

Setting Up autoMAGIC

autoMAGIC Workshop Padova

Jan Lukas Schubert

28th - 30th August 2024

Contents:

- Setting up the autoMAGIC user configuration file
- Creating a database
- Filling the database
- Adding MC sets to the database
- Adding (observation) Data to the database

Setting Up autoMAGIC

- Quite complex setup necessary
- We will set up a DB with all information we need
- Steps presented here only have to be done once in the beginning, not for each analysis
- Some steps, such as filling new data might be necessary now and then for new analyses

The autoMAGIC User Configuration File

- Specifies the paths where to find the data, the MCs, etc.
- Has to be stored in your home or in your autoMAGIC directory (!!!)
- Specifies the job submission settings

The autoMAGIC User Configuration File

```
1 # This is an example config! Your config has to be named: automagic_config.toml
2 title = "User Config for AutoMAGIC Analysis"
3
4 pic_username = 'your_username'
5 db_url = 'sqlite:///nfs/pic.es/path/to/your/local.sqlite'
6
7 [passwords]
8 user_database='magic'
9 password_database='xxx'
10 user_wiki='magic'
11 password_wiki='xxx'
12
13 [paths]
14 analysis_config = '/analysis_config/file.toml'
15
16 # Where autoMAGIC can find certain things
17 data = '/pnfs/pic.es/data/magic/Data'
18 superstar_dir_pic = '/pnfs/pic.es/data/magic/Data/SuperStar'
19 mola_path = '/data/magic/common/DataCenter/Analysis/LPOnSite/MOLA'
20 wiki = 'https://wiki.magic.pic.es/index.php/Recommended_Stereo_MC_productions'
21
22 # Paths in your setup
23 workdir = '/where/all/outputs/are/stored'
24 logdir = '/where/all/logs/are/stored/for/one/month'
25 err_logdir = '/where/logs/of/failed/jobs/are/permanently/stored'
26 tarball_dir = '/where/final/tarball/is/stored'
27 dl3_output_dir = '/where/all/DL3files/are/stored'
28
29 # Paths for special analyses
30 rep_dir = '/data/magic/common/MAGIC_lp/CCdata'
31 raw_dir = '/pnfs/pic.es/data/magic/Data/RAW'
32 calibrated_dir = '/pnfs/pic.es/data/magic/Data/Calibrated'
33 mataju_base_dir = '/pnfs/pic.es/data/magic/Data/MaTaJu'
34
35
36 [htcondor]
37 max_idle = 20
38 sleep_time = 60
39 jobs_per_node = 10
```

Let's Configure the autoMAGIC User Config

- Copy your user config into your home and call it “automagic_user_config.toml”
- Set your username
- Set the path to your Database (will be created in the next steps)
 - 'sqlite:///nfs/pic.es/your_username/database/database.sqlite'

```
1 # This is an example config! Your config has to be named: automagic_config.toml
2 title = "User Config for AutoMAGIC Analysis"
3
4 pic_username = 'your_username'
5 db_url = 'sqlite:///nfs/pic.es/path/to/your/local.sqlite'
6
```

Let's Configure the autoMAGIC User Config

- Set the current MAGIC password
- Needed for accessing some resources such as the wiki
- **NEVER** push your config file with the passwords to some public repo!

```
7 [passwords]
8 user_database='magic'
9 password_database='xxx'
10 user_wiki='magic'
11 password_wiki='xxx'
```

Let's Configure the autoMAGIC User Config

- Set the path to your analysis config file:
 - `/nfs/pic.es/user/your_username/analysis_configs/Crab_test_analysis.toml`
- The analysis config has to be changed for each analysis
- Keep the analysis config files instead of modifying them!!!

```
13 [paths]
14 analysis_config = '/analysis_config/file.toml'
```


Let's Configure the autoMAGIC User Config

- AutoMAGIC needs to find the data and some more information stored on the PIC and/or in the wiki
- These paths do not have to be changes (except e.g. for a wiki migration...)

```
16 # Where autoMAGIC can find certain things
17 data = '/pnfs/pic.es/data/magic/Data'
18 superstar_dir_pic = '/pnfs/pic.es/data/magic/Data/SuperStar'
19 mola_path = '/data/magic/common/DataCenter/Analysis/LPOnSite/MOLA'
20 wiki = 'https://wiki.magic.pic.es/index.php/Recommended_Stereo_MC_productions'
```

Let's Configure the autoMAGIC User Config

- Paths for the autoMAGIC workflow
 - In principle, you can change them but disk space is an issue!
- Please use these paths:
 - Workdir: `/pnfs/pic.es/data/magic/Legacy/AUTOMAGIC/your_name/automagic`
 - dl3_output_dir: same as workdir
 - Logdir: `/data/magic/scratch/your_username/automagic_logs`
 - err_logdir: `/pnfs/pic.es/data/magic/Legacy/AUTOMAGIC/your_name/logs`
 - tarball_dir: `/pnfs/pic.es/data/magic/Legacy/AUTOMAGIC/your_name/automagic-output`

```
22 # Paths in your setup
23 workdir = '/where/all/outputs/are/stored'
24 logdir = '/where/all/logs/are/stored/for/one/month'
25 err_logdir = '/where/logs/of/failed/jobs/are/permanently/stored'
26 tarball_dir = '/where/final/tarball/is/stored'
27 dl3_output_dir = '/where/all/DL3files/are/stored'
```

