TECHNICAL NOTES

Time-critical job submission under ECaccess

User Support

Forecasts Department

February 2015

1 Introduction

This document describes the servirce that allows users to automatically submit jobs to be run when certain points in the daily ECMWF operational forecast suites have been reached. The main purpose is to ensure that certain data is available before e.g. submitting a MARS request. This facility is running using the ECaccess environment. It is available either through the Web interface of ECaccess or with the ECaccess Web Toolkit, available on ecgate or installed locally. Note that you will need to install at least version 3.1 of the tools. This service is monitored by the operators at ECMWF.

2 Enhanced ECaccess batch service

In 2007, we have extended the existing batch service under ECaccess to provide a new facility allowing registered users to run jobs when ECMWF's operational activity has reached certain points.

2.1 Events

A database of events, also known as notifications, has been added to ECaccess. Such events can be added, deleted or modified by individual users. An event will have a name and a description. In the context of this new service for registered users, the events defined in ECaccess correspond to the points in the operational suite when some data or products are available. For example, we have defined an event called 'an12h00' with the description 'At this stage, the analysis cycle for 12:00UTC is complete.' As such, this event will not yet have any link with the ECMWF operational activity, apart maybe by its name and description. From the user viewpoint, when submitting a batch job through ECaccess, he or she will be able to subscribe (repeatedly) this job to the events available to him or her. And from the ECMWF operational viewpoint, we will send notifications to these predefined events. When ECaccess receives a notification for an event, it will release the user jobs which have subscribed to the event and submit them to the batch service on the system selected, e.g. to SLURM on ecgate or PBS on the HPCs. Finally, a little time after a notification to an event has been issued and the jobs subscribing to the event have been submitted, a 'sweeper' daemon within ECaccess will prepare a new version of the users' jobs subscribing to the event, ready for submission at the next notification of the event.

2.2 User interface

The jobs to be attached to the ECMWF operational suite will have to be submitted through ECaccess. Batch job submission is available from the ECaccess web interface or through the ECaccess Web Toolkit. We will first look at the Web interface, then at the Web Toolkit.

