3rd ACCORD STAC meeting

10 June 2022 10:00-12:00 CEST
video-conference

A Consortium for COnvection-scale modelling Research and Development

Minutes

Participants

Flat-Rate ALADIN MoU5 Representatives	Rafiq Hamdi (Belgium)
	Ilian Gospodinov (Bulgaria)
RC-LACE-MoU5 Representatives	Christoph Wittmann (Austria)
	Kristian Horvath (Croatia)
	Simona Tascu (Romania)
HIRLAM-C Representatives	Saji Varghese (Ireland) **STAC Chair**
	Xiaohua Yang (Denmark)
Météo-France Representatives	François Bouyssel
	Alain Joly
	Christine Lac **STAC vice-Chair**
ACCORD PM	Claude Fischer
ACCORD CSS	Patricia Pottier
ECMWF Observer	Steve English
Invited experts from ACCORD Management Group	Alexandre Mary (Integration Leader) Daniel Santos (System Area Leader), Daan Degrauwe & Piet Termonia (SPTR Area Leaders)

Excused: Sami Niemelä (Finland)
Absent: Vanda Costa (Portugal)

1. Opening

Saji opened the meeting and welcomed the STAC members and the invited experts from ACCORD MG. He invited the new participants to take the floor and present themselves: Steve English (ECMWF observer), Daniel Santos (System Area Leader) and Daan Degrauwe (SPTR co-Area Leaders). Saji thanked Sami for his comments on the items at the agenda, sent prior to the meeting.

2. Adoption of the draft agenda

The agenda is unanimously adopted.

3. Progress on the modernization of working practices (source code forge)

Claude recalled the conclusions of the last STAC and gave an overview of the progress and the prototyping, presented the next steps and gave a broad timeline for implementation:

- **A prototype of a cloud-based, github solution for the ACCORD source code forge exists** and was used by a small group of early testers under the lead of the IL (CY48T2). **The free-plan (no-pay) github solution seems sufficient and no specific code policy issue was identified for the time being.** The intention is to generalise the use of the new source code forge for CY49T1 (Oct'22-Jan'23).
- **The DAVAÏ testing tool exists in a scripted prototype version** that was already used by several beta-users. Its use will also be further generalised across the contributing teams for CY49T1.
- To handle the ecosystem of the many repositories for/around ACCORD code, **the exploration of bundling methods and associated tools is ongoing** and a methodology for how to use the bundling tool will be proposed this autumn.
- Communication on working practices and other system aspects has started (i.e. Local Team System Representatives network) and training was discussed at MG-level and provisioned within the 2022 budget (training to ACCORD github, to DAVAÏ).

Claude drew the attention of STAC to **the necessity to identify some dedicated manpower to help maintain the new working environment and the new tools**. One specific need is the further enhancement and maintenance of the DAVAÏ technical testing tool and the proposal is to set up a "DAVAÏ-contributors" team. The tasks of this team should be:

- to add new tests
- when relevant, to update existing tests (provide new input test data, new namelist options)
- to support the IL and the Sys-AL in assessing results of DAVAÏ tests during the integration phase and on new code releases (especially for expected non bit-reproducible results)
- to coordinate tasks at technical level with scientific experts and with MG members, under the leadership of IL and Sys-AL

As a whole team, the "DAVAÏ-contributors" should be knowledgeable of the main code components (forecast, DA components, EPS specific components) as well as of typically one system implementation of each CSC.

Claude explained that t**he implications of sharing the LAM codes within the ECMWF GIT repo** have been addressed in an IFS/ARPEGE coordination meeting. The general idea is to share the codes on the basis of a short licence, fit for this particular purpose (for testing/validation purposes only, no access by ECMWF members not part of the ACCORD consortium, no predetermined duration).

In relation to ECMWF's initiative to distribute parts of the IFS codes under an open source licensing scheme, ECMWF and MF discussed the inclusion of copyright statements in the IFS (global model) codes. The proposal that was agreed on includes two statements of copyright for ECMWF (IFS) and MF (ARPEGE). ECMWF and MF furthermore agreed that an additional, non-binding statement about the presence of LAM specific codes could be added as the IFS-ARPEGE codes include pieces of code developed by the ACCORD consortium members in order to maintain their own limited area NWP applications.

