

DeiC HPC-Forum

Meeting 2022-5

31/10/2022 @ 09.30 – 11.30

Zoom: <https://deic.zoom.us/j/2909030409>

Invited members and support:

- Members, present: Claudio Pica (SDU), Piotr Jaroslaw Chmura (KU), Hans Jørgen Aagaard Jensen (SDU), Lars Nondal (CBS), Martin Amüller (ITU), Thomas Bligaard (DTU), Svend Karlson (DTU), Dan Ariel Søndergaard (AU), Lars Sørensen (AAU).
- Members, absent (announced): Erik Bjørnager Dam (KU).
- Members, absent (not announced): -
- Support: Eske Christiansen (DeiC), Adam Stahl (DeiC).

Guests:

- ...
- ...

Minutes from last meeting (2022-4) added to meeting invitation.

Minutes by: Adam Stahl

Item	Issue	Comments Actions
1	<p>Information and approval of minutes from last meeting</p> <p>Approval:</p> <p style="padding-left: 20px;">A. Minutes from meeting 2022-4 (written comments/or important changes must be submitted directly to Adam Stahl and Hans Jørgen).</p> <p>Information items:</p> <p style="padding-left: 20px;">A. DeiC strategiseminar (ved Hans Jørgen)</p>	<p>→</p> <p>A: Minutes Approved.</p>
2	<p>On the recommendation for investment in interactive HPC with GPUs.</p> <p>Response to the outcome on this matter at the DeiC board meeting Oct. 2022, minutes for this point attached and sent to HPC Forum Friday 21. Oct. 2022</p> <p>It should be explained to the board what the GPU need of the end users is which cannot be fulfilled by LUMI-G and why.</p>	<p>→</p> <p>LUMI-G is expected to be up and running again by mid-November. Claudio Pica will send an update to the Forum later this week.</p> <p>The e-resource committee will begin the process of assessing the national HPC applications this week.</p> <p>There has been a presentation of the new DeiC "Cotainr" solution at the DeiC conference 2022. The target for the solution is midrange to advanced users of GPU's and will make it easier to run AI and ML research on GPU's. Test have shown that the technology can scale for one node, and the DeiC team is confident that it can scale to multiple nodes easily. A PowerPoint deck will be sent out with the minutes.</p> <p>One concern is that only few of the T1 users (estimated 10% of the users) will be able to use the "Cotainr" solution, because it is command line based. If this is the case the "Cotainr" solution will not be able to replace T1 GPU's.</p> <p>The "Cotainr" solution might be a good solution for</p>

		<p>midrange to experienced user, while T1 is very important for new users of GPUs, and LUMI-G is not a solution for these entry level users. DeiC needs to make it easy for entry users to apply for T1 HPC with GPU where it can make a difference to their research.</p> <p>There might also be a misconception in the DeiC board that most of the T1 users can easily get to a level where they can use LUMI, but that is not necessarily the case. It might be the case for T2 and T3 users, but not for T1 users.</p> <p>The Forum decided to send a recommendation to the board to clarify this misunderstanding. The recommendation should be ready for the DeiC Board meeting on the 9th of December 2022 (deadline 2nd of dec.).</p> <p>If DeiC board wants to push more researchers to apply HPC to their research, more resources on Type1 HPC are needed.</p> <p>There is also a large user group that runs calculations using sensitive data, which for that reason cannot be run on LUMI, including LUMI-G, for legal reasons. There is no definite time frame for running sensitive data from Denmark on LUMI, one estimate is that the scope for handling sensitive data on LUMI is 2-5 years.</p> <p>The T1 should be seen as for the broader use, and LUMI-G is for the specialized and high-level users.</p> <p>The DeiC resources are generally limited, and it should not be compared to LUMI. The focus should instead be to give resources to the users to get them started on Type1, so they can practice and eventually move on to more advanced facilities like LUMI.</p> <p>The clarification paper will be proofread and discussed at the next HPC Forum meeting.</p>
3	Discussion of the grant system for Type 1 interactive HPC	→

	<p>Should it for example be modified to 70-75% resources are allocated by universities (now 50%) and 25-30% by the national peer reviewed process.</p>	<p>Any suggestion for a reallocation process should start by clarifying what is legally possible. There are procedures already going on in the e-resource committee, and HPC Forum will have to ask the committee, how this can be handled administratively.</p> <p>One of the reasons why most T1 resources are granted locally is, that the applicants are entry level users that are not yet experienced enough to know what to apply for in their applications. Beside this, they are usually applying for few resources compared to the amount of time it takes to write the applications.</p> <p>The question which should be addressed is "To what degree it would be possible to make a simpler application system". HPC Forum can maybe discuss this problem with the e-resource committee.</p> <p>Another important question to ask here is "why T1 users should not be taught to write applications, when they will be asked to do it, as they move along"?</p> <p>It is however a misconception that T1 users will generally be moving into T2. Many T1 users will not move to traditional HPC systems because they do not need large parallel machines, but they need an interactive (and maybe secure) environment.</p> <p>It is also important to keep in mind, that 99% of the applicants will never be advanced users. There is a general trend in EU for large HPC systems to support more "interactive" workloads.</p> <p>Another important dimension is how it can become visible to users how to apply and work with the process. DeiC need to start talking more directly to the users.</p> <p>Hans Jørgen will take contact to the e-Resource Committee and arrange a meeting.</p>
4	Discussion of the structure with front offices and back offices	<p>→ The intended process is that</p>

	Initial discussion of the FO/BO split and whether HPC Forum should propose a new model for the FO/BO split	<p>users should always contact their FO in any case of support.</p> <p>At the conference 2022 it was stressed that the users should not contact BO directly but go through their FO.</p> <p>The DeiC board recently agreed upon having more focus on this in 2023, and most suddenly allocate resources for a central FO employee in DeiC.</p> <p>The agenda point will be postponed until it is known which direction the new board want to go.</p> <p>It will be important to define what the responsibilities for FO and BO is. For this we will also need to clearly defined roles. The process started out by setting up the functions and now it is time to figure out the exact roles.</p>
5	<p>Discussion of KPIs and other indicators</p> <p>At the HPC Forum meeting 28/09/22 it was decided that this discussion should be divided into three topics:</p> <ol style="list-style-type: none"> 1. Pure KPIs for centers to deliver on (in future contracts). 2. Indicators with information about whether the HPC resources are used/searched for. 3. Are the resources well used from the universities point of view - How much value do they get for the money. 	<p>→</p> <p>One should ask whether it is for the HPC Forum to come up with KPI's that are put in a service contract between consortium and a university.</p> <p>The recommendation should be pure HPC Forum lists of KPI's as references for the DeiC board in the negotiation with the vendors.</p> <p>It should be clear which stakeholders are responsible for the KPI's before they are put into the contracts.</p> <p>The KPI's put in the contracts need to be something that the vendors are in full control of.</p> <p>Eske Christiansen will investigate whether the Forum can get a list of existing KPI's within a week.</p>
6	AOB.	<p>→</p> <p>Nothing</p>

Contacts:

- Eske Hjalmer Bergishagen Christiansen (DeiC HPC Director), eske.christiansen@deic.dk, +45 9351 0048.
- Hans Jørgen Aagaard Jensen (HPC-Forum Chairman), hjj@sdu.dk