

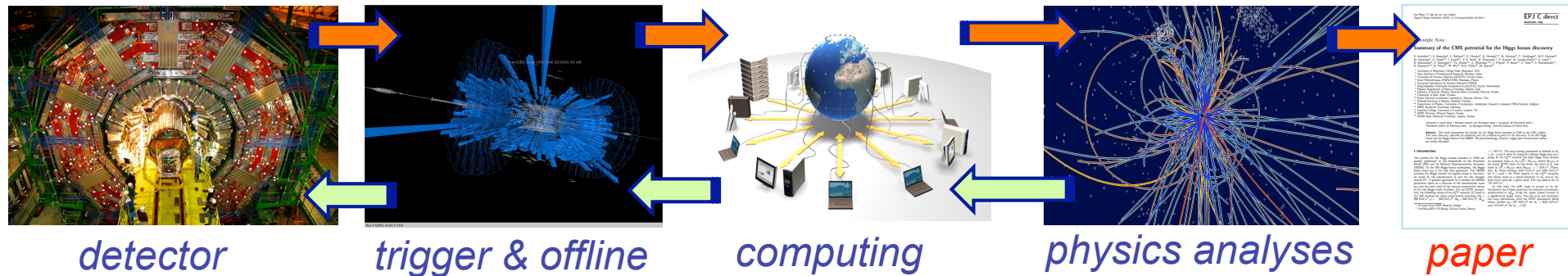
# Computing efforts supporting Physics Analyses

- reminder on the main aim of the Analysis Support Task Force
- survey of existing tools available to monitor/diagnose/communicate
- survey of how we use our GRID-like computing resources
- proposal to setup an Analysis Operation team within Computing

Jorgen D'Hondt – James Letts

*M. Calloni, D. Colling, J. Klem, J. Maes, G. Maier, S. Padhi, H. Pi, S. Sarkar, P. Van Mulders, I. Villella, et al.*

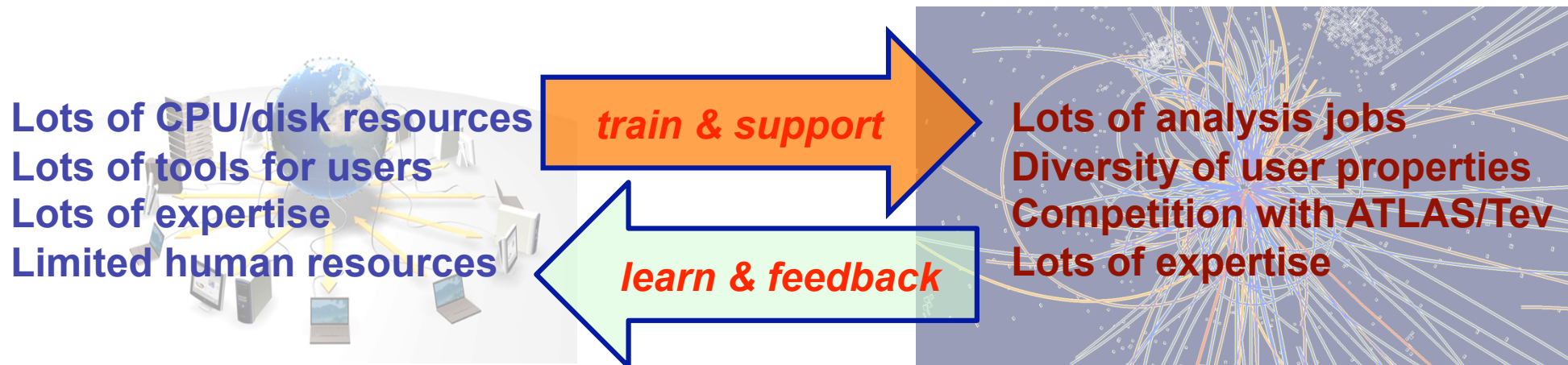
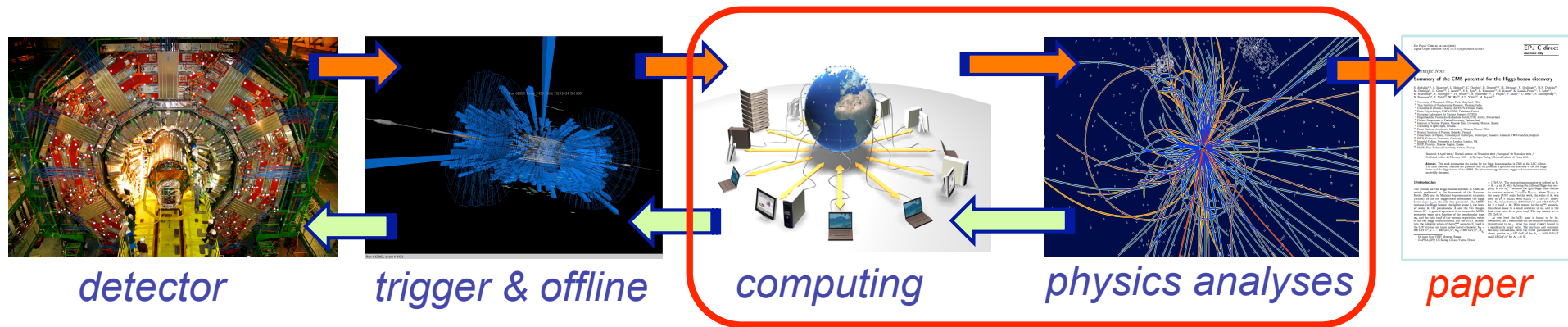
- On first sight one project has nothing to do with the other, but in reality they are only successful projects when they are coherent...



(simplified illustration!)

