DCRAB

DIPC Control and Report Application in Bash

User Documentation

Document Revision 1.0.1
29 May 2018

DCRAB v2.0

Abstract

DCRAB is a tool to monitor resource utilization in HPC environments. It works side-by-side with the job scheduler to collect runtime information about the processes generated in the compute nodes.

Excluding a few cases, the data DCRAB collects is taken from the processes which the job has started, not from the entire node. The tool is able to collect the information listed below:

- CPU used

- Memory usage

- Infiniband statistics (of the entire node)

- Processes IO statistics

- NFS usage (of the entire node)

- Disk IO statistics

Contents

Installation

DCRAB has no installation at all, you have to take the lastest version tarball available in Github (currently v2.0) and expand it in a convenient location in your system:

```
tar xzvf ./DCRAB-2.0.tar.gz
```

This will create a directory called `DCRAB-2.0` with some subdirectories:

- `/auxFiles` which is used to store the auxiliary files useful for the internal operation mode of the tool as described in 5.1 section.

- `/config` folder which only contains a folder to store the current version of the tool.

- `/docs` which contains this documentation and the `.tex` used to generate it.

- `/examples` is used to stored DCRAB's generated report examples of each version since it was created.

- `/src` folder contains the source files of the tool. Inside you will find `dcrab` script, which is the script that launches DCRAB and a folder called `/script` that contains other scripts used by the tool to work.

Also you will find the `readme.md` file and the GNU GPL 3.0 license file `COPYING`.

Using DCRAB

Data Collection

DCRAB has two operation modes: normal report operation, which may be the mode used by non-admin users, and internal report operation, which is focused to sysadmins. The second operation mode is going to be explained in 5.1 section, so firstly we will introduce here the normal report operation which will be the commonly used one.

The tool is straightforward to use, you have to add the following lines into your submission script:

```
export DCRAB_PATH=/PATH_TO_DCRAB/
export PATH=$PATH:$DCRAB_PATH/src
dcrab start

###################################
#     BLOCK OF CODE TO MONITOR    #
###################################

dcrab finish
```

Note that the variable `DCRAB_PATH` must be declared to run DCRAB. You can also make a module with lmod to load the paths in just one line like "module load DCRAB".

DCRAB will start a process in each compute node where the script runs, and will monitor the processes started by it. The script will run as normal, and meanwhile DCRAB will generate a directory report called `dcrab_report_<jobid>` where `jobid` is the job number assigned by the scheduler. This reporting directory is generated in the same folder where the job was submitted.

Inside this reporting directory, DCRAB will create the reporting file called `dcrab_report.html`, which will be named as "the reporting file" is this document from now on. Generate this file is the main purpose of this monitorization and there you will find statistics and plots to analyze visually the information collected. This report is continuous change (every 10 seconds by default) because the information collected is stored at the time it is taken. So, you could open it with a browser, at the start of the job's execution, and refresh to see what is going on with the job.

One of the goals was to provide a single file to visualize all the data monitored. So, the reporting file is completely modular and you can copy or move it to any location because every plot and image is embedded in the report.

To conclude, say that inside the `dcrab_report_<jobid>` reporting folder you will find some subdirectories, which normally are not relevant at all for the user because they are generated to guarantee the correct behaviour of the tool. One will be `/data`, which contains for each computing node the files in charge of the management of the processes asociated with the job. Another folder called `/auxFile` stores the files used for the comunication between compute nodes, and `/log`, which contains the output generated by DCRAB process in each node and DCRAB's main process output (all used for troubleshooting).

Execution Customization

In DCRAB version 2.0 is not included yet any configuration file to customize the execution, but it is included in the roadmap. However, there are some variables that could be changed inside the code until the configuration file is released:

- DCRAB_COLLECT_TIME: this variable, inside `src/scripts/dcrab_config.sh` in `dcrab_init_variables()` function, configures the time between each data collection. By default it is set to 10 seconds.

- DCRAB_NFS_MOUNT_PATH: this variable configures the path of the NFS filesystem to be monitored. It is inside `src/scripts/dcrab_node_monitoring_functions.sh` in `dcrab_node_monitor_init_variables()` function. By defaut its value is `/scratch`.

Crashed Jobs

Update the reporting file continuosly imposed some troubles but it was one of our first goals. There is no much problem for the tool when the job crashes because the report has been generating until that moment, so it brings you such a great trace of what has been occurred and it may contain relevant information about the crash.

The tool exits with code '0' if everything have run normally and exits with code '1' in case of error. To see more information about the crash you can view the logs in /log directory.

In the cases where the reporting directory has been deleted a file called `DCRAB_ERROR_<computeNodeHostname>_<jobid>` will be generated, per each compute node involved in the calculation, in the same folder as the reporting folder was (where the job was submitted). The `computeNodeHostname` refers to the hostname of the compute node that throws the error and `jobid` is the job number assigned by the scheduler to the job.

Requirements

At this version the tool have a few limitations or requirements for its use:

- Support multinode statistics but only for MPI jobs. As is explained in 3 section the execution of the tool differs between the main node and the rest of the nodes. Other nodes that are no the main node will be waiting until a process related to the job is started. In this node may be more processes of the same user in execution but, which of these processes are related to the job that the tool is monitoring? Here is the problem. To know this we need something to compare, something that make us know that the the processes in the main node and the processes in the other nodes are connected. With MPI jobs the command that start the processes have an argument called `control-port` which one can use to make this comparison and associate the processes of differents nodes in to the same job execution.

