BaHaMAS: A Bash Handler to Monitor and Administrate Simulations

Alessandro Sciarra

Goethe University (Frankfurt am Main)

Granada, 23 June 2017

35th International Symposium on Lattice Field Theory

Outline of the talk

Origin of the code

Motivation

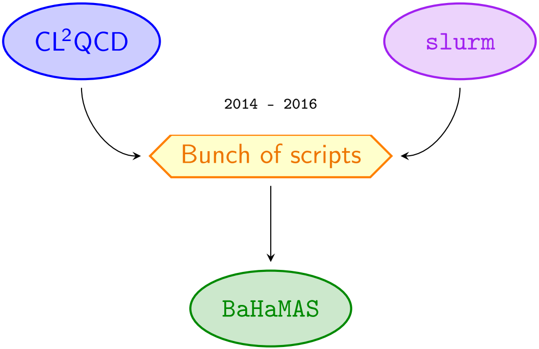
BaHaMAS in one slide

BaHaMAS at work

Summary and Perspectives

Origin of BaHaMAS as disclaimer

https://github.com/AG-Philipsen/cl2qcd



CL^{2}QCD

slurm

2014 - 2016

Bunch of scripts

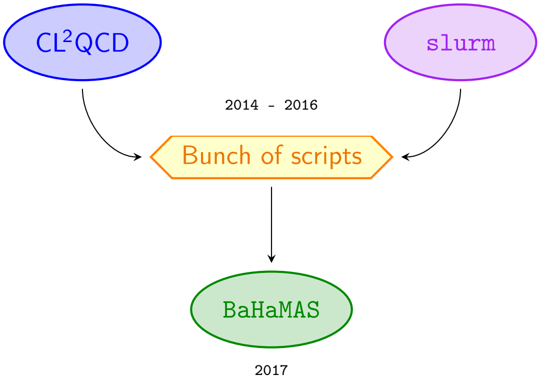
BaHaMAS

2017

https://github.com/AG-Philipsen/BaHaMAS

Origin of BaHaMAS as disclaimer

https://github.com/AG-Philipsen/cl2qcd

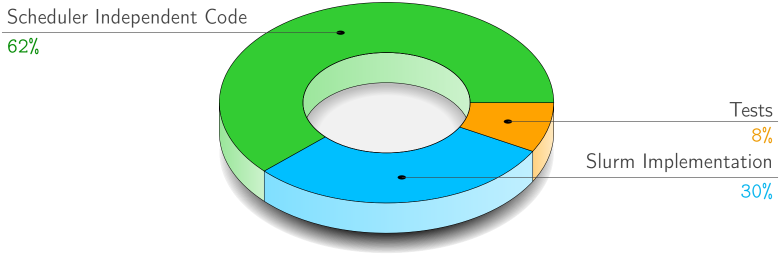


https://github.com/AG-Philipsen/BaHaMAS

Code still in an intense development phase

Origin of BaHaMAS as disclaimer

Present structure of the code



- ⚠ Better structure to extract software dependence (now hard-coded)
- ⚠ Additional scheduler and software planned to be added

Code still in an intense development phase

Daily finite temperature lattice QCD

Running simulations on a supercomputer

- A job scheduler controls the execution of programs on the cluster
- To run a simulation, a job-script file has to be produced
- The application often needs an input file as well with physical parameters
- To produce the required files by hand is the most error-prone way to proceed
- To continue a (maybe crashed) simulation is often needed
- Resuming simulations from previous checkpoints requires cleaning operations
- Working by hand is *deadly* uncomfortable and inefficient!