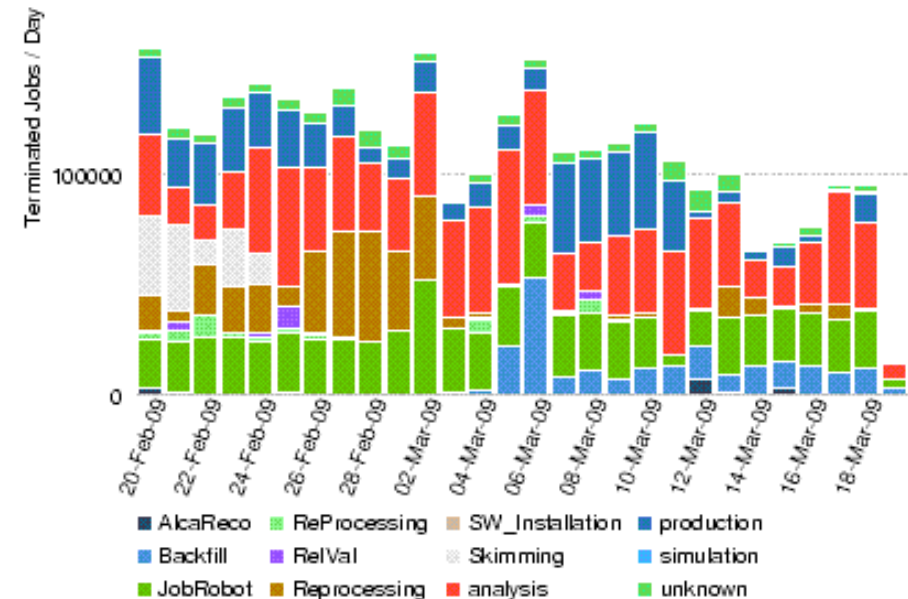
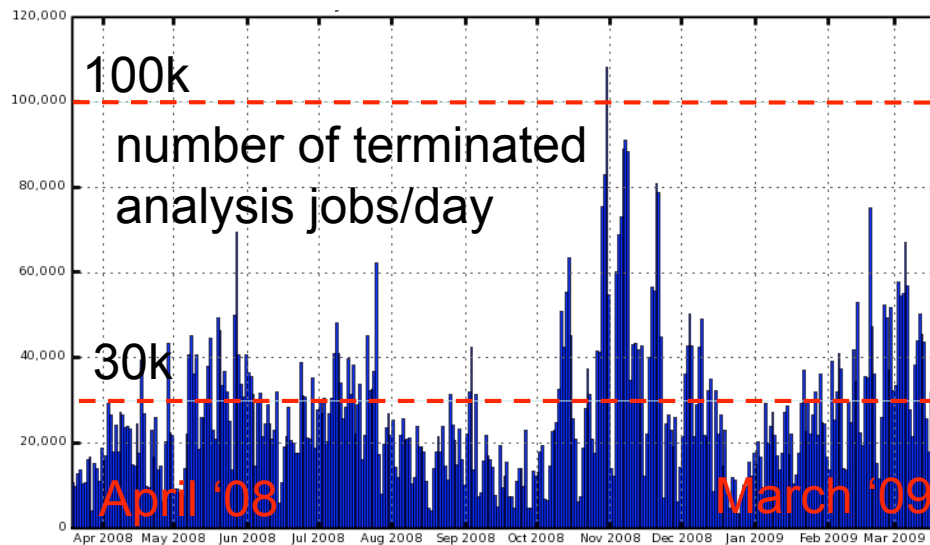
# Computing and Physics

- On first sight one project has nothing to do with the other, but in reality they are only successful projects when they are coherent...



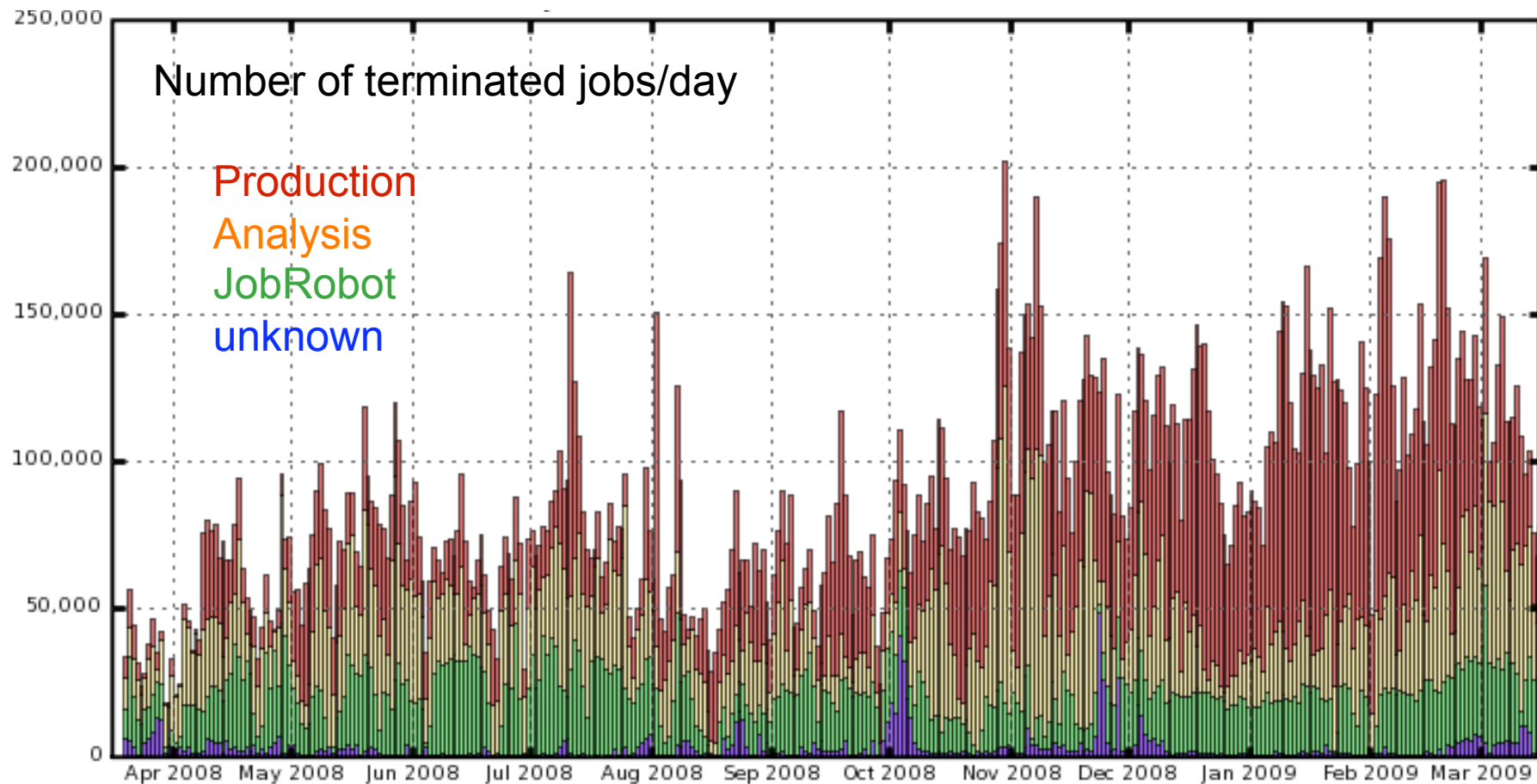
- Our success is highly correlated with the degree of coherence!**