- Only jobs under PBS scheduler are supported. To be able to monitor correctly the job, DCRAB needs explicity a couple of sentences in the batch script:

  - The option `#PBS -l mem=` must be written explicitly in the submission script for the memory data plot and the memory must be in GB.

  - The option `#PBS -l nodes=x:ppn=y` must be written explicitly in the submission script to calculate elapsed and remaining times.

Chart examples

Below you can find some examples of the charts the tool will generate for the report:
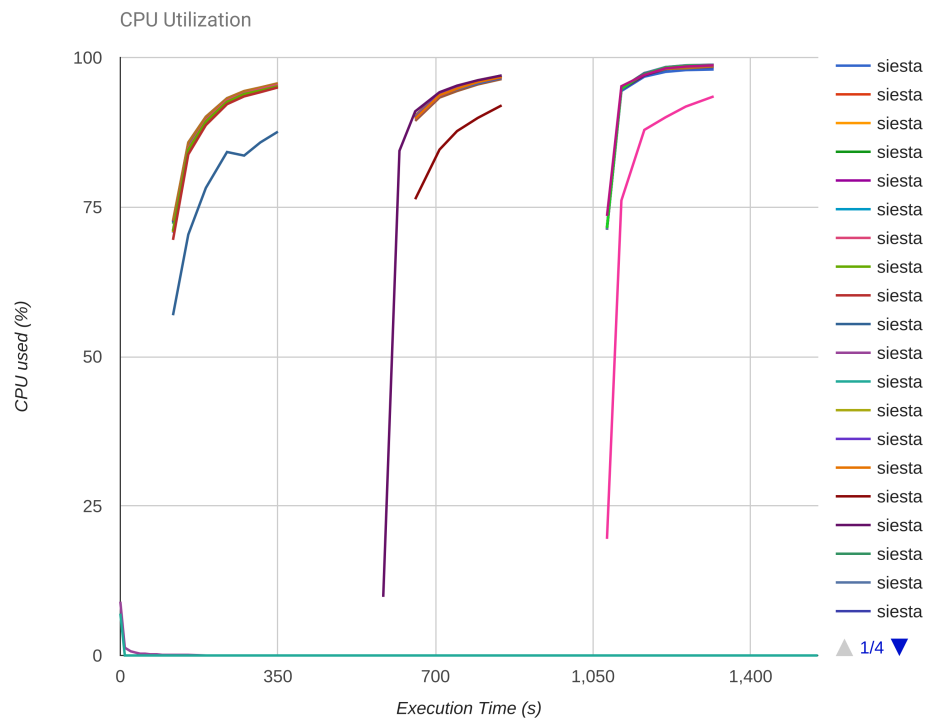
CPU Utilization