2.2.1 Web interface

When logged in on the web interface for ECaccess, e.g. on http://ecaccess.ecmwf.int/ or on your local gateway, you have the possibility to submit a new job from the left margin. The upper part of the submission page is shown in figure 1.
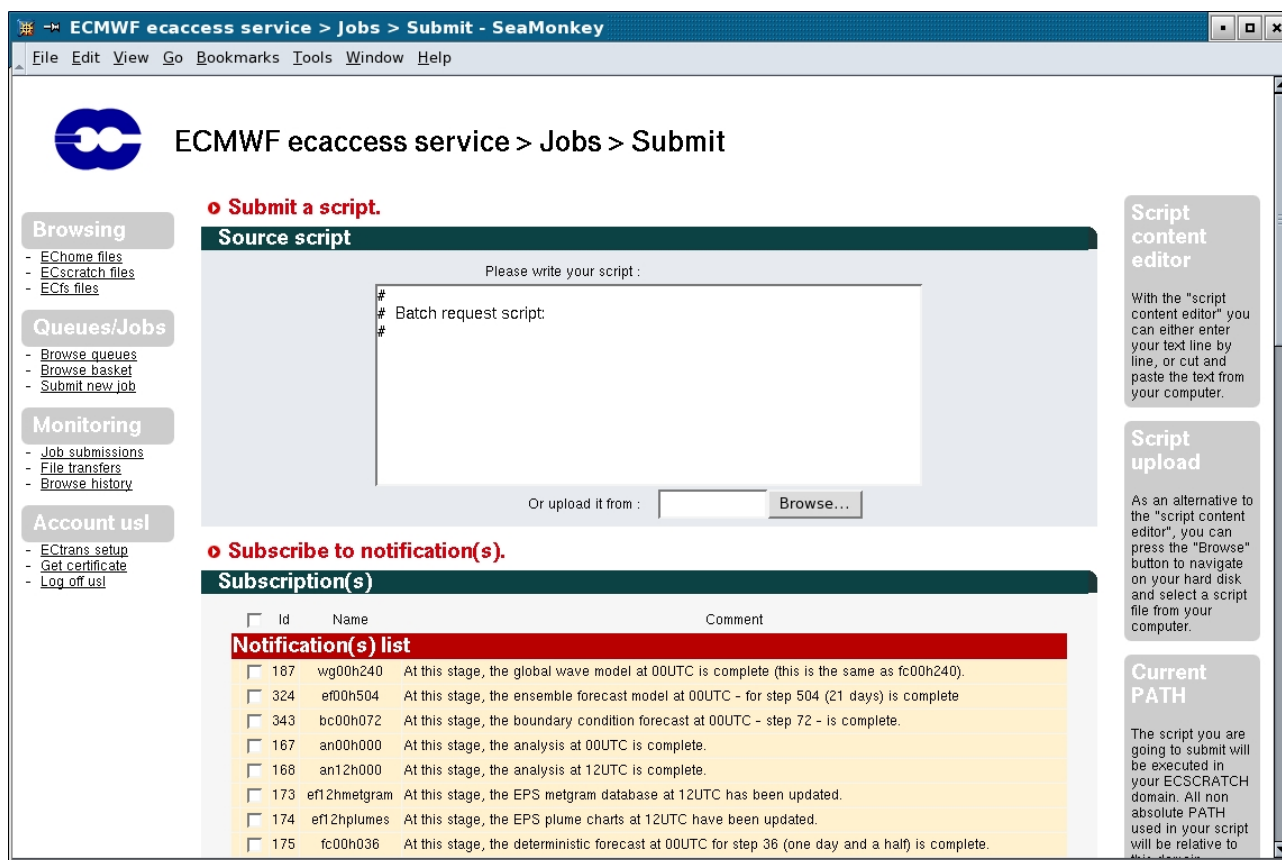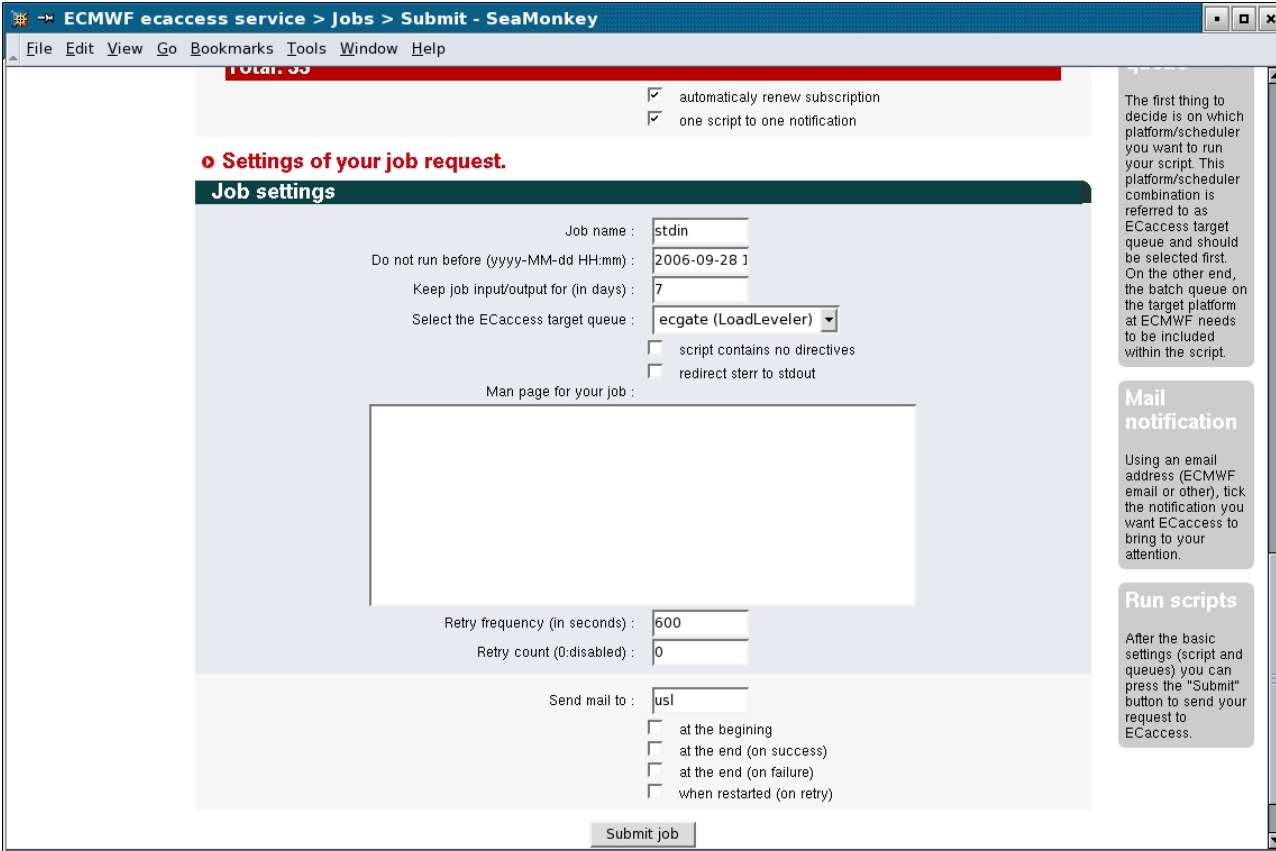

