

Wikiprint Book

Title: Envío de trabajos Matlab al cluster

Subject: TracMeteo - Enviamatlab

Version: 23

Date: 08/19/2022 10:07:54 PM

Table of Contents

Envío de trabajos Matlab al cluster	3
Configuración de correo para recepción de mensajes (job terminado, etc.)	3
Observaciones	4

Si envías un job con Matlab2009 sin usar el `enviamatlab` se puede consumir mas CPU de la asignada. `help maxNumCompThreads`

Envío de trabajos Matlab al cluster

El acceso al cluster por ssh es `mar.macc.unican.es` y los trabajos en Matlab hay que lanzarlos utilizando `'enviamatlab.sh'`, que se encarga de mandar el trabajo a la cola 'estadística', que es la cola por defecto para los trabajos de Matlab; esta cola puede ejecutar hasta 24 trabajos a la vez: nodos WN10, WN11 y WN12 (**nodes**), con 8 procesadores cada uno (**ppn**). El script `'enviamatlab.sh'` se ha añadido en un directorio del cluster (`/software/MachineIndependent/bin/`) accesible por todos los usuarios. Por tanto, si os logueais en mar y poneis `'which enviamatlab.sh'` (es decir, busca la ruta donde está `'enviamatlab.sh'`) todos deberíais ver:

```
[gutierjm@mar ~]$ which enviamatlab.sh
/software/MachineIndependent/bin/enviamatlab.sh
```

Por tanto, para ejecutar cualquier programa de Matlab basta que llaméis a este script desde el directorio donde tengáis el fichero ".m" que queráis ejecutar **SIN TENER UNA COPIA EN LOCAL**:

```
[gutierjm@mar ~]$ enviamatlab.sh script.m
```

De este modo, se generarán en la carpeta dos ficheros:

- **matlabqsub.20091123T105716**: este fichero lo genera el `enviamatlab` y será el que arranque matlab y ejecute el `script.m`. El final del nombre es la fecha de ejecución.
- **script.20091123T105716.log**: en este fichero se recoge la salida por pantalla de Matlab.

Al final de la ejecución del `script.m` se genera un nuevo fichero **script.o114037**, donde el número indica el identificador del job ejecutado.

Una vez ejecutamos la función nos aparecerá por pantalla una serie de mensajes, como se muestra en el ejemplo, en los que se muestra la información del trabajo enviado. En la última línea se muestra el identificador del trabajo (**JobId**, en el ejemplo **121176**) seguido del nombre de la máquina, `ce01.macc.unican.es` (esta máquina es el submitting machine, pero no es la que ejecuta el job; para monitorizar los jobs ver la sección de monitorización) . Este identificador nos servirá para informarnos del estado del trabajo.

```
[rodri@mar ~]$ ./enviamatlab.sh test.m
#!/bin/bash
#PBS -N test
#PBS -q macc
#PBS -l nodes=1:ppn=1
#PBS -l walltime=48:00:00
#PBS -m ea
#PBS -M rmanzanas@ifca.unican.es
#PBS -r n
export LD_ASSUME_KERNEL=2.6.18
export PATH=/software/Matlab_R2009a/bin:$PATH
matlab -logfile /vols/mar/users/rodri/ejemplo/logs/log_20091209181811.err -r "cd /vols/mar/users/rodri/ejemplo;test;" -nos
exit 0
121176.ce01.macc.unican.es
```

Por defecto, *enviaMatlab* envía los trabajos a un sólo procesador (ppn) con un límite de 750mb de memoria. Si el job consume más memoria, la cola lo interrumpe. Si se necesita correr trabajos que consuman más memoria hay una opción para reservar el doble de memoria:

```
[gutierjm@mar ~]$ enviamatlab.sh test.m -l mem=1500mb -l nodes=1:ppn=2
```

o cuatro veces más (hasta 3GB de memoria). Si necesitáis ejecutar jobs que requieran más memoria, hablar con Sixto o con Antonio.

```
[gutierjm@mar ~]$ enviamatlab.sh test.m -l mem=3000mb -l nodes=1:ppn=4
```

Configuración de correo para recepción de mensajes (job terminado, etc.)

NOTA: La primera (y sólo la primera) vez que queramos enviar un trabajo de Matlab al cluster debemos editar el fichero `.bashrc` de nuestro home de oceano, incluyendo al final del mismo las líneas:

```
export EMAIL=tuemail[at]unican.es
source /software/etc/bashrc
```

El fichero final debería ser tal que así (se ha cogido como ejemplo el `.bashrc` del home de Rodri):

```
export EMAIL="rmanzanas@ifca.unican.es"
source /software/etc/bashrc
```

Para ello, basta hacer:

```
[gutierjm@mar cluster]$ vi .bashrc
```

editar el fichero justo al final "G", incluir "i" las dos líneas anteriores. Luego presionar "escape" y escribir ":x" (comando para salir guardando) y presionar "enter".

El que no sepa, que me pregunte a mi o a Sixto, pero no tengáis copias del script 'enviamatlab' en local porque estaréis generando errores para el resto de usuarios.

Observaciones

En este apartado podéis escribir las observaciones, dificultades o aclaraciones que encontréis o creáis necesarias a la hora de usar el cluster.

- Como ya se ha indicado, para poder mandar jobs al cluster debemos tener las función `enviamatlab2009.sh` o `enviamatlab65.sh` (quien las necesite se las puede pedir a Rodri) en el mismo directorio que el script que a ejecutar. Una vez copiada/s en el directorio correspondiente y después de habernos asegurado de que está/n correctamente implementada/s, es decir, tiene nuestro nombre de usuario y las demás variables que nos identifican correctamente, podemos ejecutarla/s sin problema alguno siguiendo las instrucciones explicadas anteriormente.
- Es importante tener en cuenta que el Matlab en el que se arranca el proceso no tiene porque tener inicializada Meteolab y que, por lo tanto, debemos añadir al script una orden para que la inicialice. Bastaría con añadir al inicio del script la ruta de acceso al Meteolab que utilicemos (es recomendable tenerla siempre actualizada en nuestro home), ejecutar el programa `init` (que arranca la toolbox) y volver al directorio en el que queremos trabajar. De este modo podremos trabajar con todas las funciones de Meteolab.
- A la hora de marcar las rutas en Linux debemos usar los elementos propios de este sistema. En particular debemos usar la '/' en lugar de la '\' usada habitualmente en Windows. Es importante crear los scripts de modo que puedan ser entendidos en Linux. Para consultar dudas podéis preguntar a Chus, Antonio, etc.
- Las funciones de Meteolab ya están escritas de modo que puedan ejecutarse en Linux y Windows indistintamente (en principio).
- Siempre que tengamos algún problema con la sintaxis de un comando o no sepamos exactamente qué es lo que hace podemos acudir a su manual. Por ejemplo, en el caso del comando `qstat` (monitorización de trabajos) bastaría con teclear `man qstat`.