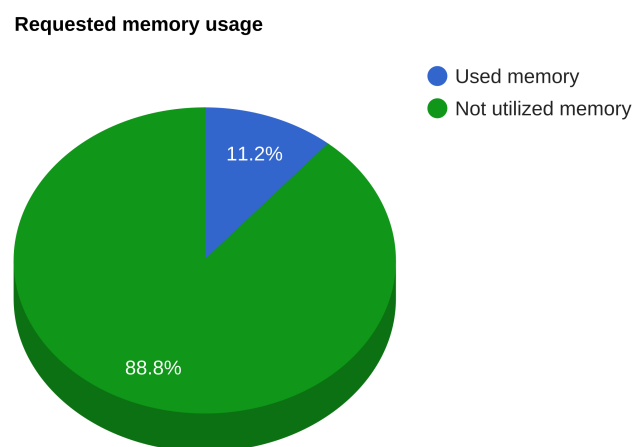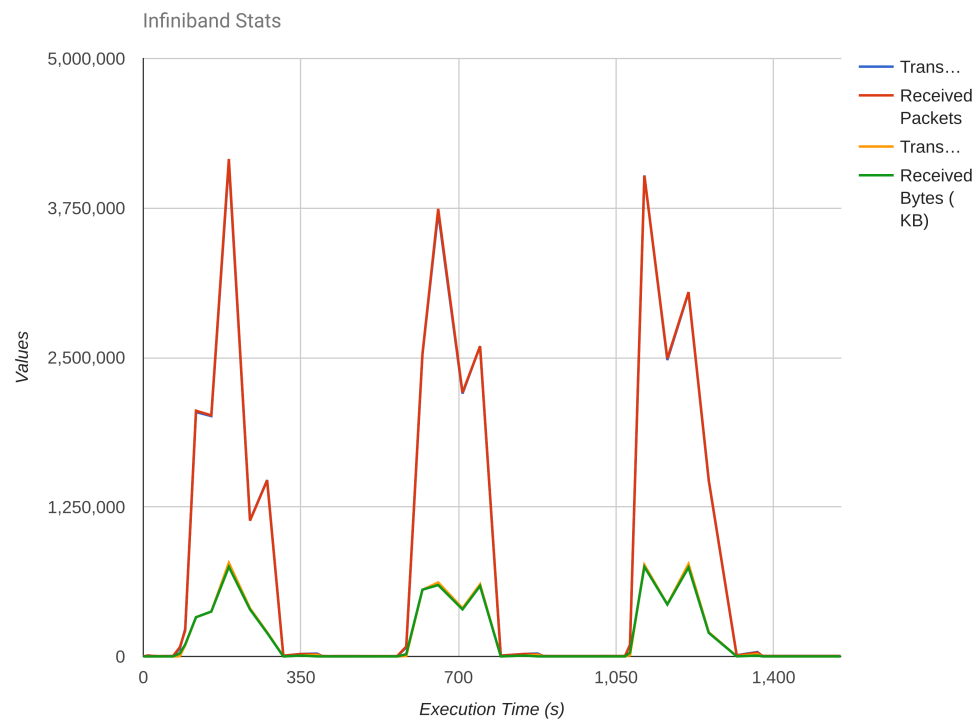

Figure 2.1: CPU usage

Requested memory usage

Figure 2.2: Memory usage

Infiniband Stats

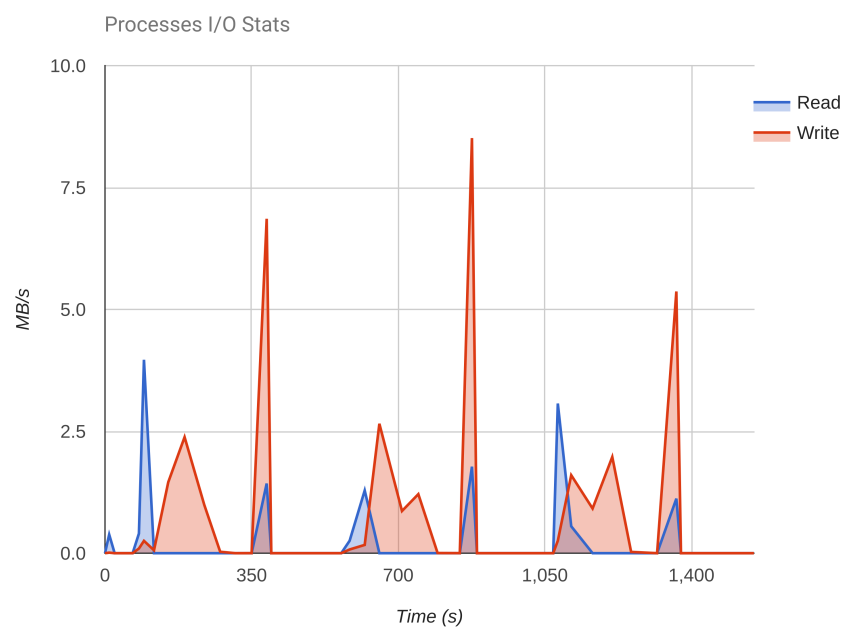Figure 2.3: Infiniband network usage by the node

Processes I/O Stats

Figure 2.4: I/O against the filesystems

6

Design and Implementation

The main idea of DCRAB was to create a tool easy to use for the users. Notwithstanding that DCRAB is upgrading an HTML file continuosly, the runtime penalty is superfluous.

The continue upgrading of the report conditioned and formed DCRAB's structure. This upgrade is a critical section of the project because all the compute nodes need to write there at the same time, so there is a race condition we had to solve. In early versions we resolved the problem with a lock, to make the writing as an atomic operation, but there was a problem with the delay of parallel filesystems when the reporting file become bigger. The error was when a compute node wrote in the report file. In those writings some data was lost because there is a little delay until the file is upgraded or refreshed with the new information for other nodes. The unique way to solve the problem was to serialize those writings (using a mark in the first line of the report), to be sure one node wrote after the previous node had written.

In its current implementation the tool makes a ssh from the master node to all other nodes. This connection starts a process in background that monitors the processes started by the job and collects information about them (every 10 seconds by default). The tool take advantage of the scheduler to know which processes are related to that job.

Statistics Collected

DCRAB collects different statistics and information which may be usefull for the user. Here is the complete list of these statistics:

- **CPU used**. The application reports the CPU usage for all the processes started by the job. This data is collected with `ps` command so it is a snapshot of the process at a concrete time. To have only a view of the main processes in the chart there is a threshold value defined to avoid trivial processes of the user. This information is very useful for applications that use OpenMP.

- **Memory usage**. This value is obtained with the processes' `/proc/<pid>/status` file where `pid` is the PID of the process. The tool collect information about Virtual Memory (displayed as VmSize), Resident Memory and Max Resident Memory (displayed as VmRSS). Also is displayed the memory requested by the user for that job in the scheduler and the total memory available in the node.

  If there are more than one node involved in the execution a bigger pie chart is generated to display the amount of memory used for the scheduler (with one node calculation this value is the same as the usage in that node so there is no reason to generate this plot). With this information the users could revise the amount of memory requested into a more real value one to not waste resources, which may also help schedulers' algorithms such as `Backfill` of Moab.