Figure 1: Job submission - Upper part

The part to include the job script has not changed. You can either type in the script, copy and paste it from another window or upload it from a local file. One important addition to make to your jobs is to add the 'set -e' command or alternatively to manage the errors in your jobs and exit accordingly - see section 3.1 for more details.

The lower part of the job submission window (see figure 2)- called subscription - allows you to attach your job to the different events available to you. Simply tick the boxes corresponding to the event(s) when you want to run your job.

By default, the jobs you attach to an event will be run automatically every time a notification is sent to the given event. If you want a job only to be run at the first next notification of an event, you can untick the box labelled 'automatically renew subscription'.

Under point 'Settings of your job request' you can customise various options for your job. The important options for this service are described below:

1. "Keep job input/output for:" - ECaccess will create one new job for subsequent notifications of an event. E.g. if you have subscribed one job to the event 'an00h000' one new ECaccess job will automatically be created and submitted every day when the ECMWF analysis for 00UTC is complete. The jobs used

Figure 2: Job submission - lower part

for the previous days will be kept in the ECaccess spool; they will be removed after the number of days specified in this field.

2. "Man page for your job:" - The ECMWF operators have utilities to monitor your jobs subscribing to any events of the ECMWF operational activity. In this page, you can give some instructions to the operators on what to do if the job fails. Operators could rerun your job (see next point) or possibly inform someone about the problem. If no instructions are given, our operators will not take any specific action on your jobs.

3. "Retry frequency and Retry count:" - With these options, you can request your job to be rerun automatically (without the intervention of the ECMWF operators) a certain number of times if it fails.

4. "One script to one notification:" - If you have ticked several events for your job, by default you will have one job running for each individual event. If you want only one job to run when all the notifications to the events have been received, you can untick the box labelled 'one script to one notification'. This option could be used if, for example, you want to extract in the same job, some epsgrams products and raw EPS data. Your single job will seem to be linked to the two events and will be submitted when a notification to the two events has been sent. Be careful, though, to submit such a job at the correct moment, before the two events occur in the Operational suite, not in between.

When you have given all the necessary information about your job, you can submit it. Your job will be taken by ECaccess and put in standby mode - status STDBY. You can monitor your jobs by selecting the link 'Job