$\sim$1500 runs [A. Sciarra et al., Phys.Rev. D93 (2016)]

$\sim$1250 runs [A. Sciarra et al., PoS LATTICE (2016)]

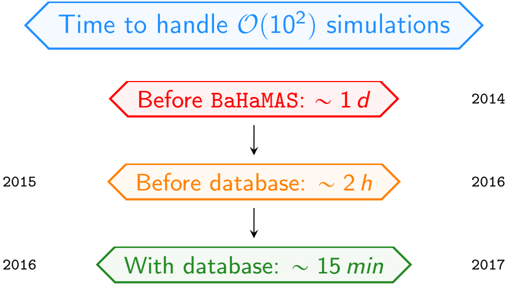
BaHaMAS at a glance

A Bash Handler to Monitor and Administrate Simulations

- Ready to go (easy interactive setup)
- Every operation is automatised: just run a script with command line options!
- Physical parameters contained in folders path and in β-file
- User friendly behaviour through smart use of β-file
- Thermalisation handled automatically using fixed naming scheme of folders
- Job-script parameters given as command line option or automatically set
- Starting, continuing and resuming simulations is straightforward
- Monitor status of simulations for different β with few keystrokes (`-l` option)
- Monitor status of *all* simulations using database functionality

BaHaMAS at a glance

A Bash Handler to Monitor and Administrate Simulations



Time to handle $\mathcal{O}(10^2)$ simulations

Before BaHaMAS: $\sim 1\,d$ 2014

2015 Before database: $\sim 2\,h$ 2016

2016 With database: $\sim 15\,min$ 2017

Even easier with *goal statistic* and *time per trajectory* fields in β-file!

⚠ Further functionality in overview and database to keep under control the remaining cost in time of the project are planned to be added

BaHaMAS at work

The β-file and the available fields

- Value of β as first, mandatory entry
- Other fields with one character prefix, in any order

PRNG seed	`sXXXX`		
Integrator steps	`iX-Y`		⚠ If needed, new ones in future
Goal statistics	`gXXXXXX`		
Time per trajectory	`tX.YZ`		
Resume trajectory	`rXXXXX`	or	`rlast`

BaHaMAS at work

The β-file and the available fields

- Value of β as first, mandatory entry
- Other fields with one character prefix, in any order

PRNG seed	sXXXX		
Integrator steps	iX-Y		⚠ If needed, new ones in future
Goal statistics	gXXXXXX		
Time per trajectory	tX.YZ		
Resume trajectory	rXXXXX	or	rlast

Two examples:

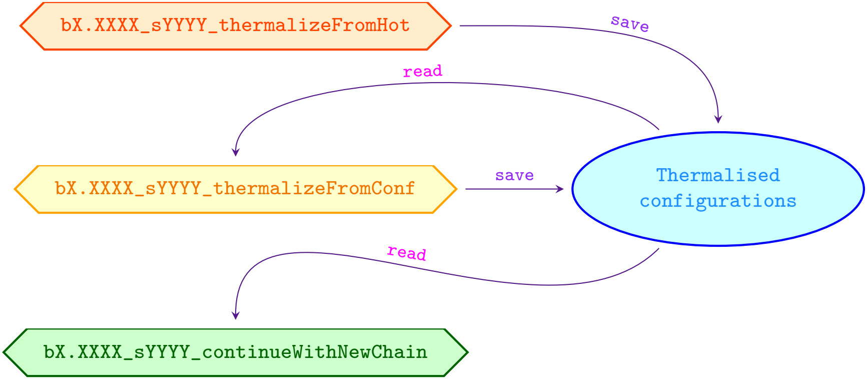
5.4280	s4280	i10-23	g95000	t170	rlast
5.4280	s5391	i10-23	g95000		rlast
5.4340	s4340	i10-21	g75000		rlast
5.4340	s5451	i10-21	g75000		rlast