Let's Configure the autoMAGIC User Config

- For special analyses, some additional paths are needed
 - Raw data
 - Calibrated data
 - Mataju-cleaned superstar data
 - .rep files (subsystem reports needed by merpp)
- Will not be discusses further here!

```
29 # Paths for special analyses
30 rep_dir = '/data/magic/common/MAGIC_lp/CCdata'
31 raw_dir = '/pnfs/pic.es/data/magic/Data/RAW'
32 calibrated_dir = '/pnfs/pic.es/data/magic/Data/Calibrated'
33 mataju_base_dir = '/pnfs/pic.es/data/magic/Data/MaTaJu'
34
```

Let's Configure the autoMAGIC User Config

- HTCondor setup configures job submission
- Depending on the DB that is used!
- Too high values → DB overloaded
- Too low values → Analysis slow

```
36  [htcondor]  
37  max_idle = 20  
38  sleep_time = 60  
39  jobs_per_node = 10
```

Working on PIC – Some Remarks

- autoMAGIC can easily write hundreds of GB of Data
 - Take care where you write your Data!
- Currently, the Legacy directory has enough disk space
- If you plan to write huge amounts of data (exceeding ~1TB) contact the PIC admins beforehand!
- The lower the data level you start on, the more disk space you will need!

Creating a Database - Some basic alembic knowledge

- Alembic: Version control for Databases
- Structure of the DB (not its contents!) is tracked
- Starting from one initial version of the DB, upgrade and downgrade commands for each DB update are defined
- To get the newest version of the DB, all upgrades have to be executed in a row

Creating a Database - Hands-On

- Go into your autoMAGIC directory
- Execute `alembic upgrade head`
- Now you should see a list of all the DB upgrades being performed

Filling the Database

- Now we have an empty DB

- Next step: Filling the DB with some pre-defined core information
 - MC periods
 - Mars Versions
 - Cleaning levels
 - Moon conditions
 - Etc.

- In the same step: Fill (selected) MC data

Fill your Database

- The initial filling of the DB is only needed once!
- `automagic fill_database`
 - `--production_name=ST0311`
 - `--view_cone_name=ringwobble`
 - `--trigger_name=standard`
 - `--superstar=true`

Some Remarks for Filling the Database

- You can select which MCs you want to fill into your DB
- **Make sure you have the required MCs in your DB before you start your analysis!!!**
- For the MCs, you can select the following options:
 - Period
 - Viewcone (ringwobble, diffuse1.5, diffuse2.5, diffuse 4.0)
 - MC Trigger (standard, SUMT, EGALSUMT, MATAJU, EGALMATAJU)
 - MC Extension
- You can add MC productions later with `automagic insert_mc_production`
 - Same options as before

More Remarks for filling the DB

- After you filled your DB, new MARS versions and DL3 converter versions will release and you don't want to start from scratch with your DB
 - You can add the new versions afterward
- You can insert new (automagic-supported) MARS versions with **automagic update_mars**
- You can insert new (automagic-supported) DL3 converter versions with **automagic update_dl3_converter**

AutoMAGIC commands

- Always helpful: Display your options with `automagic -help`
- You can do it for more complex commands as well:
`automagic create_jobs add_data_superstar_pic --help`
- If you want to see more output, use the verbose command `-v`
 - The `-v` has to be placed behind `automagic`!

Check your Database

- You can enter your DB with `sqlite3 path/to/your/database.sqlite`
- Exit your DB with `.quit`
- Sometimes commands in the following form are faster:
`sqlite3 path/to/your/database.sqlite 'SELECT * FROM table_name'`
- Moon Conditions Table:
 - `SELECT * FROM moon_conditions;`
- MARS Versions Table:
 - `SELECT * FROM mars_versions;`
- MC tables: `monte_carlo productions`, `monte_carlo sets`, `monte_carlo runs`
 - `SELECT * FROM table_name;`

Adding Observation Data to the Database

- Adding data is a bit more complex as we want to store many pieces of information in the DB

- Workflow:
 - Search for available data
 - For each observation (target position in a specific night), we create one `add_data` job
 - The `add_data` jobs are submitted to the cluster
 - On the cluster: Extract information and write them into the DB

- Two possibilities:
 - `add_data` (starting on calibrated data)
 - `add_data_superstar_pic` (starting on superstar data)

Remarks on Adding Data to the Database

- In general you should use the `add_data_superstar_pic` command!
 - Calibrated data per default not stored on PIC anymore
- If you want to do a moon data analysis, you need to fill the calibrated data!
 - Can be a bit messy in the coordination with the staging of the files by the PIC admins
- Smoother workflow for usage of staged calibrated files still to be developed
 - Currently, some human effort required for arranging the data in a way autoMAGIC can access them

Add data to your Database

- Create jobs for adding superstar jobs with
`automagic create jobs add_data_superstar_pic`
- You can manually set dates and source
 - **Please use 2018-11-01 and 2019-03-31**
- Default values will be taken from your analysis configuration file (see session tomorrow!)
- Enter a `tmux` session and submit the jobs to the cluster with `automagic submit all`

Check your Database

- Check the `jobs_add_data_superstar_pic` table in your database
- You should see some jobs in there with the `state_id` 1