- In 2008 in total 700 individuals submitted analysis jobs in CMS with an average of ~30k jobs/day and a success rate of ~55%
- Our computing model asks to scale this to 100k-200k jobs/day



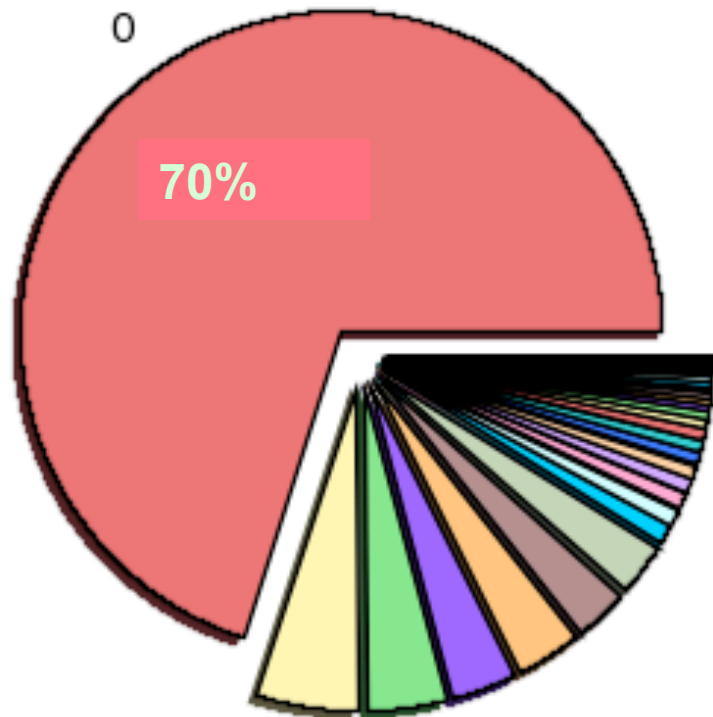
- **A joint effort from Users and Computing is needed to make this scaling efficient (both for the success rate and the human resources)**
  - ✓ Physics analysis experts should invest time to learn working correctly with our computing resources and the (future) support procedures
  - ✓ Computing experts should invest to provide adequate tools and support

- The increase in activity is already visible



# Failed analysis jobs

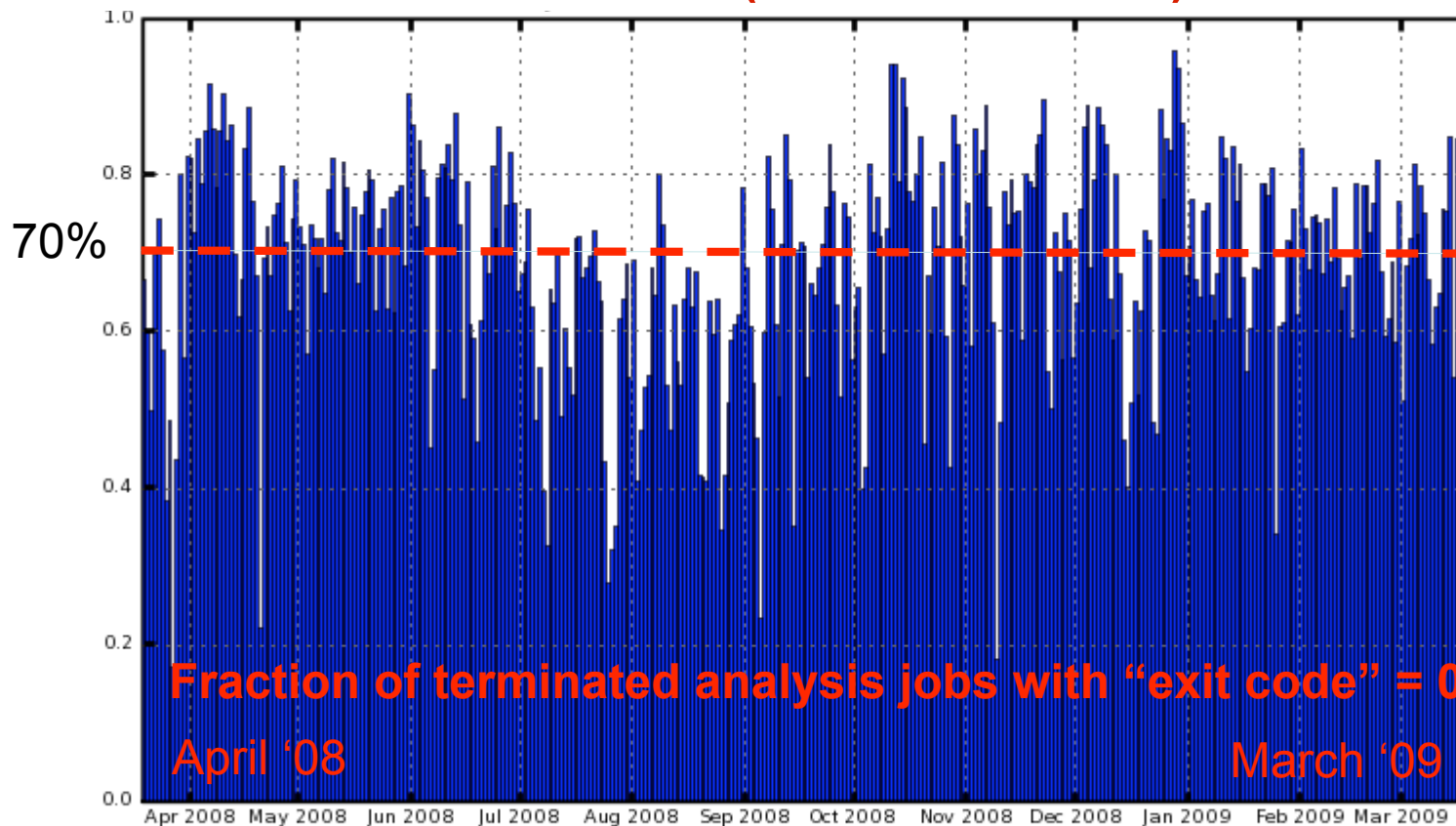
- Reasons why ~45% of the analysis jobs do not terminate successful
- And the list of causes is dynamical...



