MIPRO 2018, May 21-25, 2018, Opatija Croatia

DOI: 10.23919/MIPRO.2018.8400035

National Supercomputing in Denmark

L.K. Andersen*; B.V. Thage*; A. Syed**; B. Andersen***; P. Løngreen**; K.G. Nielsen**** and S. Pedersen*

*Danish e-Infrastructure Cooperation, Lyngby, Denmark
**Technical University of Denmark; National Life Science Supercomputing Center, Lyngby, Denmark
***The Royal Danish Library, Cultural Heritage Cluster; Aarhus, Denmark
****University of Southern Denmark, Dept. of IT-service, Odense, Denmark
lene.krol.andersen@deic.dk; birgitte.vedel.thage@deic.dk; alisyed@dtu.dk; bja@kb.dk; peterl@dtu.dk; kgn@sdu.dk; sped@dtu.dk

Abstract - National supercomputing was for the first time introduced in Denmark in 2014. Three national supercomputers were jointly funded by the Danish Government, Ministry of Higher Education & Science, the larger universities in Denmark and the Royal Danish Library. National compute facilities were to fill the gap of computing resources between local university resources, Nordic, European and Internationally. However, more importantly the goal was to give all researchers in Denmark equal access to computing resources in order to meet and qualify within the foreseeable future of Big Data. Three years have passed and the amount of national HPC users is steadily increasing every day. Workflows across university borders have been established, national payment models are in place and eScience expert hubs and centers are forming and adjusting to local needs and resources available. All in all, eScience is forming its landscape in Denmark. This paper illustrates the journey Denmark has been through in establishing and integrating national supercomputing into its research culture. Previously, such significant compute resources were strong competitive research parameters. They are now national collaborative facilities opening up for newcomers to HPC, increasing interdisciplinary research, transferring of HPC expertise between scientific disciplines etc. Furthermore, this paper addresses the challenges and successes in reaching this cornerstone for Denmark and in addition identifies the impact of national supercomputing on science in Denmark.

Keywords - supercomputing; HPC; Denmark: eScience; research impact: Abacus 2.0; Computerome; Cultural Heritage Cluster: DeiC; national collaboration

I. INTRODUCTION

This paper reflects upon the journey Denmark has been through in the establishment and integration of national supercomputing into the Danish eScience landscape and university cultures.

A. *The History of Danish e-Infrastructure Cooperation (DeiC)*

DeiC (Danish e-infrastructure Cooperation) was established in 2012 as an organisation under the Danish Ministry of Education and Research. The organisation was the result of a merger of the supercomputing activities within the Danish Center for Scientific Computing (DCSC) and the national research and education network. The latter belonged to the Danish universities and UNI-C.

B. *The Danish Research Network*

The first computing systems of the Danish universities were based at regional computing centers. In 1985, these centers were merged into a joint organisation, named UNI-C. Around 1984, the Danish universities started connecting their local area networks nationally in order to communicate with each other. The connections were established via bridges: On one side of the bridge the local area network was connected via Ethernet, on the other side a telco line connected to the other university. Eventually, TCP/IP routers were added to the architecture, making UNI-C the first to manage a large network in Denmark based on Internet technology. The network, known as DENET, was also connected to the Internet. Later the network was organised under the Danish name 'Forskningsnettet' (i.e. The Research Network). Since 2003, the Danish research network has been financed by the universities through the national Danish budget. Until 2012, a network secretariat ran the administration while operations were outsourced to UNI-C.

C. The National eScience Competence Center

In April, 2011, the Danish Ministry of Higher Education and Science published a roadmap for research infrastructure [1]. This roadmap was based on input from six panels covering different areas of science and research. The roadmap stipulated that DeiC should be established as a merger of the DCSC and the research network. Furthermore, DeiC should establish a national eScience competence center in order to catalyse the knowledge of eScience. It was also decided that DeiC should be involved in the area of research data management.

D. Supercomputing

When DeiC was formed back in 2012; 6,7 million EUR (i.e. 50 million DKK) were transferred from the part of the national budget earmarked for research infrastructure. The money was used for establishing the three national DeiC HPC (High Performance Computing) centers. The DeiC national HPC centers are based at two individual universities and at the Royal Danish Library, established in cooperation with them (Table I). But the centers are available for all researchers at Danish universities and research organisations. Hence, the model is a combination of a centralised and decentralised organisation.