Saji thanked Claude for his very clear presentation and Alexandre and Daniel for their effort on modernising the working practices. Daniel pointed out that an important goal for now is to help all Members to reach the same level of knowledge on the new tools. Saji opened the floor for questions

and comments.

Many STAC members commented on their appreciation of the work already done and of the ambitious plans.

Rafiq asked for clarification about the meaning of testing and the impact of DAVAÏ on the centralised phasing in Toulouse. He also pointed out the need for documentation. Claude answered that the testing during the integration process aims at systematic technical QA (component-wise checking, assess bit reproducibility, non-regression, check that a contribution does not break other configurations). The build of a new T-cycle will remain under the leadership of the IL, but DAVAÏ offers the means for more collaborative work from distance, while building a new code version more incrementally and continuously. The material prepared for training (webinars) will be the basis for a more complete documentation.

Xiaohua appreciated these ambitions and asked about the possible use of DAVAÏ for Continuous Deployment (CD) on top of Continuous Integration (CI). Xiaohua also asked about the DAVAÏ implementation at ECMWF. Claude answered that the testing with DAVAÏ will focus on component-wise testing, typically for a single set of input data. The specific needs for CD may be different, and Claude was considering that even the definition of what is meant by CD could be clarified. In this spirit, MG is brainstorming about the handover of new code versions from ACCORD to specific teams or groups who will install them, and how this handover could be seen from the ACCORD perspective. Daniel added that indeed CD would be a long term goal only. Regarding the porting on the ECMWF environment, Daniel explained that they are working on a portable DAVAÏ that runs on different platforms, the main one being ECMWF, and they will use the special project granted at ECMWF for that (SPFRACCO).

Christoph noted that these new working practices could be a gamechanger within the ACCORD collaboration, with the ability of getting all Members on board, thus the training, being the key for its success. Training (GIT, DAVAÏ) indeed is offered by ACCORD but it should also be accepted by the teams (with the help of the LTMs and LTSRs).

Kristian wondered about the decision procedure during the code integration process and the feedback to the teams (not only to the contributors). Claude answered that the MG will gather and monitor the proposals for code contributions. This continued effort by MG will be complemented by a specific Call for Contributions issued by the IL. Both the input by the MG and by the IL (on the Call) will be discussed at MG level. Claude added that indeed, proposals could be refused. Potential reasons for refusal could be lack of relevance w/r to the RWP, too late announcement, a contribution based on an old base version or not passing systematic testing, or problems in the code design. Daniel added that information on contributions (content, status: accepted, rejected, reasons, ..) will be public and accessible to all teams through GIT and the functions offered by the forge. In this environment, there also will be templates for contributors, and tools for reviewers (MG and/or expert scientists in the area).

Kristian commented that he would approve to go for an on-pay solution for the source forge, if needed. Daniel answered that the free-plan currently fills all needs, but an on-pay solution may be considered if further functionalities are needed.

François pointed to the growing importance of the bundling and asked about its main impacts on working practices. Alexandre answered that he foresees no technical problems but a need for changes in nomenclature:

- a code release would no longer contain all needed components, based on a stand-alone cycle version (i.e. CY48T1_op1) but rather a combination of an ecosystem of things: the IFS/Arpege/LAM cycle version, the version of OOPS, the version of SUREFX, …
- possibly a scripting system reference, a testing tool (harp version) reference, etc. would be needed later on as well
- A version control to keep track of all references of the ecosystem codes should be done and a list of what are the versions compatible together should be provided.

Following these discussions, **STAC agreed on recommendations for the Assembly (see Annex I).**

STAC also made a few suggestions to take onboard of the next steps planned by MG:
- **@PM/IL/SysAL: in the document about the new working practices, to explain more precisely the meaning of "testing"; to explain the decision procedure about contributions during integration;**
- **@SysAL: to promote the offered training (GIT, DAVAÏ) and the portable DAVAÏ at ECMWF to LTSRs.**

4. Progress on code adaptation

Claude recalled the outcome of the last STAC and the actions taken since then: the communication and the collaboration across all teams involved (ECMWF, MF, ACCORD) was enhanced; the RWP2022 was revised (SPTR Work Package) and approved by the 3rd Assembly; the CSC-L and the MG have been working on identifying experts for the physics code refactoring.