- ✓ In total ~25% of all submitted jobs are aborted or cancelled, these are not taken into account in this chart
- ✓ Of the ~75% remaining (=terminated) jobs the “exit code” over the last year is shown
- ✓ ~30% of them have a non-zero “exit code”
- ✓ In total we had 9,8M jobs submitted

- Monitoring the jobs to find failures and act pro-actively, and diagnose the cause to communicate to developers is a key point to increase the efficiency of our computing and human resources

- Evolution of the success rate (zero “exit code”) of terminated jobs



- Translate this into the fraction of CPU or “WallClock time” wasted
- A key number to quantify the efficiency of the use of our computing resources: average is ~80%, but with significant ~10% fluctuation

# Analysis Support Task Force

- **The Task Force was setup in October 2008 to sketch the workplan/ model of a future Analysis Operation team by March 2009**
- **Short overview of main charges of this Task Force**
  - ✓ Evaluate how we can adapt the link between Computing and Physics to optimize the use of our computing and human resources
  - ✓ Evaluate the tools and information needed to improve our understanding of failures of CRAB submitted jobs
  - ✓ Evaluate the tools needed to provide user support to pro-actively address issues with user jobs and TIER sites
  - ✓ Evaluate the human resources needed to run Analysis Operations
  - ✓ Long version: <https://twiki.cern.ch/twiki/bin/view/CMS/AnalysisSupportTaskForceCharge>
- **This Task Force has mainly a survey and exploration character**
- **Started with a kick-off workshop on the 16<sup>th</sup> of December 2008, followed by weekly meetings**
- **Twiki page:** <https://twiki.cern.ch/twiki/bin/view/CMS/AnalysisSupportTaskForce>



# Survey of Monitoring tools

- **Survey of existing tools useful for Analysis & Analysis Support**
  - ✓ CRAB – the CMS job submission tools (strong development team)
  - ✓ CRAB Feedback mailing list for support
  - ✓ DashBoard to monitor the jobs in all stages and to archive the information
  - ✓ Ticketing system (savannah) deployed but not heavily used
  - ✓ Education & documentation via tutorials and workbooks
  - ✓ PhEDEx to place data according to the needs and to manage disk resources
  
- **We have identified several areas where we would like to develop extensions of the functionalities of these (and other) tools, in order to serve our purpose with an Analysis Operation team**
- **Details:** <http://indico.cern.ch/getFile.py/access?contribId=25&sessionId=5&resId=0&materialId=slides&confId=54515>

# Survey of Monitoring tools

## DashBoard (current web interface to job monitoring in CMS)

<http://dashboard.cern.ch/cms/>

- **Archives and reports on a specified collection of CRAB jobs at various stages of the job submission, execution & completion, with “exit codes” for different features**
    - ✓ *the tool to monitor the performance of our system*
    - ✓ *worked with DashBoard developers to be able to have quick access*
  - **There is no direct access however to the users log files**
    - ✓ *would be needed to diagnose the observed failures*
    - ✓ *possible via the CRAB Server at the end of the job*
    - ✓ *possible with WM Core DashBoard API for real-time reporting*
  - **Define thresholds or metrics which our system should pass**
  - **One can also perform data mining within the information to search for “Association Rules” between feature and cause**
- ➔ **an essential tool as basis for the development of monitoring tools**

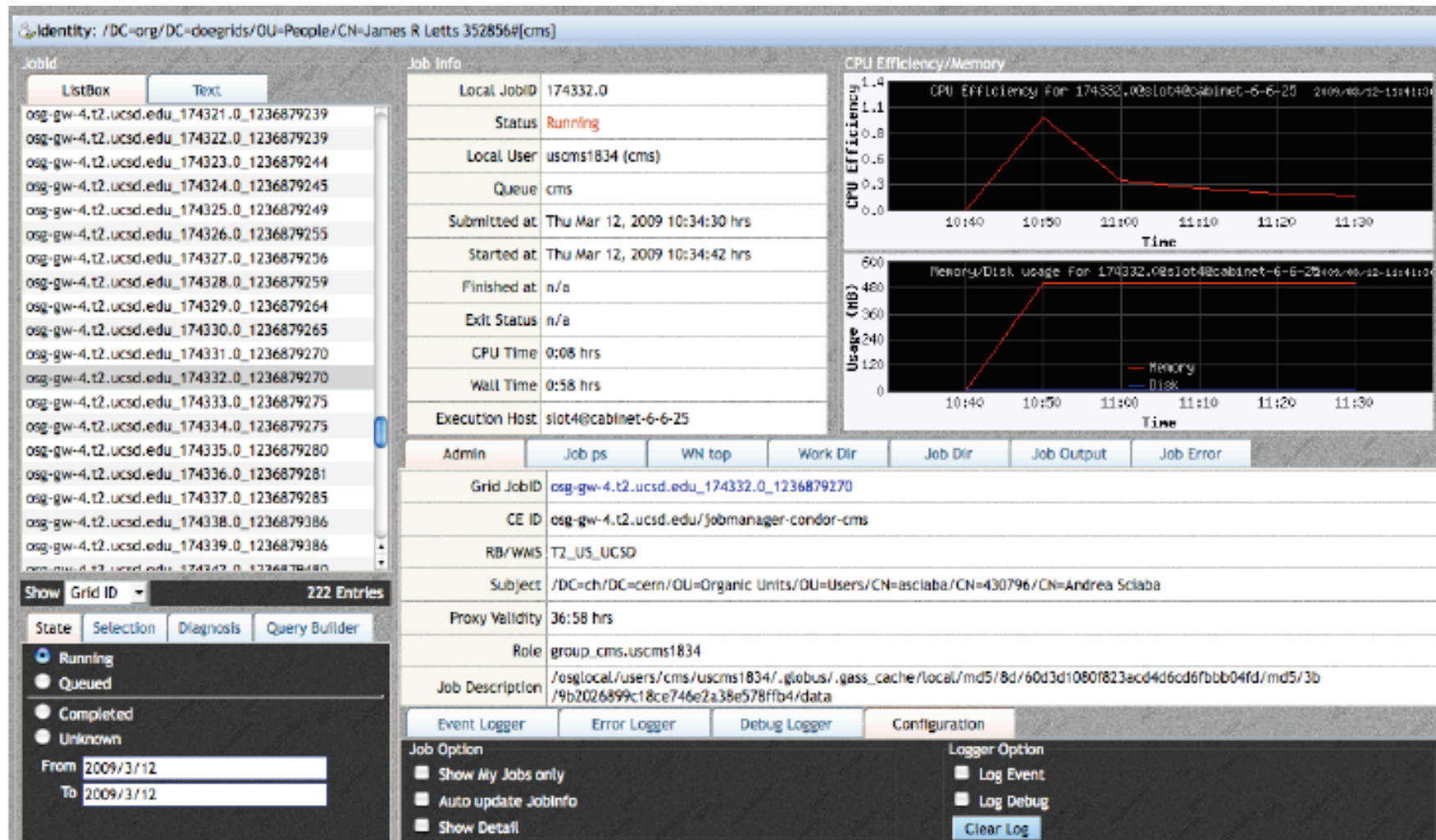
## Reference site(s) (non existing tool today)