E. ABACUS 2.0

The DeiC National HPC Centre at SDU, hosts the ABACUS 2.0 supercomputer, a state-of-the-art solution for academic HPC suitable for a wide range of research and technological applications. All researchers at Danish universities and all academic users can benefit from the available resources under the same conditions. Access to the facility is granted on a pay-per-use base and all academic users, independently of their host institution, pay the same price. This allows all researchers at Danish universities to take advantage equally of the partial public funding to the national facility at SDU. The ABACUS 2.0 supercomputer

TABLE I. DeiC National High Performance Supercomputing (HPC).

HPC	*DeiC National HPC Centre, SDU* *ABACUS 2.0*	*DeiC National Life-Science Supercomputer* *COMPUTEROME*	*DeiC National Cultural Heritage Cluster* *KAC*
Host/Location	Southern University of Denmark, Odense (SDU)	Technical University of Denmark, Lyngby (DTU)	Royal Danish Library, Aarhus (KB)
Cores Storage	14,000 CPU 1 PB	19,000 CPU 8 PB	<500 CPU 0.1 PB
Access model	Pay-per-use	Pay-per-use Subscription model	Subscription model
Special features for data handling	Non-restricted data	Non-restricted data Sensitive data Proprietary data	Non-restricted data Sensitive data Copyright data*

* Special access to a collection of webpages, newspapers, radio and TV broadcasts > 4 PB data.

is currently one of the most power-efficient HPC installations in the Nordic countries and offers one of the best pricing options for academic users in Europe, also thanks to the national funding from DeiC. Users are billed according to how many node hours they are allocated on the ABACUS 2.0 supercomputer. For small projects (less than 240k corehours) resources must be spent within the agreed time frame (currently 8 months) or they will be lost. Larger projects can request access via a regular call every four months and last for a length of one year. It is possible to move up to 20% of the allocated resources to the next allocation period. Access to the HPC facility, the allocated resources as well as the appropriate user support are provided within the scope and the limits of the HPC Centre policies and regulations.

The ABACUS 2.0 HPC facility provides a pay-per-use service open to all, where national academic users are given priority. In case the requests for access exceed the available capacity,the final decision on access to the HPC facility remains with the SDU eScience Steering Committee. The Committee allocates resources at the facility taking into consideration the requests received and according to the HPC center policies and regulations.

Requests for access to the HPC facility for private/commercial users are allocated depending on the available resources and within the limits of the service agreement with DeiC.

The price for academic access to HPC resources is calculated based on the full cost of ownership to operate the facility. This includes:

- the cost of the HPC hardware and the cost of the fast, parallel GPFS storage system;
- the cost of the infrastructure needed to host the HPC facility at SDU, including: the server rooms, the cooling system, the required electrical components to connect to the power grid (e.g. transformers), etc.;
- the cost of the electricity for the whole HPC facility (HPC hardware, storage system, cooling system, etc).

Note that the pricing model at the national HPC facility at SDU includes the costs for the storage in the cost of node-hours, i.e. users are not required to pay separately for the use of the fast storage system. Furthermore, the final price is based on this full cost calculation, that is very low for Danish researchers (i.e. 1.1 Euro Cents per physical corehour; slim nodes).

F. Computerome

The DeiC National Life Sciences Supercomputing facility "Computerome" is built as a collaborative and a dedicated state-of-the-art computing platform for health care and life sciences. The system represents the Danish infrastructure in ELIXIR[1] and Nordforsk project; Tryggve[2].

Approximately 2% of Denmark's Gross Domestic Product (GDP) is related to life sciences. There is a large foundation based on pharma with a tradition of Public-Private Partnership. The demand for High performance infrastructure to support the growth has lead to a national strategy for HPC infrastructure dedicated to life sciences.

The architecture of Computerome is very unique and it is the largest system of its type in the world. Computerome has more than 1300+ users and completed 197M jobs in 2017. The architecture is tailored specifically for the workloads in healthcare and the field of life sciences. The system is built using the Security by Design principle and optimised to put a higher priority on throughput on storage, memory, reliability and accessibility than raw number of floating point operations (FLOPS), The metric that traditional HPC systems are designed around.

One of the key features of Computerome is the ability to house multiple tenants working with sensitive data, compliant with strict national regulations. Currently, Computerome consists of about 19,000 CPU cores and 8PB of storage.

Computerome hosts computing for both non-sensitive and sensitive research projects. The offering for non-sensitive data is Computerome's HPC system, with configuration aimed at data-driven research typical in biosciences. The system also offers a secure cloud computing platform both as Platform-as-a-Service (PaaS) and Software-as-a-Service (SaaS) variants, with several security profiles including extremely high-security setups use cases. Within the private cloud platform researchers are free to create their own "supercomputers on demand", comprising of hundreds of nodes (thousand of CPU cores) and petabyte-scale storage. At the same time maintaining complete freedom to define any environment they require, such as commercial analytical platforms, Next Generation Sequencing (NGS) analysis pipelines, user management schemes, queuing systems etc. The environment is built with certified, high performance and high security components designed for processing of large volumes of data.