A work plan was established to make progress on the important and urgent issue of code adaptation. Two source-to-source (s2s) translators have been under consideration: LOKI is being explored, however "fxtran" seems to better fit ACCORD current needs. Regarding the code adaptation itself a stepwise approach is taken:
- a significant work was done for cleaning and simplifying the ARPEGE physics codes at MF, which provides good guidance, especially for the ALARO CSC (concrete refactoring tasks will soon be discussed);
- at high-level of GP codes, an Open-MP loop strategy has been implemented (single versus split mode) and at low-level GP codes, the current proposal is to not implement the smart data structures explicitly in the "scientific code", but to add them by the s2s tool;
- LAM spectral transforms ready for NVIDIA GPUs.

The next steps are to discuss and plan the concrete refactoring for each CSC (specific Meso-NH code refactoring is ongoing, and its outcome will enable to plan more in details for AROME and HARMONIE/AROME low-level codes); OpenACC directives are to be evaluated, following the strategy in place with Open-MP. Spectral transforms versions ready for other GPUs will be coded (AMD GPUs).

Claude explained that there is a very large intersection between ACCORD RWP tasks (from SPTR) and DEODE WP2 tasks, with the difference that the WP2 tasks are drafted with the aim of delivering a result by June 2024. ACCORD work for refactoring the physics codes towards parser-readiness does not lie in this intersection. For DEODE, some specific work (implementing the OpenACC directives, porting the codes to CPU-GPU machines, collaboration with BSC, ...) will accelerate ACCORD progress (use of s2s, code profiling, access to GPU machines, …).

Claude acknowledged the involvement of the ACCORD teams in the refactoring work per CSC, with an increase of staffing. In DEODE, besides the staff from ACCORD teams, there will be staff from partner institutes and staff to be hired.

Claude pointed out that communication and collaboration inside ACCORD (regular meetings) but also outside ACCORD (ECMWF, MF) have significantly increased and will have to be maintained at a high level.

Claude proposed an analysis of the opportunities and risks of DEODE, with respect to management aspects:
- opportunity to hire additional staff and also to secure existing staffing,
- risk that not all CSC-specific codes can be refactored at a similar speed, thus difficulties and frustration across the teams involved either in SPTR tasks or in the CSC teams might arise,
- risks if DEODE is not funded: no contribution from FMI and BSC; no extra staff hired; slower progress without the incentive to deliver for DEODE.

Saji thanked Claude for his very comprehensive presentation, congratulated the PM and the MG for the excellent progress and opened the floor for questions.

Steven asked about the key factors that led to the choice for fxtran rather than LOKI. Daan answered that they currently rely on fxtran as there is more knowledge on fxtran in ACCORD, fxtran also is faster and feeds the current needs. The ACCORD SPTR team will repeat the same exercise with LOKI, in order to compare with fxtran and to learn more about the potential of LOKI (LOKI's learning curve is found rather steep). Piet added that the SPTR team plans to implement in LOKI some of the features present in fxtran, like automated encapsulation and loop reverting for the GP CSC codes.

Rafiq proposed to warn ALARO colleagues (at the upcoming ALARO training) to pay attention to the interaction between ALARO and SURFEX when refactoring ALARO. Claude explained that there will be no specific refactoring inside SURFEX, only perhaps at the interface between SURFEX and the atmospheric computations, thus at the level of the data layout, the code itself remains computed on the CPU part. Daan added that the cost of keeping SURFEX on CPU is not big, and besides this the amount of data to be transferred between CPU and GPU for the surface remains small (this data offloading remains cheap for the surface). Daan noted that the issue for ALARO may be the complexity brought by the use of two surface schemes (SURFEX and the old ISBA code).