submission' under topic 'Monitor' in the left margin. The monitoring page is shown in figure 3.
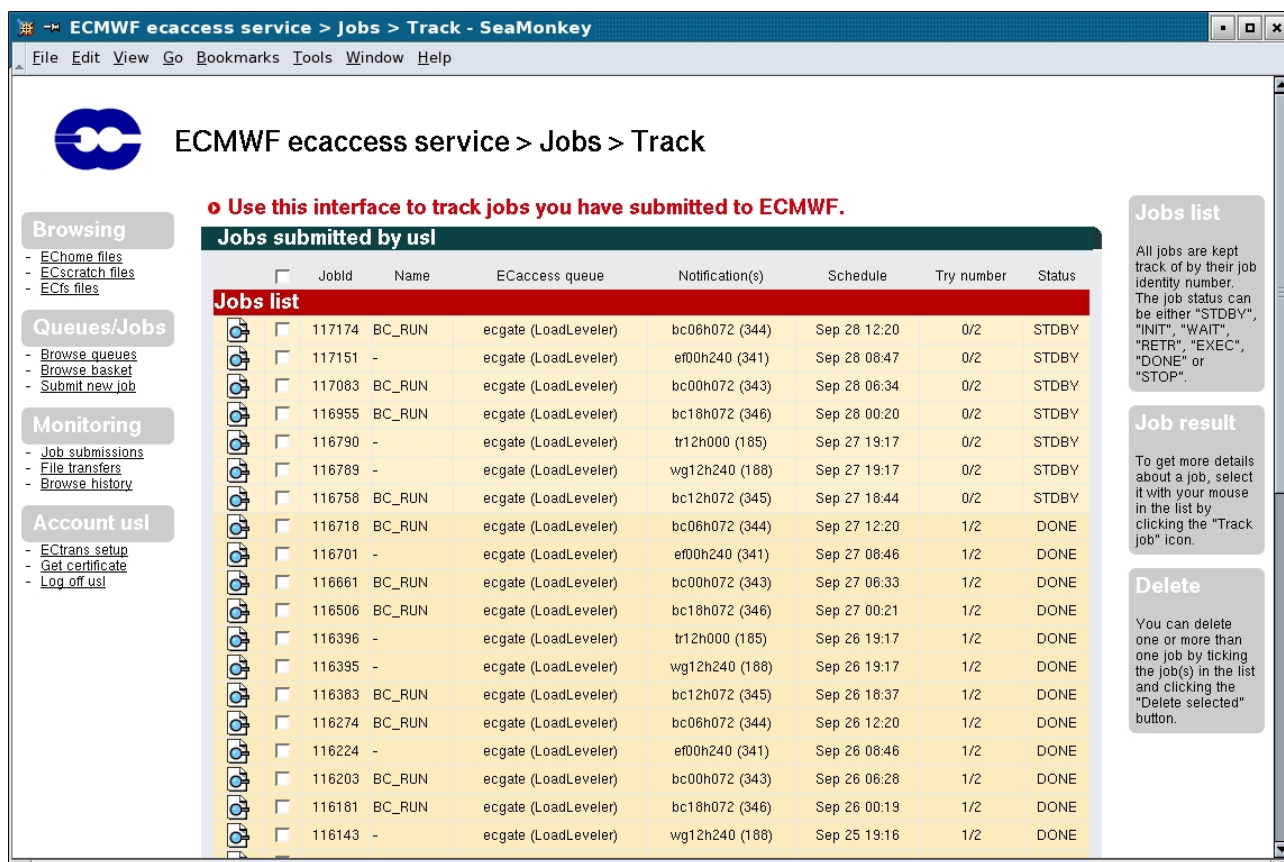
Figure 3: Job monitoring

In this page you will see all your jobs submitted through ECaccess, both those with subscriptions to some events of the operational suite as well as other jobs. You will also see the jobs due for later schedule, as well as those which have already run for the previous notification of some events. Please note that the name of the job is also shown, when available. You can delete jobs from this page. See section 3.4 for more details.

2.2.2   ECaccess Web Toolkit

The same functionality as described above for the web is available through the ECaccess Web Toolkit. These mre available on ecgate and the HPC systems at ECMWF. They may also available on your local systems. Please refer to the ECaccess documentation for further details on the ECaccess Web Toolkit:

http://software.ecmwf.int/wiki/display/ECAC/Web+Toolkit+-+The+full+featured+client

ecaccess-event-list

The command ecaccess-event-list has been added to the ECaccess Web Toolkit to allow you to list the events available to you.

```
ecgate{~uid/}: 1 -> ecaccess-event-list
167   an00h000    At this stage, the analysis at 00UTC is complete.
201   an06h000    At this stage, the deterministic analysis at 06UTC is complete.
168   an12h000    At this stage, the analysis at 12UTC is complete.
202   an18h000    At this stage, the deterministic analysis at 18UTC is complete.
221   bc00h012    At this stage, the boundary condition forecast at 00UTC - step 12 - i
343   bc00h072    At this stage, the boundary condition forecast at 00UTC - step 72 - i
222   bc06h012    At this stage, the boundary condition forecast at 06UTC - step 12 - i
344   bc06h072    At this stage, the boundary condition forecast at 06UTC - step 72 - i
223   bc12h012    At this stage, the boundary condition forecast at 12UTC - step 12 - i
...

ecgate{~uid/}: 2 -> ecaccess-event-list an00h000
     Event-id: 167
         Name: an00h000
       Public: yes
        Owner: <owner>
      Comment: At this stage, the analysis at 00UTC is complete.
```