- **Infiniband statistics (of the entire node)**. There is counted the number of packets and MB of data transfered and received over Infiniband in the entire node (this values can not be collected per processes). This information can be used to improve or revise certain parts of the code, reducing the amount of data transferred over the network, comparing this code section with high network activity levels. This values are taken from the counters available in `/sys/class/infiniband/mlx5_0/ports/1/counters`.

- **Processes IO statistics**. The I/O (Input/Output) made by the processes (regardless of the type of the filesystem). This information can be useful to see if a process is writing more than expected which could be a bottleneck in the program. This data is collected from `/proc/<pid>/io` file where `pid` is the PID of the process.

- **NFS usage (of the entire node)**. The I/O (Input/Output) made by the processes on the NFS filesystem determined by `DCRAB_NFS_MOUNT_PATH` variable. This value, as the Infiniband value, can not be measured by process and must be a collected from the entire node statistics. The information is collected from `mountstats` command.

- **Disk IO statistics (of the entire node)**. The I/O (Input/Output) made by the processes on the local disks. This data is obtained from `/proc/diskstats` file so this value is from to the entire node and not only of the current job.

Code Structure

This are the scripts that compose DCRAB. Here does not appear the internal report operation scripts, which are dcrab_internal_report_generation.sh and dcrab_internal_report_functions.sh, because they do nothing in the normal operation mode and will be explained in 5 chapter.
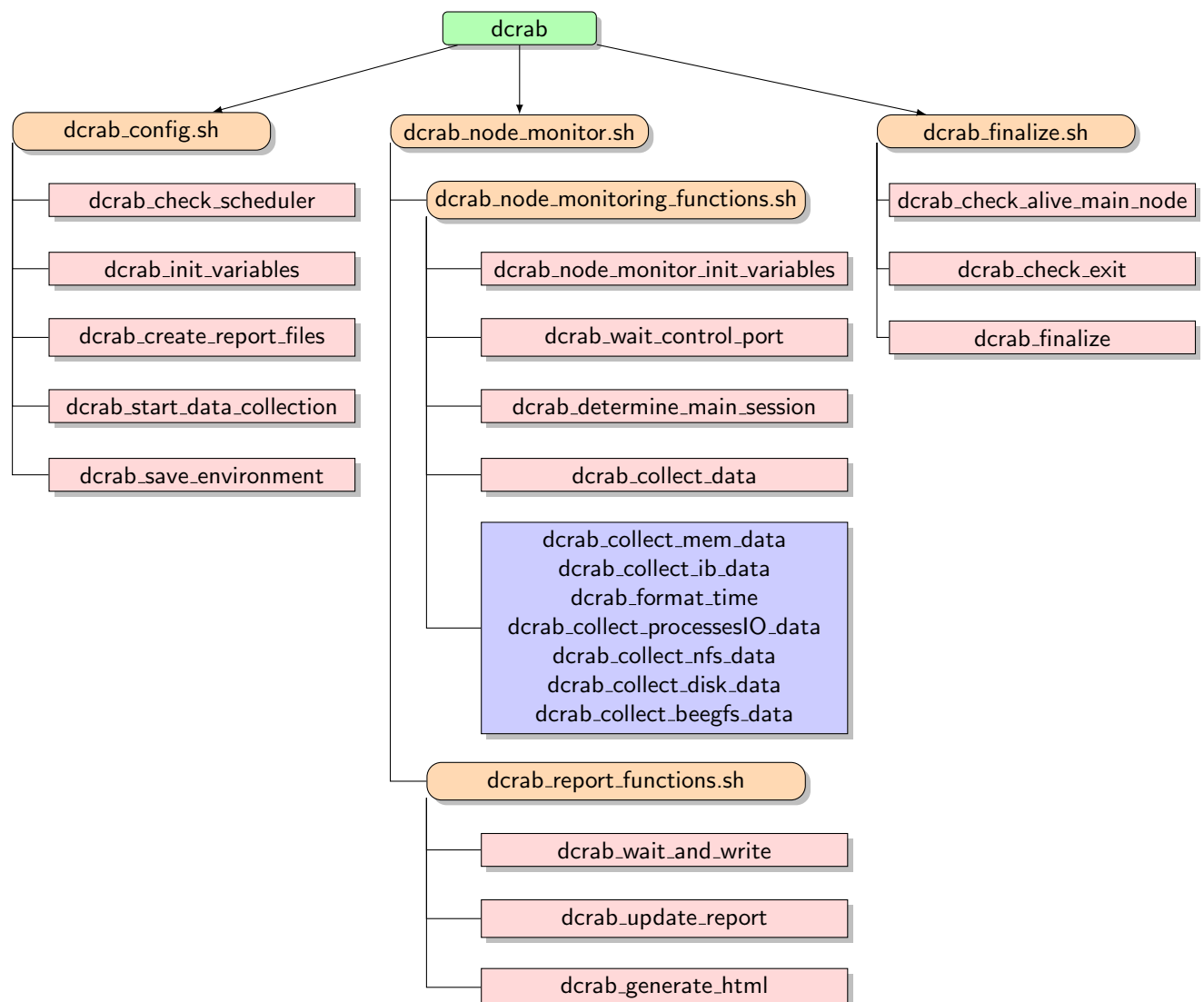