5.8770		i4-9			rlast
5.8810		i4-9			r23000

BaHaMAS at work

The thermalisation phase and the naming scheme

/pathToSubmitDisk/namOfTheProject/NfX/muiX/massXXXX/ntX/nsXX

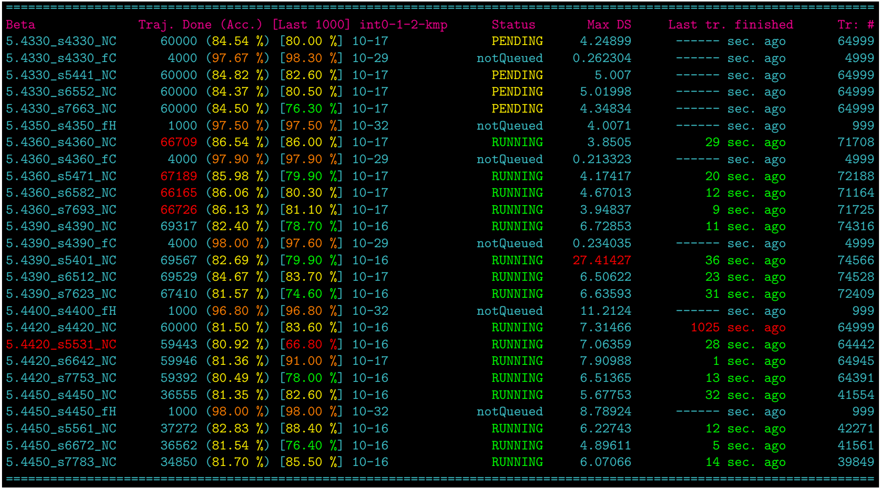

bX.XXXX_sYYYY_thermalizeFromHot
save
read
bX.XXXX_sYYYY_thermalizeFromConf
save
Thermalised configurations
read
bX.XXXX_sYYYY_continueWithNewChain

Main physical parameters contained in and extracted from path!

BaHaMAS at work

Monitor the status of the simulations in real time (`--liststatus` option)

(Example from previous to latest version of BaHaMAS)



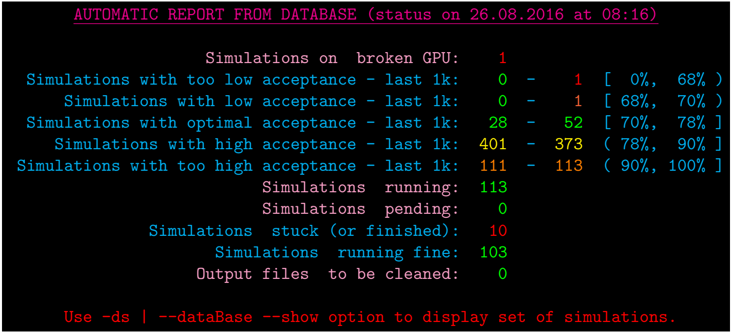
```
====================================================================================================================
Beta                Traj. Done (Acc.) [Last 1000] int0-1-2-kmp        Status     Max DS     Last tr. finished     Tr: #
5.4330_s4330_NC       60000 (84.54 %) [80.00 %] 10-17                PENDING    4.24899     ------ sec. ago      64999
5.4330_s4330_fC        4000 (97.67 %) [98.30 %] 10-29              notQueued   0.262304    ------ sec. ago       4999
5.4330_s5441_NC       60000 (84.82 %) [82.60 %] 10-17                PENDING      5.007     ------ sec. ago      64999
5.4330_s6552_NC       60000 (84.37 %) [80.50 %] 10-17                PENDING    5.01998     ------ sec. ago      64999
5.4330_s7663_NC       60000 (84.50 %) [76.30 %] 10-17                PENDING    4.34834     ------ sec. ago      64999
5.4350_s4350_fH        1000 (97.50 %) [97.50 %] 10-32              notQueued     4.0071     ------ sec. ago        999
5.4360_s4360_NC       66709 (86.54 %) [86.00 %] 10-17                RUNNING     3.8505         29 sec. ago      71708
5.4360_s4360_fC        4000 (97.90 %) [97.90 %] 10-29              notQueued   0.213323    ------ sec. ago       4999
5.4360_s5471_NC       67189 (85.98 %) [79.90 %] 10-17                RUNNING    4.17417         20 sec. ago      72188
5.4360_s6582_NC       66165 (86.06 %) [80.30 %] 10-17                RUNNING    4.67013         12 sec. ago      71164
5.4360_s7693_NC       66726 (86.13 %) [81.10 %] 10-17                RUNNING    3.94837          9 sec. ago      71725
5.4390_s4390_NC       69317 (82.40 %) [78.70 %] 10-16                RUNNING    6.72853         11 sec. ago      74316
5.4390_s4390_fC        4000 (98.00 %) [97.60 %] 10-29              notQueued   0.234035    ------ sec. ago       4999
5.4390_s5401_NC       69567 (82.69 %) [79.90 %] 10-16                RUNNING   27.41427         36 sec. ago      74566
5.4390_s6512_NC       69529 (84.67 %) [83.70 %] 10-17                RUNNING    6.50622         23 sec. ago      74528
5.4390_s7623_NC       67410 (81.57 %) [74.60 %] 10-16                RUNNING    6.63593         31 sec. ago      72409
5.4400_s4400_fH        1000 (96.80 %) [96.80 %] 10-32              notQueued    11.2124     ------ sec. ago        999
5.4420_s4420_NC       60000 (81.50 %) [83.60 %] 10-16                RUNNING    7.31466       1025 sec. ago      64999
5.4420_s5531_NC       59443 (80.92 %) [66.80 %] 10-16                RUNNING    7.06359         28 sec. ago      64442
5.4420_s6642_NC       59946 (81.36 %) [91.00 %] 10-17                RUNNING    7.90988          1 sec. ago      64945
5.4420_s7753_NC       59392 (80.49 %) [78.00 %] 10-16                RUNNING    6.51365         13 sec. ago      64391
5.4450_s4450_NC       36555 (81.35 %) [82.60 %] 10-16                RUNNING    5.67753         32 sec. ago      41554
5.4450_s4450_fH        1000 (98.00 %) [98.00 %] 10-32              notQueued    8.78924     ------ sec. ago        999
5.4450_s5561_NC       37272 (82.83 %) [88.40 %] 10-16                RUNNING    6.22743         12 sec. ago      42271
5.4450_s6672_NC       36562 (81.54 %) [76.40 %] 10-16                RUNNING    4.89611          5 sec. ago      41561
5.4450_s7783_NC       34850 (81.70 %) [85.50 %] 10-16                RUNNING    6.07066         14 sec. ago      39849
====================================================================================================================
```

