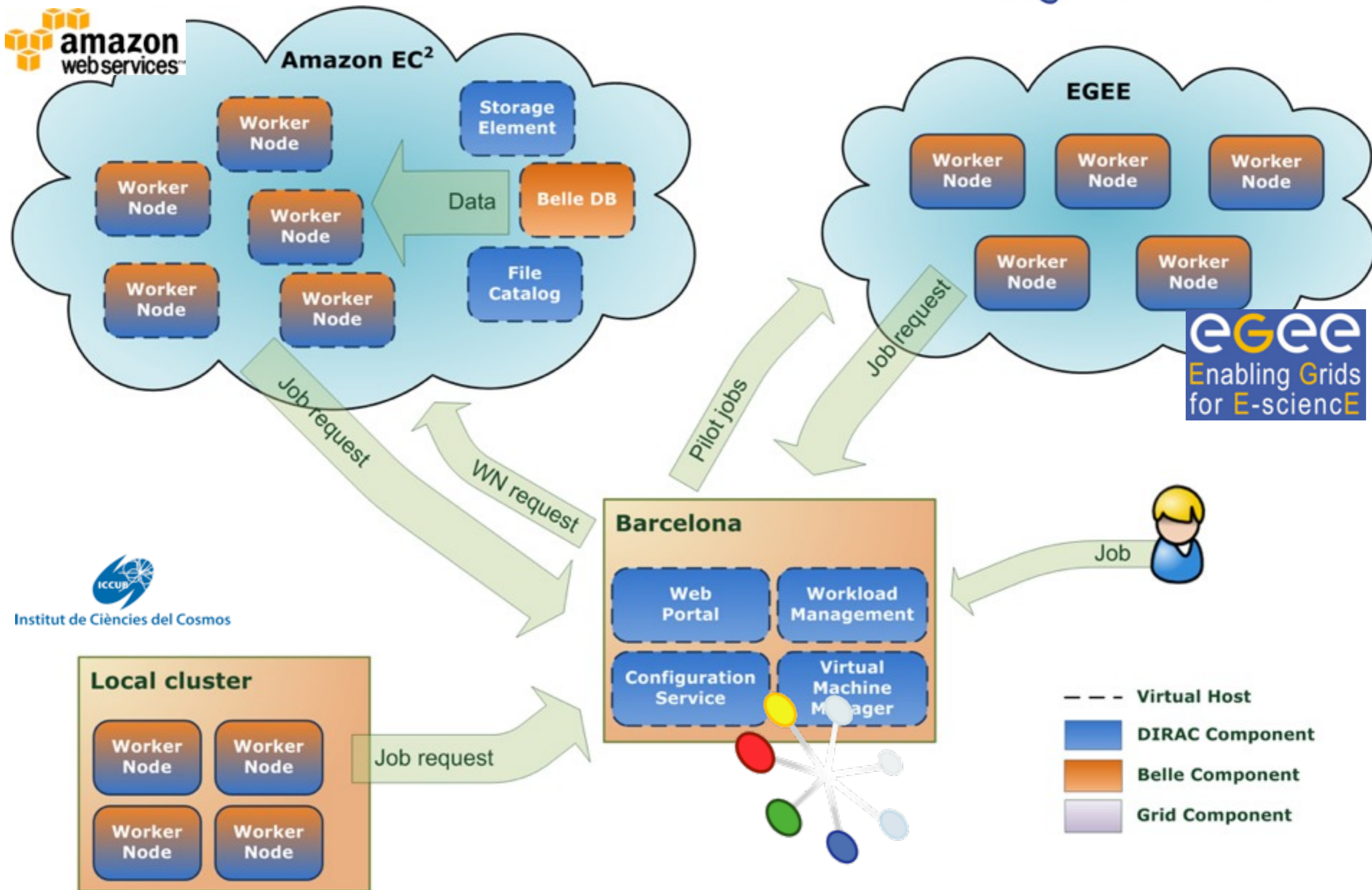
- input data pre-uploaded to Amazon SE VM.