Note that either the event number or name can be used with the ECaccess Web Toolkit.

ecaccess-job-submit

When you have found the notification to which you want to attach your job, you can use ecaccess-job-submit
to submit your job. This command has been enhanced to allow you to attach your jobs to some event of the
ECMWF operational suite. The relevant options for this service are "-eventIds", "-noRenew" and "-manPage".

```
ecgate{~uid/}:3 -> ecaccess-job-submit -help

 Usage:
    ecaccess-job-submit -version|-help|-manual

    ecaccess-job-submit [-debug] [-distant] [-bufsize *length*]
    [-scheduledDate *date*] [-noDirectives] [-gateway *name*] [-remote
    *location*] [-transferOutput] [-transferError] [-transferInput] [-keep]
    [-eventIds *list*] [-sterr2Stdout] [-noRenew] [-mailTo *email*]
    [-onStart] [-onSuccess] [-onFailure] [-onRetry] [-jobName *name*]
    [-manPage *content*] [-lifeTime *days*] [-retryCount *number*]
    [-retryFrequency *frequency*] [-queueName *name*] *source*

Arguments:

...
    -ni, -eventIds *list*
          Allow giving a *list* of event-identifiers to subscribe to with
          the Job. The list should be separated by ';' or ','. Only one
          job will be launched when all the events in the list have been
          reached. To submit the same job to multiple events, one will
          need to submit the job to each event separately.

    -ro, -noRenew
          The job subscriptions to events will not be renewed.

    -mp, -manPage *content*
```

```
                   Allow giving the man page *content* which will be displayed to
                   the ECMWF operators in case of problems with your Job (e.g. what
                   to do or who to contact).

      ...
```

The option '**-eventIds**' allows you to give the event(s) to which your job should be attached.

By default, the jobs you attach to an event will be run automatically every time a notification is sent to the given event. If you want a job to be run only at the first next notification of an event, use the option "**-noRenew**".

You can give some short instructions to the operators with the option '**-manPage**'.

Note that there is no equivalent option with the 'ecaccess-job-submit' command to the tick box "one script to one notification" of the web interface. If you want to submit one job to multiple events, you will have to run multipe ecaccess-job-submit commands.

A sample job submission attached to the event an00h000 could look like follows:

```
ecgate{~uid/}:4 -> ecaccess-job-submit -queueName ecgate -eventIds
an00h000 -mp "nothing to be done" -retryCount 1 -retryFrequency 300 job.cmd
6746919
ecgate{~uid/}:5 -> ecaccess-job-list 6746919
      Job-Id: 6746919
    Job Name: job.cmd
       Queue: ecgate
        Host: ecgb.ecmwf.int
    Schedule: Feb 04 08:25
  Expiration: Feb 11 08:25
   Try Count: 0/2
      Status: STDBY
    Event-Ids: an00h000 (167)
```

Note the status 'STDBY' for the job. This job will remain in standby mode up until the ECMWF operational activity will have produced the analysis for the 00Z run of the HRES forecast. Note that the jobs submitted through ECaccess in standby mode will only be visible through ECaccess. They will not be visible using the usual batch service commands, e.g. squeue for Slurm on ecgate.