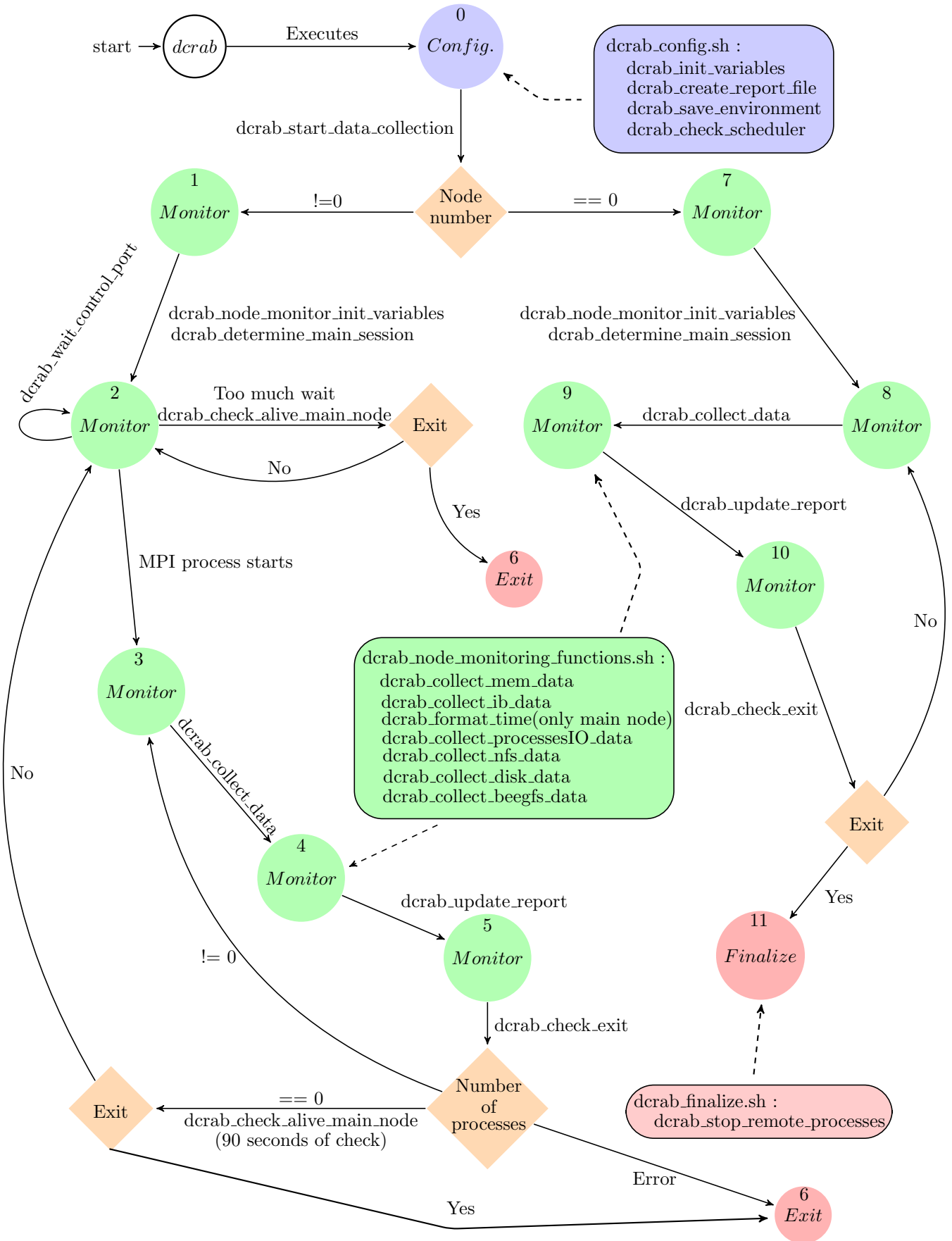Figure 3.1: DCRAB's code structure

Figure 3.2: DCRAB's state machine

Above, in the figure 3.2, is presented a bit modificated state machine which presents the states where DCRAB passed through. Each node represents the main steps made by the tool and each edge an operation to change between states. There are boxes attached to some states to point the functions executed on them.

- The first node, which is named as `dcrab`, represents the script executed with `dcrab start` command.

- The second node, named as `Config.` and with the number 0, represents the configuration made by the tool where different functions are executed (attached box in blue color). It starts in each node the data collection but here the execution differs between the main node (where the node number is equal 0, represented as `==0`) and the other nodes that are not the main node (where node number is not equal 0, `!=0` way).

- The sixth nodes were made to represent the stop of DCRAB an will be ommited in below explanations.

- Left part, no main node execution, when the node number is non-zero:
  - The first of the five states that compose the monitoring, which is marked as 1, is the first vertex to analize. These `Monitor` nodes represent DCRAB processes started in background in each compute nodes. The edge with the next vertex represents the initialization of needed variables to make correctly the data collection and creates the minimal environment to work. Furthermore, it will find the session of the processes created by the job (the tool takes advantage of the scheduler used to identify it).