- Deploy a few reference sites where the user can test his failed job before contacting the computing support structure
  - ✓ *user survey (see later) indicates the user is very positive on this*
- The candidate TIER-2 site(s) should pass a pre-defined performance metric before being considered, hence dynamical
- Small standard data samples to be stored for tests
- Install enhanced real-time monitoring and diagnose tools
  - ✓ *JobMon software is being deployed and tested*
  - ✓ *Interfaces to LSF, PBS, CONDOR & CONDOR Quill DB*
  - ✓ *Broader access for admins & developers*
- Analysis Operations could have access to the log files of failed jobs, hence they deal only with log files from the reference site(s)

→ a good idea to follow in a layered support structure

# Survey of Monitoring tools

- Example of the JobMon tool with a CONDOR Quil DB batch system



Identity: /DC=org/DC=doegrids/OU=People/CN=James R Letts 352856[cms]

JobID

Listbox	Text
osg-gw-4.t2.ucsd.edu_174321.0_1236879239	
osg-gw-4.t2.ucsd.edu_174322.0_1236879239	
osg-gw-4.t2.ucsd.edu_174323.0_1236879244	
osg-gw-4.t2.ucsd.edu_174324.0_1236879245	
osg-gw-4.t2.ucsd.edu_174325.0_1236879249	
osg-gw-4.t2.ucsd.edu_174326.0_1236879255	
osg-gw-4.t2.ucsd.edu_174327.0_1236879256	
osg-gw-4.t2.ucsd.edu_174328.0_1236879259	
osg-gw-4.t2.ucsd.edu_174329.0_1236879264	
osg-gw-4.t2.ucsd.edu_174330.0_1236879265	
osg-gw-4.t2.ucsd.edu_174331.0_1236879270	
osg-gw-4.t2.ucsd.edu_174332.0_1236879270	
osg-gw-4.t2.ucsd.edu_174333.0_1236879275	
osg-gw-4.t2.ucsd.edu_174334.0_1236879275	
osg-gw-4.t2.ucsd.edu_174335.0_1236879280	
osg-gw-4.t2.ucsd.edu_174336.0_1236879281	
osg-gw-4.t2.ucsd.edu_174337.0_1236879285	
osg-gw-4.t2.ucsd.edu_174338.0_1236879386	
osg-gw-4.t2.ucsd.edu_174339.0_1236879386	
osg-gw-4.t2.ucsd.edu_174340.0_1236879480	

Show Grid ID 222 Entries

State Selection Diagnosis Query Builder

- Running
- Queued
- Completed
- Unknown

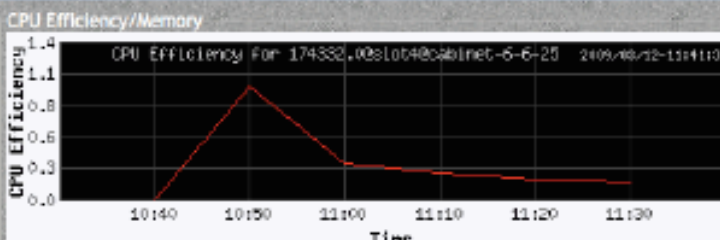
From 2009/3/12 To 2009/3/12

Job Info


Local JobID	174332.0
Status	Running
Local User	uscms1834 (cms)
Queue	cms
Submitted at	Thu Mar 12, 2009 10:34:30 hrs
Started at	Thu Mar 12, 2009 10:34:42 hrs
Finished at	n/a
Exit Status	n/a
CPU Time	0:08 hrs
Wall Time	0:58 hrs
Execution Host	slot4@cablnet-6-6-25

CPU Efficiency/Memory

CPU Efficiency For 174332.0@slot4@cablnet-6-6-25 2009/03/12-11:11:00



Memory/Disk usage For 174332.0@slot4@cablnet-6-6-25 2009/03/12-11:11:00



Admin Job ps WN top Work Dir Job Dir Job Output Job Error

Grid JobID	osg-gw-4.t2.ucsd.edu_174332.0_1236879270
CE ID	osg-gw-4.t2.ucsd.edu/jobmanager-condor-cms
RB/WMS	T2_US_UCSD
Subject	/DC=ch/DC=cern/OU=Organic Units/OU=Users/CN=asclaba/CN=430796/CN=Andrea Sclaba
Proxy Validity	36:58 hrs
Role	group_cms.uscms1834
Job Description	/osglocal/users/cms/uscms1834/.globus/.gass_cache/local/md5/8d/60d3d1080f823acd4d6cd6fbbb04fd/md5/3b/9b2026899c18ce746e2a38e578ffb4/data