Column containing the time needed to perform one trajectory omitted for readability reasons

BaHaMAS at work

The database report of the simulations (--database --report options)

(Example from previous to latest version of BaHaMAS)



```
       AUTOMATIC REPORT FROM DATABASE (status on 26.08.2016 at 08:16)

                     Simulations on  broken GPU:    1
  Simulations with too low acceptance - last 1k:    0   -    1  [  0%,  68% )
      Simulations with low acceptance - last 1k:    0   -    1  [ 68%,  70% )
  Simulations with optimal acceptance - last 1k:   28   -   52  [ 70%,  78% ]
     Simulations with high acceptance - last 1k:  401   -  373  ( 78%,  90% ]
 Simulations with too high acceptance - last 1k:  111   -  113  ( 90%, 100% ]
                           Simulations  running:  113
                           Simulations  pending:    0
               Simulations  stuck (or finished):   10
                      Simulations  running fine:  103
                    Output files  to be cleaned:    0

      Use -ds | --dataBase --show option to display set of simulations.
```

- Better naming of each entry in the report
- Distinguish between finished and stuck simulations (using goal field in β-file)
- New entries about overall statistics on the project

To understand which simulations are problematic cannot be easier!

Summary and Perspectives

At the moment:

- BaHaMAS is a tool to *efficiently* run LQCD simulations on supercomputers
- Just one single, user friendly script
- Production of configurations fully automatised
- Monitoring of ongoing simulations straightforward
- Comfortable handling hundreds of runs at the same time

Coming soon:

- Better structure of the code to extract software dependence
- New functionality in --liststatus option and in database
- Handling of future development soon using the gitflow scheme
- New scheduler/software probably included

Any suggestion, feedback or bug report is welcome!

Thank You