  - This second `Monitor` node, marked with the number 2, is where the tool is ready to collect the data. This vertex is the first that represents tool's execution because the steps done before were just to intialize data structures.
    The first step to explain here is the loop in its left side, which executes the function `dcrab_wait_control_port()`, and represents the wait that the node must done until a MPI job is started and the execution passes to the next vertex
    On the right part there is drawn a loop with a condition: after a long wait for a MPI job to be started the tool checks if the main node is still executing the job or not (which takes normally about 90 seconds to ensure its state). If it is not still executing the job the tool will stop the execution but in the other case, where the condition is satisfied, the tool will reset its counters and will wait again for the MPI job. This steps are made because a user could insert other no MPI launcher commands inside the 'start' and 'finish' statements of DCRAB that may last for a long time, as huge copies with `cp` for example.

  - The next `Monitor` vertex with number 3 represents the state where the tool knows that a MPI process has been started. It has obtained the session of that MPI processes in the previous step so now it will collect the data related to these processes and move to the next step.

  - The fourth state represents the data collection. Here all the modules functions are executed and the data is extracted in the way this functions determine. Once this step is done the report is updated and moves down to the next vertex.

  - This last node, the fifth one, represents the state where the tool has collected and has written the data into the report. Here a few checks are made to ensure the correct behaviour of the tool.
    An important comprobation is made to count the number of MPI processes related to the job that are in execution. If an error occurred in that comprobation the tool will stop.
    The edge marked as `!=0` is satisfied if there are still MPI processes in execution. It goes back to the third state to make the loop again after a certain sleep (controlled by `DCRAB_COLLECT_TIME` variable).

The last case, marked as `==0`, means that all MPI processes have been finished. Here will be done another time the comprobation to check if the main node is still executing the job (normally this takes like 90 seconds) with `dcrab_check_alive_main_node()` function. In that comprobation if the main node has already finished the tool will stop but in the other case, the tool will go back to the second step, and will wait until a new MPI job comes.

- Right part, main node execution, when the node number is equal zero:

    - The first vertex, marked with the number 7, is drawn to represent the same steps as the vertex number one, but in this case another data structures will be initialized. This is because the main node must make some other operations in the moment of upgrading the report.

    - The second node, marked with the 8 number, represents the state where all the initializations have been made. It goes to the next node making the data collection.

    - The ninth node is were all the data is collected. As the main node, in this step it makes some extra operations inside the modules' functions and also has the resposability of updating the elapsed time inside the report, which is made by the `dcrab_format_time()` function. Finally, it goes towards the next vertex making the report updating.

    - The node number 10 is drawn to represent the state where the main node has made all its changes in the report. Here it checks if the tool must be stopped or not. If the conditions are satisfied the tool will go to the last node `Finalize`, if not, another loop will be done going back to the eighth state.

    - The last node, called `Finalize`, is reached when some conditions are satisfied as mentioned. Here DCRAB finishes and is going to kill all DCRAB instances inside the compute nodes (those processes launched in background).

Creating New Module

The process to create a new module is straightforward. It could be divided into two main steps, the first one the step where you must make the module to collect wanted data, and the second one modify the html report to display the information collected. The second step is not a must so you can decide not to do and store the collected data into a file. You need to add a few lines in `dcrab_node_monitoring_functions.sh` for the first part, and other lines into `dcrab_report_functions.sh` script to configure the display.

Below are presented the steps to insert a new module and the best way to name variables and functions in charge of it, however, the names are only a recommendation to preserve the coherence in the code.

Step 1: Data collection

The first thing to think about, in case you want to plot the collected data, is the type of chart you are going to use. DCRAB is going to upgrade the report file continuously by adding new data to the charts every loop, so you need to know how this data need to be inserted in the report. The charts we used are made with the Google Visualization API so take a look to the documentation to choose one of them.

Once you have choosen one chart type, you have to understand the way you could insert data into them. Normally it consists in generate a string with the correct structure of the new point of the chart to be inserted. We normally store this data string in variables named `DCRAB_MODULENAME_DATA`.

The changes that must be made in `dcrab_node_monitoring_functions.sh` script are:

- In the function `dcrab_node_monitor_init_variables()`: the variables needed for your module must be added where the `## NEW MODULE CODE ##` line is, that is, at the end of other variable declarations. You should use the naming we have used in all the variables of the tool which starts with `DCRAB_`. A variable `DCRAB_NEWMODULENAME_DATA` should be created, which will be used to insert the data of the new module into the report file.

- Create a function, which should name as `dcrab_collect_newModuleName_data()`, after the last module function defined, where the line `## ADD NEW MODULE'S FUNCTION HERE ##` is. Collect the data and fill up the variable `DCRAB_NEWMODULENAME_DATA` must be the purpose of this function. As we have pointed out, remember to generate properly the data on `DCRAB_NEWMODULENAME_DATA` variable. For example, in most of the cases we fill up the variable with a new point of the charts, which is translated to generate a string with the information of that point using the data collected (you could understand better looking how the data is inserted at the examples in `/examples/normalReport/` or in the Google Visualization API Documentation).

- In the function `dcrab_determine_main_session()` you have to add at the beginning, where the string `## ADD INITIALIZATION OF NEW MODULE VARIABLE HERE ##` is, the initialization of the `DCRAB_NEWMODULENAME_DATA` variable if needed. After that, at the bottom of the same function, add where the string `## ADD THE CALL TO THE NEW FUNCTION HERE ##` is, the call to the function `dcrab_collect_newModuleName_data()` created.