Event Logger Error Logger Debug Logger Configuration

Job Option

- Show My Jobs only
- Auto update JobInfo
- Show Detail

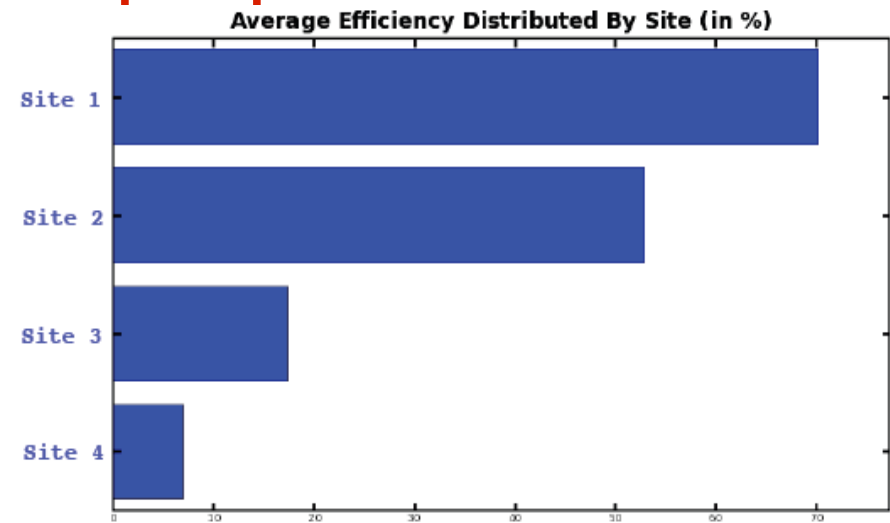
Logger Option

- Log Event
- Log Debug

Clear Log

# First examples of Monitoring

- **Perform same task at different sites to compare performance**
  - ✓ *to identify data access inefficiencies*





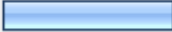


- **Understanding jobs with an unknown “exit code”**
  - ✓ *detailed comparison between DashBoard CPU usage accounting and local batch system accounting: found to be consistent to ~10% level*
  - ✓ *during this study it was found that 2% of the jobs were killed because they reach the threshold of maximum 36h running (site dependent)*
  - ✓ *because these jobs are long compared to usual jobs they take ~50% of the resources being wasted*

→ **these studies indentify the main areas to consider for improvements**

# Survey of User properties

## Survey of how people perform physics analyses

- In total 226 people participated (about 50-70% of the active users)
- **Results:** <https://twiki.cern.ch/twiki/bin/view/CMS/AnalysisSupportTaskForceSurvey>
- This is taken as guidelines on how the user would like to proceed

1. For how many months or years have you been doing analysis in CMS?			
		Response Percent	Response Count
less than 6 months		10.2%	23
between 6 months and 1 year		11.9%	27
<b>between 1 year and 3 years</b>		<b>47.8%</b>	<b>108</b>
between 3 years and 5 years		19.5%	44
more than 5 years		10.6%	24
		<i>answered question</i>	<b>226</b>
		<i>skipped question</i>	<b>0</b>

*Thanks to many people who have contribute (especially Ilaria Villella)*

**... and 91% have used CRAB**



# Survey of User properties

## Volume of data produced during last 6 months?

< 1TB		50.0%
1 - 10TB		27.6%
10 - 100TB		2.6%
> 100TB		1.3%
I don't know because I have never bothered		11.5%
I don't know how to track this		7.1%

*Typical properties of user analysis via CRAB*

## Typical amount of jobs per task?

< 10		7.7%
10 - 100		48.1%
100 - 1000		39.1%
1000 - 10000		5.1%
> 10000		0.0%

**An average “analysis round” takes**

<b>few hours</b>	<b>31%</b>
<b>1 day</b>	<b>14%</b>
<b>2-5 days</b>	<b>39%</b>
<b>1 week</b>	<b>7%</b>
<b>&gt;1 week</b>	<b>7%</b>
<b>&gt;1 month</b>	<b>1,5%</b>

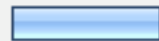
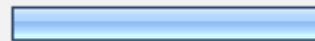
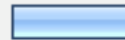

## Category 1: Do the users use CRAB as a black box?

- **In general NO, which is positive!**
  - ✓ *17% of the CRAB users tried to hack CRAB to include new or better functionalities*
  - ✓ *these users do not answer differently on the questions however*
- **68% know they are using the latest version (CRAB\_2\_4\_4)**
- **50% uses the CRAB server**
  - ✓ *80% of the remaining users claim they do not need it*
- **74% knows you cannot run jobs at T1's**
- **63% often black/white lists sites, they get information from**
  - ✓ *prior experience* 53%
  - ✓ *availability & reliability* 21%
  - ✓ *available CMSSW version & datasets* 14%
  - ✓ *whitelist the local T2* 6%
  - ✓ *via information on CRAB hypernews* 3%
- **Those who do not black/white list either don't know how to do this (33%) or mention it doesn't help them (66%)**
- **All this does not correlate with the period the user spend in CMS**



## Category 2: For what purpose they use CRAB?

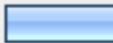
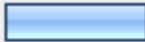




It is found that the application of CRAB is not correlated with the answers on the other questions

13. Do you use CRAB to (multiple answers possible)			
		Response Percent	Response Count
Produce simulated events		43.4%	66
Perform my analysis on existing samples		90.8%	138
Make private PAT tuples		33.6%	51
Make private skims from PAT tuples		19.7%	30
Why don't you use the official ones, please elaborate			31
			<b>answered question</b> 152
			<b>skipped question</b> 74

*"CRAB is a very nice tool, with simple commands, in practice it's very helpful for non-experts on grid like myself"* – a participant to the survey

# Survey of User properties

## How do you stage-out your output files ?

Return them back to my machine (return_data = 1)		27.9%
Copy them remotely (copy_data = 1) to a T2 site		36.1%
Copy them remotely (copy_data = 1) to a T3 or other site		8.2%
Copy them remotely (copy_data = 1) to my CASTOR area because I'm used to it and it's more convenient		15.0%
Copy them remotely (copy_data = 1) to my CASTOR area because I don't know which T2 I could write to		4.8%
Copy them remotely (copy_data = 1) to my CASTOR area because I'm not aware of other way to stageout than to CASTOR at CERN		8.2%

Those who produce >1TB of data

CASTOR @ CERN  
28%

# Survey of User properties

## Do you publish your output files in private DBS?

Yes		22.6%
<b>No, because I don't need this functionality</b>		<b>54.1%</b>
No, because I was not aware of this functionality		3.4%
No, because I don't know how to do that		12.3%
No, because I am writing my output files to my CASTOR area at CERN so the publication is not supported		7.5%

This does not correlate with the type of jobs, being analysis jobs on existing samples, to produce simulated events, making private PAT tuples or making private skims from PAT tuples.

