

CRAB Status

Stefano Lacaprara

INFN LNL

Joint Offline & Physics Meeting
CERN, 23 October 2007



Current Status

- We are today supporting both sl3 and sl4 with two separate branches
- CRAB 201 for sl4 and 154 for sl3

CMS policy for sl3/sl4

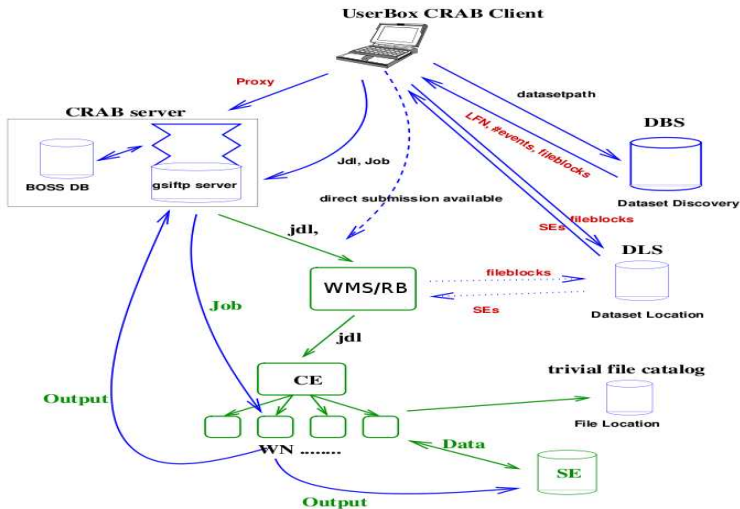
- CMSSW < 1.5.0 **SL3**
 - CMSSW \geq 1.5.0 **SL4**
 - try always to use the latest release of a given serie (1_X...)
- On SL3, correct environment is: UI, eval '...', crab.csh
 - **Support for CRAB Server in both releases**
 - More info on techicalities on Offline-Computing Workshop's talks last week.

Copy to SE

- coping large ($> \mathcal{O}(10\text{ Mb})$) via output sand box is not viable: you **MUST** use copy to remote SE
- **More robust copy to remote SE** Lot of problems lately: should be gone (almost)
- Working on a backup solution in case of copy failure.
 - copy output to SE close to CE
 - provide user with a crab command to copy it from close SE to final destination asynchronously
- **Storage management for User**
- Where is user supposed to copy its output? T2? Do the T2 knows about that? Do they have space dedicated? Do the user know about that?
- A task for AnalysisOps (see below)



CRAB Server WorkFlow



Server

- CRAB standalone submits jobs directly to the grid (RB/WMS/condor_g)
- With server, CRAB delegate jobs handlign to the server, which in turn submits, check status, get output, etc ...
- CRAB interface for server is **exactly** the same as for standalone: `crab -create -submit -status -kill -get ...`
- Just add to your `crab.cfg`
 - `server_mode = 1`
 - `server_name = crabas.lnl.infn.it/data1/cms/`
- Interaction with server is much faster, since it acts as a sort of cache for all actions

- Moreover: automatic resubmission, get-output, mail notification, status querying, ...
- Server is provided as a *service* to the analysis community
- Daily operation are followed by a team of people
- More info available on Twiki

<https://twiki.cern.ch/twiki/bin/view/CMS/Tutorial12jun>

Use the server!

StandAlone crab will continue, but if you have many jobs ($\mathcal{O}(100)$), please use the server.

We need feedback!

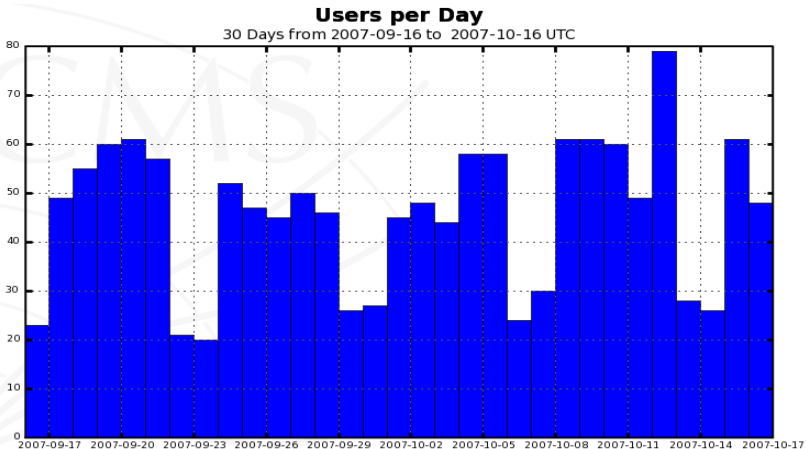


CRAB usage (from dashboard)