- Finally, you must repeat the previous step inside the `dcrab_collect_data()` function in the same way.

Step 2: Display

This second step is only needed for display the data into the report, so do not do it if you do not want to visualize the data.

The changes that must be made in `dcrab_report_functions.sh` script are:

- We previously store in a file the commands used to make the upgrade in the report and after we execute them at once. The function `dcrab_write_data()` is in charge of this operation and this is where you need to put your own lines: at the bottom of the function, where the string `## NEW MODULE OPERATIONS ##` is. Be sure to append the command in the file named `DCRAB_COMMAND_FILE` which will be execute later.

- In the function `dcrab_generate_html()` a few modifications need to be make: 1) At the message `## NEW MODULE CODE 1##` you have to add the data of the chart. 2) Where the message `## NEW MODULE CODE 2 ##` is you have to add the options of the chart as defined in Google Visualization API. 3) At `## NEW MODULE CODE 3 ##` you have to create the chart object, call to the draw function and create a listener to be executed when ready event is fired (necessary to the correct drawing of the chart). 4) Add the tab of the new module at `## NEW MODULE CODE 4 ##`. Copy the structure of the above tabs to generate the new one because the names of `<li>` tag's id, button tag's id and `tabChanges()` function's arguments must be named as there. 5) Where `## NEW MODULE CODE 5 ##`. is you need to add the chart definition. The required lines to insert correctly the chart are shown below:

```
printf "%s \n" "<div id=\"processesIOChart\" class=\"chart\" style=\"disp-
lay:block;\">" >> $DCRAB_HTML
printf "%s \n" "<div class=\"overflowDivs\">" >> $DCRAB_HTML
i=1
while [ $i -le $DCRAB_NNODES ]; do
   printf "%s \n" "<div class=\"inline\">" >> $DCRAB_HTML
   printf "%s \n" "<table><tr><td>" >> $DCRAB_HTML
   printf "%s \n" "<div style=\"width: 1100px;\" class=\"header\">$(echo
$DCRAB_NODES | cut -d' ' -f $i)</div>" >> $DCRAB_HTML
   printf "%s \n" "</td></tr>" >> $DCRAB_HTML
   printf "%s \n" "<tr><td>" >> $DCRAB_HTML

   ####################################
   ## NEW MODULE CHART DEFINITION ##
   ####################################

   printf "%s \n" "</td></tr>" >> $DCRAB_HTML
   printf "%s \n" "</table>" >> $DCRAB_HTML
   printf "%s \n" "</div>" >> $DCRAB_HTML
   i=$((i+1))
done
printf "%s \n" "</div>" >> $DCRAB_HTML
printf "%s \n" "</div>" >> $DCRAB_HTML
```

Internal Report Operation

This DCRAB's operation was implemented for sysadmins. DCRAB could monitor some aspects that the scheduler is not able to collect, so this operation mode allows the sysadmins to collect some of the data that only could be catch during the job's execution. Currently, at the 2.0 version, some of the modules added to this operation mode (infiniband and disk modules) are used in the normal mode, so the functions to collect the data are the same.

As the normal operation, it needs a folder to store the necessary files to make the communication between the compute nodes. That file path is set with the variable `DCRAB_REPORT_DIR`, as the normal report operation, but it will point to another path and must be defined. This report directory will have the same structure as the normal operation one that is explained in 2.1 section but it will only execute the modules added in this internal operation.

In this operation there is no real-time report generated because the user does not request it, so its not necessary. However, the data monitored is going to be stored in a file to generate the report later. That file, which is set with the `DCRAB_IREPORT_DATA_FILE` variable, is going to be written only by the main node. The rest of the nodes will store its own data in the folders used to write and the main node will collect their data to write into the file pointed by `DCRAB_IREPORT_DATA_FILE`.

The idea of this mode is to be transparent for the user. If the user decide to not execute DCRAB in his/her job the sysadmin could insert in his/her submission script the lines described in 5.1 to execute DCRAB in internal operation mode. For this purpose was created the file /auxFiles/dcrab PBS comprobation.sh.

Activate Internal Report

The way to activate this operation mode is similar to the normal one:

```
export DCRAB_PATH=/PATH_TO_DCRAB/
export PATH=$PATH:$DCRAB_PATH/src
dcrab istart

###################################
#    BLOCK OF CODE TO MONITOR    #
###################################

dcrab ifinish
```

Notice that an extra 'i' must be added to the `start` and `finish` commands.

Code Explanation and Customization

There is not to much to explain in this part, because the core of the internal operation is the same as the normal operation one, however, there are some details that must me pointed. The scripts which compose this operation are:

- `dcrab internal report functions.sh` with two functions: 1) `dcrab internal report init variables()` to initialize the variables needed for the report and 2) `dcrab write internal data()` to write into the file that will store the data collected, which is pointed by `DCRAB_IREPORT_DATA_FILE`. This function is going to be executed only by the main node.

- `dcrab_internal_report_generation.sh` which is used to generate the internal report with all the files created per job (with the `DCRAB_IREPORT_DATA_FILE` variable). This script is adapted to our particular case, so you could create your own one to visualize the data collected with this internal operation mode in the way you want to. However, we will detail the variables you need to change to adapt this script to your own scenario in case you want to use it (in 5.3 section).