Two-factor authentication is mandatory for all projects. Integration with Citrix Access Gateway allows locking environments down, with no internet access or possibility to "copy-paste" data out of the system.

The National Life Science Supercomputing Center has started the process of establishing a new system, which is planned to arrive in 2018. This new system will still be called *Computerome* and offer five times the capacity of the current installation. Built to handle production-grade services at a national level, using state of the art data management system to enforce the FAIR principles. The system will automatically offer full data lifecycle service.

G. Cultural Heritage Cluster

The Cultural Heritage Cluster is by far the smallest cluster in Denmark's national e-infrastructure. It would not be fair to call it an HPC given the size of the installation. Nevertheless, this specialised installation plays an important role as it fills a gap for specific type of research and specific types of researchers.

With the additional funding for HPC, at the establishment of DeiC, also came a requirement from the funders for DeiC to work on broadening

[1] https://www.elixir-europe.org

[2] https://neic.no/tryggve/

the use of the HPC technology among new user groups, especially researchers from the humanities and social sciences. The Royal Danish Library holds very large collections of digital cultural heritage (e.g. a web archive with 900 Tbytes of data, radio/TV collections with more than 3 Pbytes of data). Legal complications are an issue when moving such large collections, where some data also holds personal sensitive data, alongside the practical implications by moving Pbytes of data over a network. Hence, this led to the establishment of what today became the Cultural Heritage Cluster inside the Royal Danish Library.

The Cultural Heritage Cluster is a small infrastructure consisting of 10 nodes and a capacity of 360 CPU cores. The core technologies on the cluster are Hadoop and Spark, on top of which are access interfaces, currently in the form of RStudio and Jupyter Notebook.

The Cultural Heritage Cluster had an unfortunate start with the selection of a commercial platform provider, which were to deliver the platform and the end user interfaces. After 18 months of struggle with both the basic platform and serious problems with the e.g. handling of access rights in use cases working with personal sensitive data, DeiC stopped the cooperation. Since the spring of 2017, the Royal Danish Library has been working on a new platform, all built on open source tools with Hortonworks Data Platform as the central component. Now, the platform is fully functional and the first pilot project is running at full speed. The very first results are promising and the cluster is already improving the analysis processes for the users. The Cultural Heritage Cluster is fx. capable of statistical computing a 2 billion rows dataset within seconds, a task that would have taken hours or even days on a single laptop or server.

Apart from working with the very large digital collections of the Royal Danish Library, users are very welcome to bring other data to the platform. This could be any kind of data, but the platform and the knowledge of the library concentrates its focus on data in the form of text, images, audio and video.

Due to the delay of the establishment of the Cultural Heritage Cluster the payment model is still to be deployed. During 2018, a number of pilot projects with researchers from the humanities and the social sciences are planned to be carried out. Based on the experiences of these projects, a suitable payment model will be developed. Currently, the initial thoughts around payment models are that it is most likely to be based on a combination of a fixed time based payment (e.g. per week of access alias a subscription model) and a variable element based on usage (e.g. based on CPU-seconds and storage).

II. PAYMENT MODELS

There is a financial risk for a university to host a national supercomputer. The installation is partly subsidised through national funding alongside partner universities' co-funding, however, there is still a significant running cost on a fairly short-term lifespan hardware facility (i.e. four years), which must be able to recover both its existing costs and the costs for the next generation of hardware installation. This risk can potentially create local financial pressure. Each host of a national HPC system was to establish their own payment model in order to reach local sustainability. However, with every new installation follows a start-up period where test-runs, workflows, support set-up, service descriptions etc. need to be developed and integrated. Hence, three national HPC centers have given rise to three different payment models.

The payment model of ABACUS 2.0 benefits its users in terms of dedication to the amount of compute resources specified and applied for. However, this model challenges new HPC users, due to their inexperience within estimating the needed compute time for a given compute job. To meet this challenge of primarily new HPC users; overestimated and unused compute hours are re-allocated on-to the following allocation period, as described in the 'Supercomputing' subsection for ABACUS 2.0. Another challenge of the administration of ABACUS 2.0, is to collect the payment from the researchers up-front, before the compute job is initiated on the HPC. Hence, a research grant is paid to the granting institution through midterm or at the end of the given project period when deliverables are met. However, in order for ABACUS 2.0 to meet its local financial criteria, payment of allocated compute hours must be in place before the compute job is initiated.