# Survey of User properties

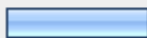

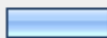

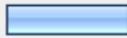
## Category 3: Why don't they use CRAB? (~10% of all users)

- 6% of all users run their analysis interactively
- 3% of all users use another tool to run their analysis on the GRID
- The users who do not use CRAB have arguments like:
  - ✓ *too slow* 60%
  - ✓ *not reliable* 50%
  - ✓ *difficult to deal with failed jobs* 30%
  - ✓ *don't need it due to adequate alternative* 30%
- 60% of the non-CRAB users have followed a tutorial however




Maybe the outcome of these survey can convince the remaining 10% users to use CRAB.

## Category 4: What do the users think about the CRAB performance?

- In general the “analysis round” time depends on the amount of data, the sample analyzed and the GRID performance. It takes weeks however to get CRAB understood.
- The whole chain is mentioned to be much more reliable now.
- It can be painful when dealing with a huge number of jobs (book-keeping aspects)
- A fraction of jobs still fail most of the time, and it takes “90% of the time for the 10% of jobs that are not running properly”

Reasons to use the CRAB server (50% of users)		
I have never tried without using the server		45.3%
I want to make use the "tools" which are only available when using the server, e.g: receive an email when jobs are done		20.0%
I want to make use the "tools" which are only available when using the server, e.g: "unlimited" size of input sandbox		32.0%
I want to make use the "tools" which are only available when using the server, e.g: easier tracking of problems		24.0%
I want to make use the "tools" which are only available when using the server, e.g: automatic resubmission of failed jobs		38.7%

## Reasons to NOT use the CRAB server (50% of users)

I never heard of it		20.3%
I don't see any advantage in using it		17.2%
<b>I never tried it because my jobs run correctly most of the time without using the server</b>		<b>46.9%</b>
I never tried it because I don't know where I can find information on how use it		15.6%



# Survey of User properties

## Category 5: How do the users deal with the failure of CRAB jobs?

	always	most of the time	sometimes	never
Yes, by looking in .stderr, .stdout and .lof file	62.3% (91)	28.1% (41)	8.2% (12)	1.4% (2)
Yes, by asking a colleague	5.6% (7)	13.5% (17)	66.7% (84)	14.3% (18)
Yes, by using the using CMS Dashboard Interactive view for monitoring: <a href="http://dashb-cms-job.cern.ch/dashboard/request.py/jobsummary">http://dashb-cms-job.cern.ch/dashboard/request.py/jobsummary</a>	6.3% (8)	7.9% (10)	35.7% (45)	50.0% (63)
Yes, by using the Dashboard Task monitoring: <a href="http://dashb-cms-sam.cern.ch/dashboard/request.py/taskmonitoring">http://dashb-cms-sam.cern.ch/dashboard/request.py/taskmonitoring</a>	6.4% (8)	11.2% (14)	34.4% (43)	48.0% (60)
Yes, by searching in twiki pages available through: <a href="https://cmsdoc.cern.ch/cms/ccs/wm/www/Crab">https://cmsdoc.cern.ch/cms/ccs/wm/www/Crab</a> or <a href="https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideCrab">https://twiki.cern.ch/twiki/bin/view/CMS/SWGuideCrab</a>	10.9% (14)	22.7% (29)	47.7% (61)	18.8% (24)
Yes, by using the Grid Analysis Job Diagnosis guide: <a href="https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookGridJobDiagnosisTemplate">https://twiki.cern.ch/twiki/bin/view/CMS/WorkBookGridJobDiagnosisTemplate</a>	1.7% (2)	8.5% (10)	15.4% (18)	74.4% (87)
Yes, by sending an email to the crabFeedback hypernews	4.0% (5)	10.4% (13)	44.8% (56)	40.8% (51)
Yes, by searching in the threads of the crabFeedback hypernews if the problem already occurred	7.9% (10)	23.0% (29)	49.2% (62)	19.8% (25)
Yes, by opening a Savannah ticket	0.0% (0)	0.8% (1)	14.4% (17)	84.7% (100)
No, I just resubmit the jobs which failed and when they fail again, I look into the problem	8.5% (10)	17.8% (21)	50.0% (59)	23.7% (28)
No, I run my jobs interactively when it does not work with CRAB	4.4% (5)	6.1% (7)	29.8% (34)	59.6% (68)

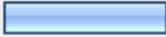
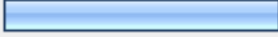


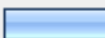

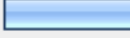


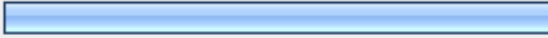

# Survey of User properties

- ~90% looks at the standard output and standard error log files
- ~70% of the users test their code interactively before sending via CRAB
- Few use the available monitoring tools (eg. DashBoard), mostly because they are unfamiliar with them → *importance of streamlining education & documentation*
- About 50% of the users use the CRABFeedback hypernews to browse for answers or ask for help
  - ✓ ~80% of them report they receive useful feedback from this list
- About 50% of the users simply resubmits the jobs and only investigate when they fail twice
- It seems that the way people proceed for failed jobs is divers...
  - ✓ 73% would use a Reference Site to debug a job
  - ✓ 78% would use a snapshot tool of a running job (Real-time Job Monitoring)