- few bug fixes

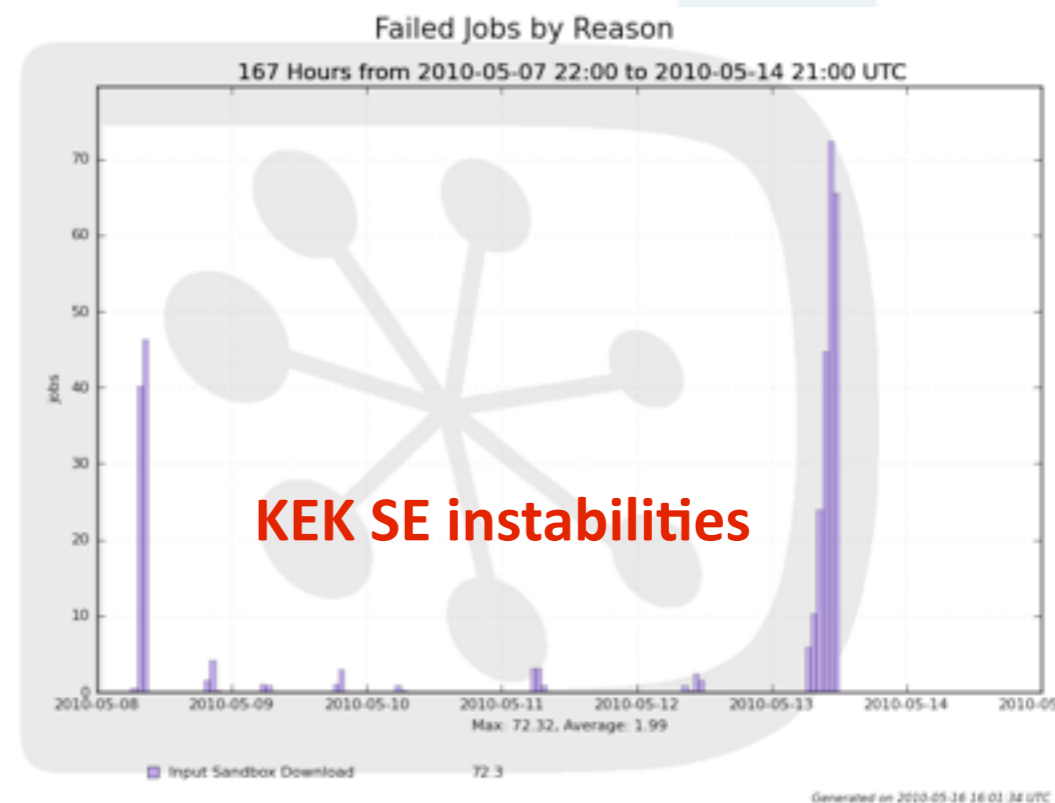
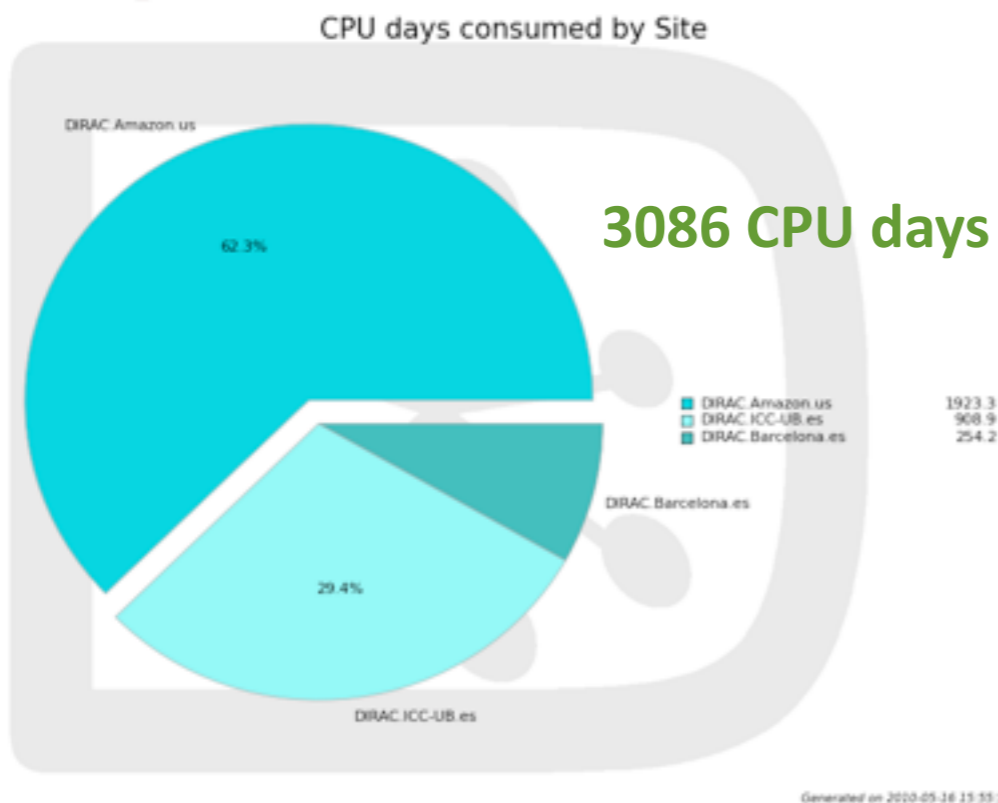