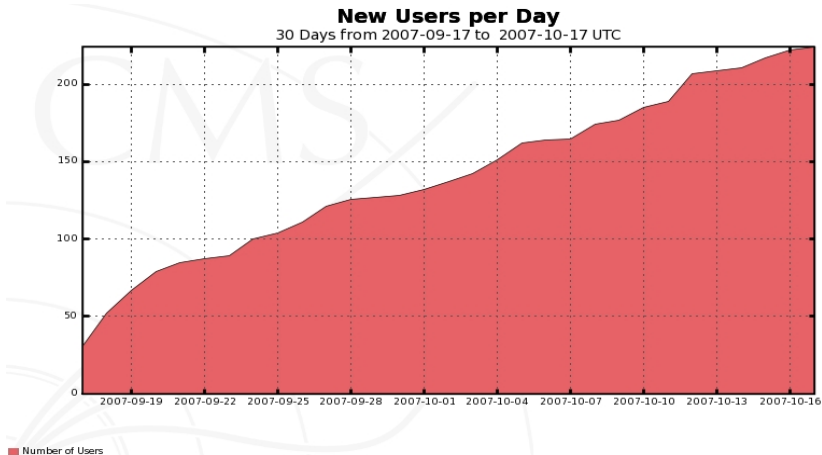
Reached ~ 19 *kjob/day* during test, and ~ 27 *kjob/day* last week!

CRAB users per day



Average different users per day 46 (including we)

CRAB unique users vs time



242 different users from 16th Sept. to 16th Oct.

User support

- With so many users, user support is more and more critical
- Lot of activity on the feedback list

We are trying to further improve the support on the list

- shift to have one person responsible for the support at any time;
 - try to give a preliminary answer asap, within the day, even just to say: *we are looking at your problem*
 - whoever first answer to a help request, is responsible to carry on the thread until a solution
-
- It's a FTE task! peaks of 50 mail per day from different users
 - Many problems are outside what crabbers can do (site problem, WMS problems, etc ...): we try to involve proper people to fix

CRAB Documentation

- FAQ, Twiki, help page, tutorial: all linked from CRAB web page, renewed and updated
<http://cmsdoc.cern.ch/cms/ccs/wm/www/Crab/>



The screenshot shows the CRAB web page with the following content:

- Source:**
CRAB_2_0_0 has been released. The tarball can be found here [CRAB_2_0_0.tar](#) or [CRAB_2_0_0.tar.gz](#)
At CERN it has already been installed and can be used by sourcing the usual [/afs/cern.ch/cms/ccs/wm/scripts/crab/crab.sh](#)
- What's new:**
This release is intended to support **both** the submission modality:
Direct job submission: the normal CRAB standalone which allow direct submission to the Grid.
Job submission through the server: the new CRAB feature enabling the job submission to a CRABServer which take care to perform actions [submission, check status, get-output, error handler, resubmission...] for the users.
The switch between the two modality is easy and fast. For more further informations about the CrabServer see also: [the help page](#)
[Direct link to [the how-to](#) use the server.]
- Fixed bugs and improvements:**
 - dashboard messages related to dataset
 - listmatch for glite and gitecoill (for MC production)
 - copy output with amrep and log-op "old style"
 - output/andbox check size
 - increased timeout for glite
 - fixed bug in white and black list (CE and SE)To be used with the server:
 - user outputdir parameter
 - kill and submission by range
 - crab.ctg output_dir usage
- Updates Tags:**
 - Boss: 4_3_8-643_1
 - DBS: 1_0_7
- Warnings:**
 - To submit CRAB jobs from slc3 please use CRAB_1_5_3
 - CRAB_2_0_0 works only with the UI 3.1.0-1 version
At CERN you have to be logged in lxplus and source environment files following this order:
1. source /afs/cern.ch/cms/LCG/LCG-2/Utils/sl_env.csh
2. eval `source`/runtime-csh from your scram area
3. source /afs/cern.ch/cms/ccs/wm/scripts/Crab/crab.csh
 - For CRAB local installation:
the CRAB installation must be followed before installing LCG environment

Analysis Operation

- As analysis activity is increasing, we need someone to be responsible of it
- Manage resources and its allocation (servers, WMS, sharing between production/analysis, etc ...)
- Find resource for users (eg at T2) and notify users about that
- Check daily status of sites (on dashboard) and react if something is not working
- The issue has been raised to management, and discussion is ongoing



Conclusion

- CRAB (and the grid behind) is still much used;
- Need effort by us, sites, etc in order to improve the stability of the system: situation has much improved wrt the past;
- still problems, we are putting our best effort on user support;
- crab standalone is (and will be) supported for daily work: but if you submit $\mathcal{O}(100)$ jobs, **consider using the server**;
- **Do use copy to remote SE for large output**
- Analysis Operation task is beeing discussed to improve resource usage and support for final users



Continous
Meeting
Society