Internal Report Customization

There are some variables which may be customize in the internal report:

- `DCRAB_IREPORT_DATA_FILE` inside `dcrab_internal_report_init_variables()` function of `dcrab_internal_report_functions.sh` script. It is used to store the data monitored. It will be only written by the main node.

- `DCRAB_REPORT_DIR` inside `dcrab_init_variables()` function in `dcrab_config.sh` script. In the case of the internal report it is used to configure the path to the reporting file.

And here the variables defined in `dcrab_internal_report_generation.sh` which you could customize to adapt the script to your case and generate the internal report:

- `DCRAB_IREPORT` which is the path to the internal report file.

- `DCRAB_IREPORT_DATA_DIR` which define the path where the files pointed by `DCRAB_IREPORT_DATA_FILE` are been generated. For example, if `DCRAB_IREPORT_DATA_FILE` is defined as `/home/user/dcrab/data/JOBID` the `DCRAB_IREPORT_DATA_DIR` variable must be defined as `/home/user/dcrab/data`.

- `DCRAB_IREPORT_DATA_BACKUP_DIR` define the directory of the data backup. This variable's usage is explained deeper in 5.4 section.

- `DCRAB_NUMBER_OF_BARS` which defined the number of bars to appear in the plot.

- `DCRAB_NUMBER_OF_CHARACTERS` is used to control the width of the progress bar displayed during the execution of the script (it is just an esthetic feature, you do not need to change it).

Auxiliary Scripts

There are a few more scripts implemented to make some operations on the internal report:

- `dcrab_internal_report_clean.sh` which recollects all the data generated by all the DCRAB jobs into one file to decrease the number of folders. Furthermore, it makes a tarball of all those files and their report folder and will save it in the directory pointed by `DCRAB_IREPORT_DATA_BACKUP_DIR` variable. It also removes the bad files, which may be generated from exited or crashed jobs.

- `dcrab_PBS_comprobation.sh` is used to insert into users' jobs script the required lines to execute DCRAB in internal operation mode on PBS scheduler type file. This script modify the file entered as the first argument and prepares it to submit to the scheduler. It also checks a few scenarios where the user forgets some lines of DCRAB and trys to fix the file. Example of usage:

```
./dcrab_PBS_comprobation.sh script.pbs  --> Modifies script.pbs file as below
```

```
BEFORE                              AFTER

#!/bin/bash                         #!/bin/bash
#PBS -q parallel                    #PBS -q parallel
#PBS -l nodes=1:ppn=24              #PBS -l nodes=1:ppn=24
#PBS -l mem=150gb                   #PBS -l mem=150gb
#PBS -l cput=1000:00:00             #PBS -l cput=1000:00:00
#PBS -N daniTestJob                 #PBS -N daniTestJob
                                    export DCRAB_PATH=/home/user/dcrab/software
module load QuantumESPRESSO         export PATH=$PATH:$DCRAB_PATH/src
cd $PBS_O_WORKDIR                   dcrab istart

mpirun -np 24 pw.x < ausurf.in >& OUT   module load QuantumESPRESSO
                                    cd $PBS_O_WORKDIR

                                    mpirun -np 24 pw.x < ausurf.in >& OUT

                                    dcrab ifinish
```

Script examples

```
#!/bin/bash
#PBS -q parallel
#PBS -l nodes=2:ppn=24
#PBS -l mem=180gb
#PBS -l cput=760:00:00
#PBS -N Heusler_AgMnSn_64pts

export DCRAB_PATH=/scratch/user/DCRAB
export PATH=$PATH:$DCRAB_PATH/src
dcrab start

cd $PBS_O_WORKDIR

module load SIESTA/4.1-b3-foss-2017b

export NPROCS='wc -l < $PBS_NODEFILE'

mpirun -np $NPROCS --map-by ppr:24:node siesta  AgMnSn.fdf >& OUTPUT

sleep 200

mpirun -np $NPROCS --map-by ppr:24:node siesta  AgMnSn.fdf >& OUTPUT2

sleep 200

mpirun -np $NPROCS --map-by ppr:24:node siesta  AgMnSn.fdf >& OUTPUT3

sleep 200

dcrab finish
```

```
#!/bin/bash
#PBS -q parallel
#PBS -l nodes=1:ppn=24:mediumsize
#PBS -l mem=192gb
#PBS -l cput=1000:00:00
#PBS -N JOB_NAME

export DCRAB_PATH=/scratch/user/DCRAB
export PATH=$PATH:$DCRAB_PATH/src
dcrab start

export LSCRATCH_DIR=/lscratch/$USER/jobs/$PBS_JOBID

mkdir -p $LSCRATCH_DIR
cd $PBS_O_WORKDIR
cp -r * $LSCRATCH_DIR
cd $LSCRATCH_DIR

module load program/version

export NPROCS=`wc -l < $PBS_NODEFILE`

mpirun -np $NPROCS program >& OUTPUT_FILE

export RESULTS_DIR=$PBS_O_WORKDIR/RESULTS
mkdir -p $RESULTS_DIR
cp -r * $RESULTS_DIR
rm -rf $LSCRATCH_DIR

dcrab finish
```