The payment model of Computerome is very flexible and is based on a "pay as you go" principle. The users are charged for the "CPU hours" the system uses for working on their jobs and the "TB" storage they use. There is no pre-reservation needed as the workload in life sciences is highly unpredictable and comes in spikes.

However, it has been a challenge for many groups, independent researchers and individual clinical departments to take full advantage of the infrastructure due to this payment model. Most of the users consider this type of infrastructure a basic requirement for their work. Therefore, a different payment model is brought up for discussion by the users at the national level.

The payment model of the Cultural Heritage Cluster needs to match the most diverse landscape of users, hence a sharply defined payment model is a challenging task to establish. Users will be running a very diverse number of jobs on a very diverse type and amount of data. It is difficult to foresee individual users usage pattern. On one hand, the platform needs to be shared between a number of users simultaneously given the size of the cluster. On the other hand, the payment model needs to take into account that some users will run very big jobs using many cluster resources whereas others will work very explorative with maybe smaller jobs on smaller data sets, thus only utilising the cluster resources within normal working hours. On top of that, researchers from social sciences and humanities are (overall) not very used to working with large technical infrastructures. Therefore, they are not used to having internal budgets for buying access to technical infrastructure nor are they used to having that kind of budget in their external funding applications. Hence, both cultural and behavioral changes are needed in order to reach a focus where technical infrastructure is a natural part of the individual research budgets.

III. TURNING COMPETITION INTO COLLABORATION

Denmark holds eight universities and even with its fairly small size of a country (5.7 million inhabitants) the competitive strength between universities lives well. Up until 2012, where national funding was distributed between local compute facilities, the hardware was seen as a competitive asset for a university. Hence, when the idea of national compute resources was introduced in Denmark, it was also the beginning of a new era and a shift in the competitive mindset of the universities. They were to change a culture from, the more hardware the stronger competitive asset, to now letting go of that value proposition, and instead integrating external and national compute resources when the maximum of local capacity is being reached. Instead of strengthening local compute capacity through national funding, the time was now right to pool the money, join efforts and strive for larger and shareable compute resources, in order to prepare Denmark for the future of Big Data.

An unforeseen competitive barrier did arise through the researcher's preferred choice of compute resources. However, not due to a deselection of less HPC capacity, but instead a protests against the (external) compute resources now had to be visible in research budgets. Local compute resources have for historic reasons always been financially covered through the university's overhead-percentage, hence this financial post was invisible to the researcher. In situations where compute resources are bought through external platforms (eg. at a national HPC Center), the cost suddenly becomes visible. From a researcher's viewpoint, this means an extra cost of his/her research budget, since he/she is still paying the same overhead percentage to the university. However, from the university budget's point of view, this means a reduction of costs, since this is no longer covered through the university's overhead-percentage. Thus, accidentally the new national HPC price models became unattractive to the researcher's HPC operation conditions. However, alongside the university and researcher commitment slowly became more and more established and engaged, with the fact of national supercomputing facilities, a more supplementary compute landscape started to find its way, in a symbiosis of local, national, Nordic, European and even commercial compute resources. A landscape navigating amongst available compute resources matching the variety of HPC users at various levels, and at the same time meeting the needs arising from the Danish research landscape [2].

IV. THE ESCIENCE LANDSCAPE OF DENMARK

Alongside national supercomputers were launched in Denmark around 2014-15, the majority of the universities started to organise themselves with a focus on building stronger eScience expertises both across faculties and departments. Different concepts of eScience organisations developed around the eight universities in Denmark. Some of the larger universities established their own eScience Centers; i.e. Technical University of Denmark (DTU) and Southern University of Denmark (SDU). University of Copenhagen (KU) was the first to establish their own eScience Center within the Faculty of Science even before national supercomputing was introduced in Denmark. Aarhus university (AU) established an eScience committee with representation across scientific disciplines, Roskilde University Center (RUC) established their own HPC hub with a dedicated focus on the natural sciences. Aalborg University (AAU) established an HPC committee focussing on allocating their HPC users on to national HPC. The IT University of Copenhagen (ITU) did already have an overall eScience profile by being a dedicated university of IT. Copenhagen Business School (CBS) has continued with their existing organisation.