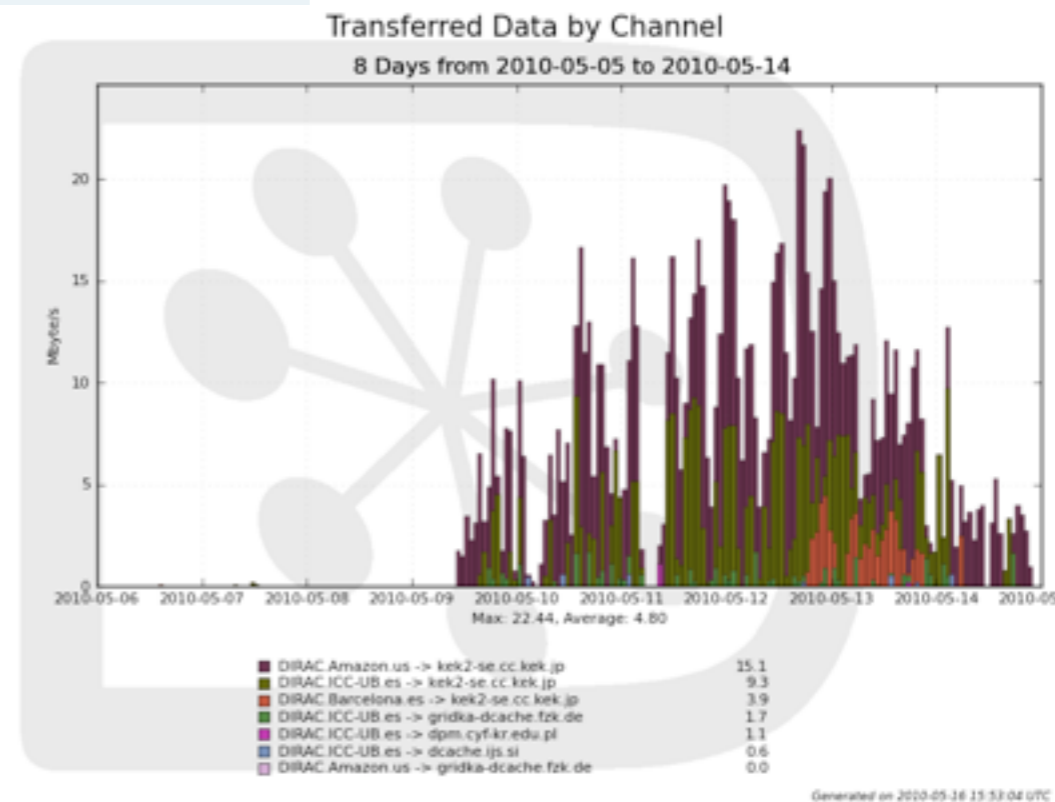
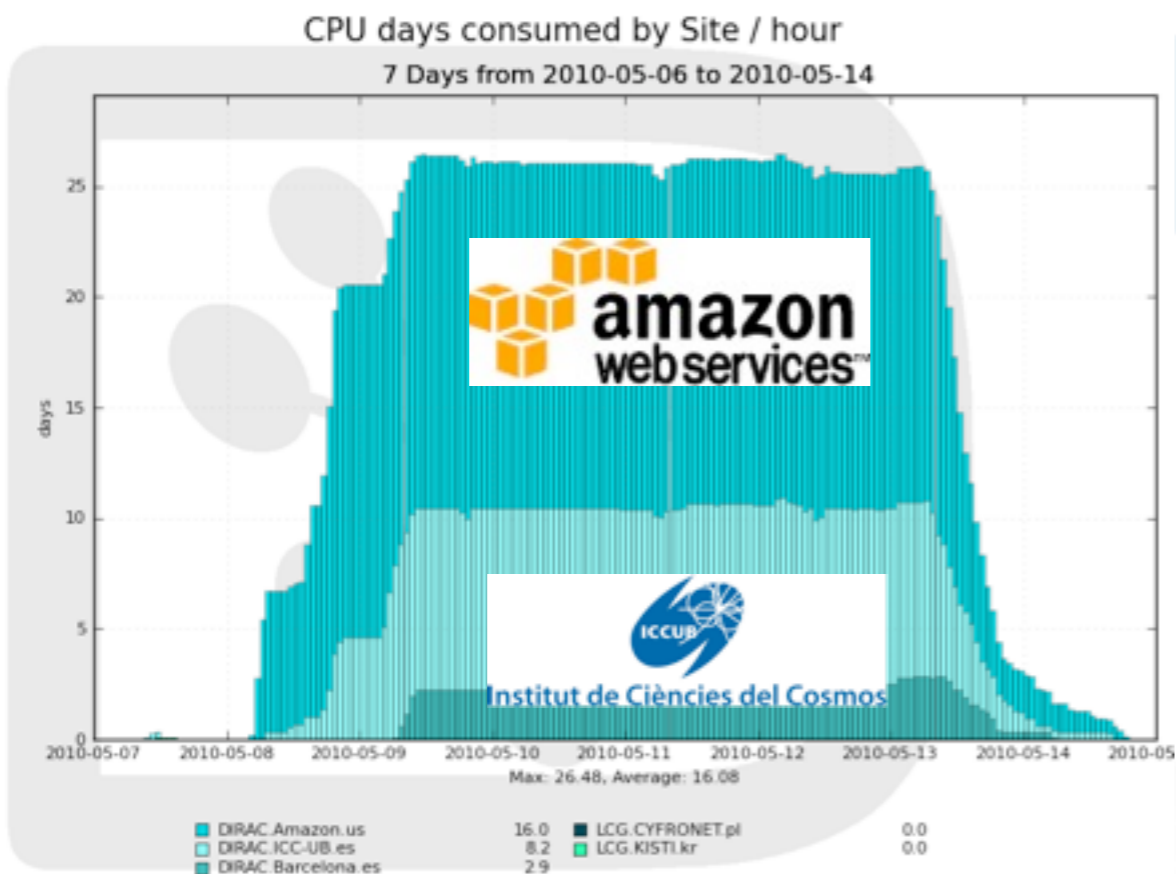