Xiaohua asked how the scientific changes in the physics code (i.e. radiation code in HARMONIE) are coordinated with the technical changes (refactoring). Alexandre gave examples of CY48T3, in which there were conflicts between refactoring (based on CY48T1) and scientific changes (CY48T2). These conflicts have been solved but, in the future, it would be better to avoid this kind of situation of rephasing a bunch of scientific changes as a block on top of refactoring.
Claude proposed that refactoring first enters a new cycle, then the contributors of scientific changes re-phase their changes on top of the refactoring. The CSC teams should be encouraged to regularly re-phase their developments and use a very recent version of the code. Claude gave the example of the cloud-aerosols-radiation group that has decided to go for a two-stage cleaning of their codes, with a first step of CSC-transversal cleaning (based on CY46) and then so-called "forward phasing" to CY48 to continue their common work. This group is aware that they will also have to phase with respect to the re-factoring that is going to enter for code adaptation purposes.

STAC agreed on recommendations for the Assembly (see Annex I).

5. Proposal by MG for rolling reporting by Area

Claude presented the MG proposal to implement a rolling reporting by area. The objective is to enable MG as well as the teams themselves to share key information on the progress of the tasks described in the RWP. Given the large size of most areas, the MG suggests this material to be updated in a collaborative manner, frequently, per Area, by MG members, the WP Co-leads (from the RWP) and key scientists. Apart from the MG members, attention is therefore drawn to the fact that this editing can require some specific manpower effort by WP Co-leads and key scientists, to prepare the relevant reporting material and edit the document. This specific work load should be balanced by the fact that only synthetic material is expected, and ideally any material should be taken out of otherwise existing reporting practices (e.g. done anyway in local institutes or in groupings).

Many STAC members supported the implementation of such a reporting.

Christoph proposed to ask for quarterly updates of the rolling reporting, following the same rhythm as the quarterly declarations in the Common Manpower Register.

Christoph would like more precise guidelines, in order to avoid too much inhomogeneity if everybody just copy-pasted from their local reports. Rafiq added that the final document should not be too large or too technical for the MG: the guidelines should be more clear about the size of the reporting per team.

Kristian stressed the difficulty of reporting on the deliverables as they appear in the RWP2022. Claude explained that the MG is working on a new way of describing the goals of the tasks and defining their deliverables in the Work Packages of the RWP2023. These changes will be explained at STAC in the autumn, when the RWP2023 is presented.

STAC took note of the new reporting practice by Area and encouraged key staff to contribute. STAC also proposed the MG to make the guidelines for contributors more precise about the format and the size of their reporting and the importance of summarising their main results. Christoph and Rafiq are invited to send Claude any further concrete proposals to improve the guidelines.

6. Date of the next STAC meeting

At the autumn meeting, STAC will be proposed to review the realisation of the RWP2022, the proposed RWP2023, the interactions between DEODE WPs and the RWP2023, and the progress on code adaptation.
STAC agreed to plan for a 2 (consecutive) half-day meeting on 2, 3 or 4 November 2022. Patricia will distribute a pool for the choice of the exact dates, the attendance (in-person or video), and the proposal for hosting (the host should be prepared to organise a hybrid meeting as not all participants would be able to attend in-situ, should the decision to have an in-situ meeting be taken).

7. AOB

None.

8. Closing

Saji thanked the MG and all teams for their efforts and progress and highly appreciated the enhanced communication and collaboration inside ACCORD and with other entities.
Saji thanked the STAC members and the invited experts for the fruitful discussions during the meeting and closed the meeting at 12:45.

Annex I: STAC recommendations

Modernization of working practices

- STAC acknowledges the progress made so far and recommends to continue establishing methodologies and tools based on the current choices.
- STAC supports and encourages the MG, especially PM+IL+Sys AL, to analyse the needs for dedicated, expert staffing to help setting up the new methods and maintaining the tools ("support team").
- STAC recommends the Members to consider staff contribution to the support team activity.

Code adaptation

- STAC supports the work plan proposal.
- STAC encourages the inclusive approach w/t to all 3 CSCs.
- STAC recommends the ACCORD members to secure staffing for the priority topics of SPTR, especially in case DEODE is not funded.
- STAC recommends continuing the enhanced communication and collaboration across the CSC teams in ACCORD.

Rolling reporting

- STAC takes note of the new reporting practice by Area.
- STAC encourages key staff to contribute.