3   Notes on new service

3.1   Job status

One advantage of the new service for time-critical jobs submissions under ECaccess is that the ECMWF operators are monitoring your jobs submitted via this system. Also, you can request your jobs to be rerun automatically on failure. However, ECaccess will only be able to show the correct status of your job or possibly rerun your job if it has correctly been notified about the exact status of your job by the batch service, e.g. Slurm on ecgate. It is therefore your responsability to notify correctly the batch service about errors occuring in your job. By default, an error in your job will not be reported to the batch service; the execution of your job will continue and it will finish as if it had completed successfully. One stratight way to stop the

execution of your job as soon as there is an error is to use the `set -e` command, in the ksh or bash. With this command, your job will stop and exit abnormally as soon as an error occurs. If you use the csh, it will be slightly more difficult to notify the batch service about an error. If you want to use the `-e` option, you will have to run your job like:

```
#@ shell = /bin/csh
#@ queue
csh -ex toto.csh
if ( $status != "0" ) exit 1
```

If you want a finer control over the errors, you can include some specific tests in your jobs and, for those important tests, exit with a non zero return code, e.g.

```
mars request
if [[ $? -ne 0 ]]; then
 echo mars request failed
 exit 1
fi
```

If you request that your job is restarted after a failure, you will have to make sure that it can be rerun. For example a job doing, e.g.

```
set -e
mkdir $SCRATCH/data
```

cannot be rerun, as the directory $SCRATCH/data will already have been created during the first run. There are different ways to avoid such problems. One way would be to switch `set -e` off in some parts of your script, e.g.

```
set +e
mkdir $SCRATCH/data
set -e
```

Another option is to use the conditional excution statement, e.g.

```
set -e
mkdir $SCRATCH/data || true
```

Another way to avoid this particular problem is to work in $SCRATCHDIR.

Please note that you can submit your job to ECaccess without setting up what is suggested above. Your jobs will run normally but, without this job control, the ECMWF operators will not notice any errors with your jobs and ECaccess will fail to resubmit your jobs, even if you requested some retries.

3.2   Monitoring by the operators

With the new system, the operators now have a specific interface to monitor user jobs submitted through this system. This will allow them to identify various problems, e.g. a general problem with one system at ECMWF

or a failure to send a notification to an event, leaving all user jobs waiting to be run. When such problems occur, the operators will try and take corrective action. In case several jobs have failed, apparently linked to a general problem, our operators will be able to restart these jobs, after the problem has been fixed. Our operators will also usually have access to the users job and job output files (see section **??** below), as well as to some instructions you may have given for them, when submiting the job through ECacces. If no instructions are given, our operators will normally ignore any specific failure of your jobs. Note that our operators will be unable to correct something in your job or under your account; they cannot edit your jobs. Rather then asking the operators to rerun the jobs or to notify someone when a failure occurs, we recommend you to use the automatic job resubmission facility on failure or the email notification option, available with ECaccess or with the batch service. Instructions to the operators should be clear and simple.

3.3   Environmental variables

Before submitting the job, the following environment variables are set by ECaccess, are passed to your job and can itherefore be used within your job:

MSJ_BASETIME forecast base time

MSJ_STEP forecast time step

MSJ_YEAR year of the run

MSJ_MONTH month of the run

MSJ_DAY day of the run

MSJ_EXPVER version number of data archived in MARS (if relevant)

MSJ_EVENT event name

These environmental variables will help you to access the operational data, e.g. to build the correct date to extract the data from MARS.

3.4   Changes in job or suppression of jobs

If you have to make some changes to any of your ECaccess Time Critical jobs, you will have to cancel the existing job in standby mode and submit the new version of the job. The job name shown with 'ecaccess-job-list' or through the web interface should help you in identifying the correct job to delete. Similarly, to remove a job from the system, you will have to remove the job in standby mode.

4   Job examples

One Time Critical batch job example is available under:

http://www.ecmwf.int/services/computing/job_examples/ecgate/.

5 Help and Support

For initial help on implementing your jobs into the new system, please contact your usual contact point in User Support. For problems with the 'operational' runs of your jobs, we recommend you to leave some instructions to our operators in the man pages and we would preferably ask you to contact User Support during working hours. If you need to contact ECMWF outside working hours, for urgent issues, please contact the calldesk. When talking to calldesk or operators at ECMWF about one of your time-critical jobs, please do not forget to report the UID used and also the event your job is attached to. Our operators should then be able to identify your jobs. They will be able to tell you if the job has been submitted or not and if it failed. They will only be able to resubmit your job, not to edit or fix it.