# Next steps



# Cloud + Local: Phase II



# Results (II)

- Phase II (local + cloud integration):

- production ready:

- 7% of Belle production in 6 days
- 170 M evt (~3.6 TB)
- 3100 CPU days

- proven interoperability:

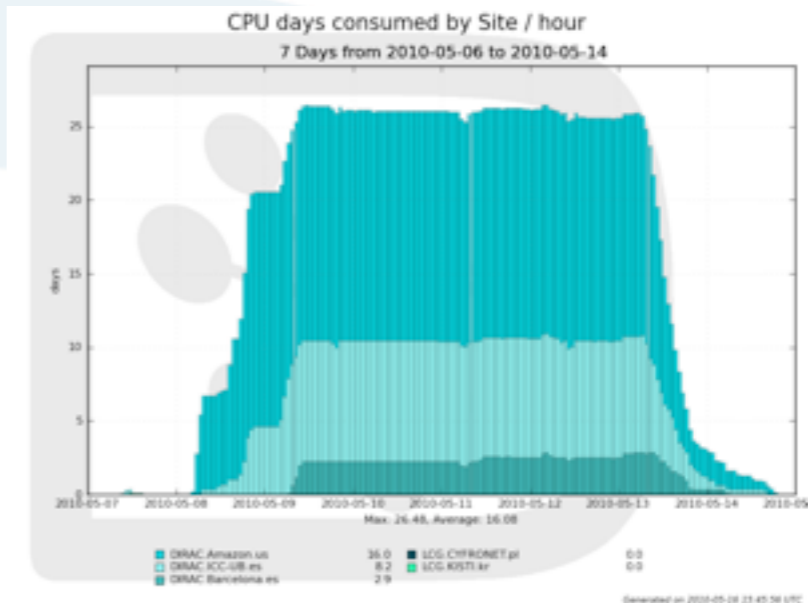
- 60 % cloud resources / 40 % local resources
- transparent integration of heterogeneous resources
- > 95 % efficiency in CPU usage

- improved cloud cost:

- 0.20 USD / 10k evt

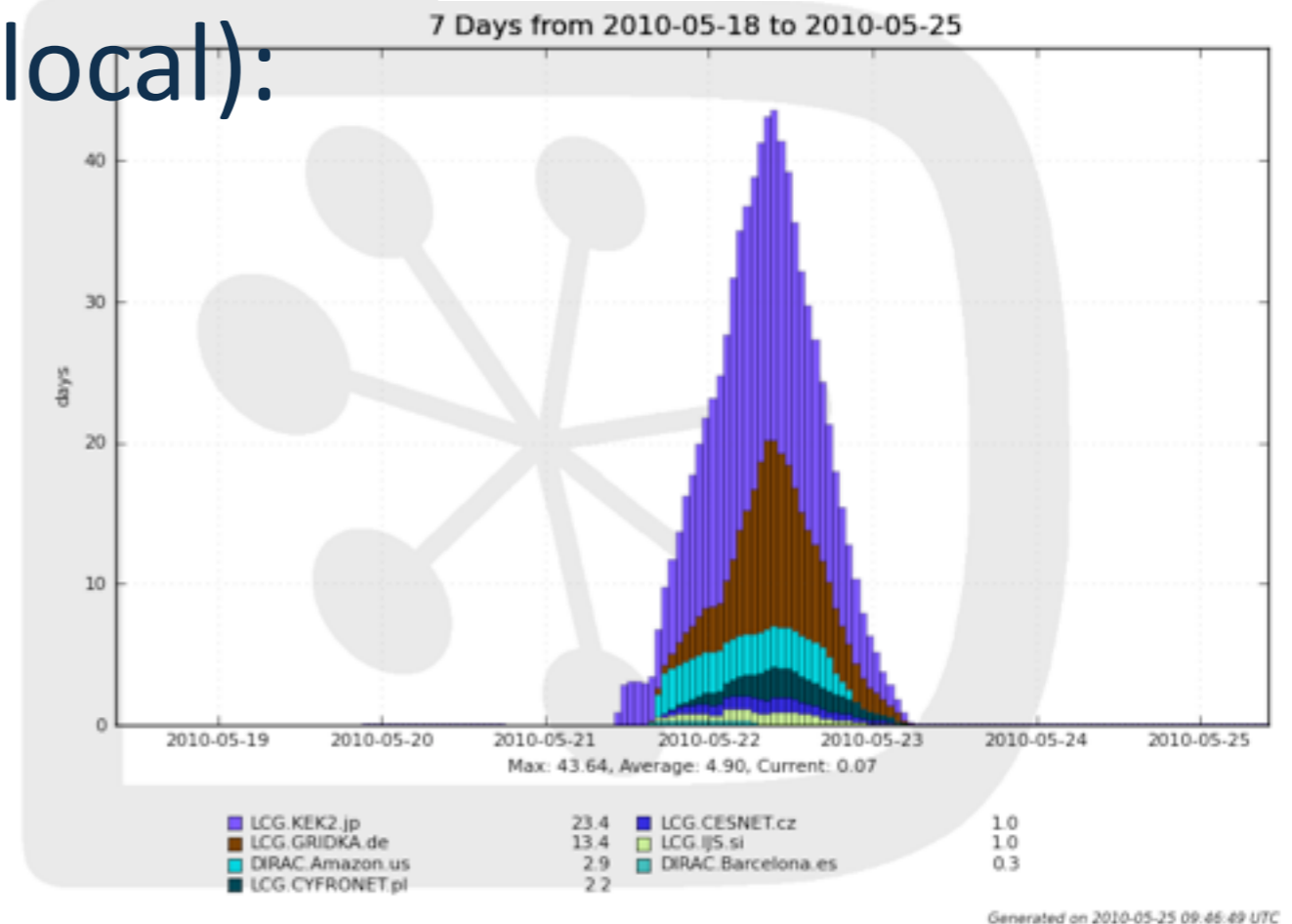
- input data downloaded from KEK SE

– issues with stability of grid Storage



# Results (III)

- Phase III (grid + cloud + local):
  - Currently running:
    - 800 CPU days
    - Full interoperability:
      - cloud + grid + local
      - stability issues on grid SE.
      - main effort to make sure grid environment is appropriate
  - Other issues:
    - Jobs can be very long, 1-30 days, only the shorter ones are appropriate for the grid



# Summary and Outlook



- **DIRAC** is a flexible framework to implement any distributed computing model.
- **DIRAC** allows integration of any kind of computing resources: grids/local/clouds/...
- **DIRAC** brings many added values: interactive web portal/monitoring/accounting/...
- Being considered by Belle II, MAGIC, PIC, ILC,...
- Are you interested?
  - contact: [dirac.project@gmail.com](mailto:dirac.project@gmail.com)