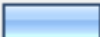
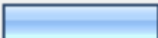
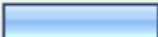


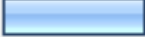


# Survey of User properties

## When my CRAB jobs fail it is most often due (3 possible answers):

related to a bug in my code (including pset.cfg)		20.5%
related to a bug in my crab.cfg (e.g. wrong output filename, ...)		34.9%
during creation		3.4%
during creation because the input sandbox size is too big		9.6%
during submission		13.0%
during stageout		39.0%
because the data is not available		16.4%
because the site is not available		20.5%
because the CMSSW version is not available		9.6%
<b>because the jobs are aborted</b>		<b>69.9%</b>
issues related to Grid certificates		2.7%

## Category 6: Do the users find enough adequate information?

What is your experience with the twiki pages with CRAB documentation?		
I did/do not know (most of) them		8.7%
I know they existed, but I never tried to use them		14.5%
I try to use them, but I do not find the solution for my failed jobs in there		23.1%
I find the information to debug my failed jobs there most of the time		23.1%
They are not up to date		5.2%
I don't find my way in there		4.6%
Most of the things I need are in there		20.8%

**Only 43% is positive, hence an area to improve**

# Survey of User properties

indicate up to 3 areas where you think major improvement are needed

	less important	important	very important
Grid overhead	26.2% (27)	<b>50.5% (52)</b>	23.3% (24)
Simplicity	22.0% (26)	34.7% (41)	<b>43.2% (51)</b>
Reliability of whole chain	6.3% (9)	27.5% (39)	<b>66.2% (94)</b>
Job monitoring	17.6% (22)	<b>44.0% (55)</b>	38.4% (48)
Job error handling	9.2% (12)	43.8% (57)	<b>46.9% (61)</b>
Possibility for debugging	11.9% (14)	<b>45.8% (54)</b>	42.4% (50)
Documentation	14.0% (17)	<b>47.1% (57)</b>	38.8% (47)
Tutorial	<b>48.1% (50)</b>	35.6% (37)	16.3% (17)
Support	24.7% (24)	<b>43.3% (42)</b>	32.0% (31)
Speed and through-put	28.0% (30)	<b>40.2% (43)</b>	31.8% (34)
Limited output storage space	29.9% (32)	<b>35.5% (38)</b>	34.6% (37)
Everything is fine	<b>46.0% (23)</b>	42.0% (21)	12.0% (6)

To increase the success rate we proceed in 4 coherent directions:

- 1. Development of tools to monitor & diagnose failures pro-actively**
  - *Extend the existing tools to allow for real-time monitoring*
  - *Develop metrics to judge the quality of the performance*
  - *Close collaboration with developers (DashBoard, JobMon, ...)*
- 2. Setup of a monitoring task & communication lines**
  - *Proposal to include this into foreseen computing shifts*
  - *Strong communication line to developers to solve the problems and to users to pro-actively steer their use of the computing resources*
- 3. Deploy a layered support system**
  - *Install and maintain some reference sites, design system to get access to the logfiles... force users to first try their jobs on these sites*
  - *Streamline the twiki pages and hypernews like support*
- 4. Prepare for a pro-active use of prior information (“phase II”)**
  - *Design, develop and test an automatic procedure to include prior information into the CRAB system*
  - *This should reduce the workload of support & increase the success rate*

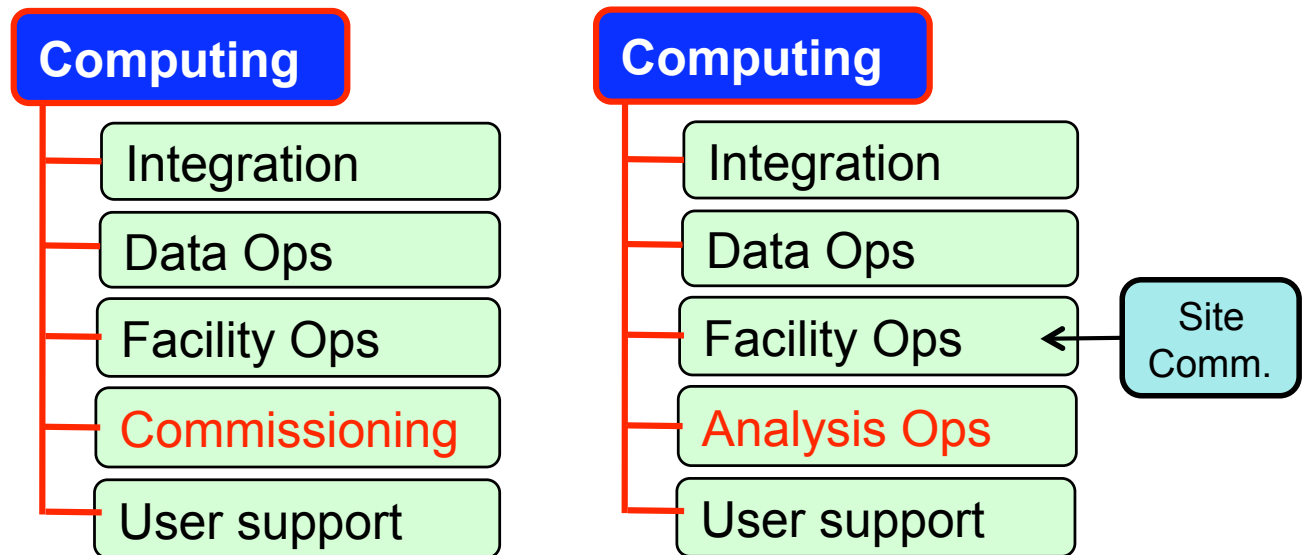
**Small human investment from the collaboration, but the return is worth it.**

- Analysis Operations will monitor the performance, discover problems and report to developers/users/CRAB. It will also provide, deploy and test all the extra tools needed for this.
- It will form somewhat the glue between the different stakeholders.

## Potential Future

### Organizational Change:

*In the continuing transition from Commissioning to Operations, the computing project proposes to change the level-2 structure*



- We will not jump blind in this adventure (eg. Task Force), but we will increase step-by-step the effort... because there will be a learning curve for all of us, aim to be fully operational by autumn 2009.
- *Today you should still send your questions to the appropriate forums!*