The Danish eScience landscape mirrors self-developing communities of principles and practice deploying ICT tooling, software and e-infrastructures, which indeed challenges the speed of new discoveries. The eScience landscape of Denmark builds on each university's organisation and mission. Denmark keeps a high priority within protecting the variation of the existing research landscape, benefitting from the diversity of each university and their individual strategies.

DeiC eScience Competence Center was established by the end of 2014 (two years later than the foundation of DeiC), as a national platform for all various eScience parties to share expertise, best practices and challenges. Thereby, creating cross fertilisation over disciplines, i.e. a crucial eScience collaboration platform bridging the various and differentiated eScience Communities across universities throughout the country. The staff of the DeiC eScience Competence Center support and collaborate with the Danish eScience communities in order to enhance their integration of tools and competences benefitting from both national and international resource opportunities.

The eScience diversity within such a small country as Denmark naturally challenges national decision making between different interests and university strategies. The strength of building national HPC facilities alongside a national collaboration platform (i.e. DeiC) is that it creates equal access for both beginners and experienced HPC users across all disciplines, within the national landscape of computing access.

The local compute units and environments around the Danish universities play a very important role in the whole life cycle of eScience and HPC competences. Hence, if local HPC reaches its level of capacity for an individual researcher, new and larger HPC requirements are to be made accessible for both the individual researcher alongside his/her local HPC support.

V. DEIC NATIONAL SUPERCOMPUTING IN DANISH RESEARCH

An analysis was made in order to track HPC publications generated from DeiC National HPC facilities in the period 2015-2017. The purpose was to illustrate the impact of HPC in Danish research. This bibliometric analysis is based on input from the HPC users (500+). The users were asked to identify the papers that have been enabled by DeiC National HPC. This resulted in collection of 290 publications across a broad range of scientific disciplines distributed in 160 different scientific peer-reviewed journals. The total of 290 scientific publications corresponds to approximately 1.3 % of the yearly amount of scientific publications in Denmark [3]. It should be noted that in the remaining 98.7% there is additional HPC related papers from local HPC clusters that are not accounted for in this analysis.

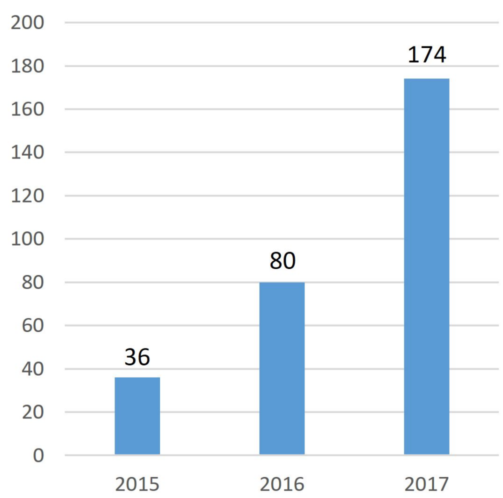

Figure 1. Number of scientific publications (n=290) using DeiC National HPC in the period from 2015 to 2017.

TABLE II. JIF ranking of scientific publications that used DeiC National HPC in the period 2015 to 2017.

JIF Interval	*n Pub*	*% Pub*	*Average JIF in the interval*
0 - 2.49	43	15	2
2.5 - 4.99	117	40	4
>5	130	45	13

Abbreviations: n Pub, Number of publications; % Pub, Percentage of publications; JIF, Journal Impact Factor.

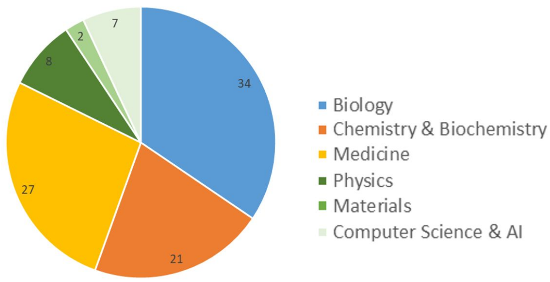

Figure 2. Distribution (%) of 290 scientific publications into research areas that used DeiC National HPC from 2015 to 2017.

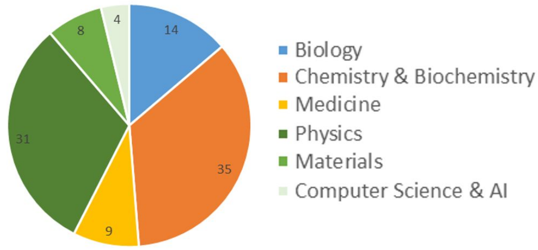

Figure 3. Distribution (%) of 80 scientific publications that used Abacus 2.0 HPC from 2015 to 2017.

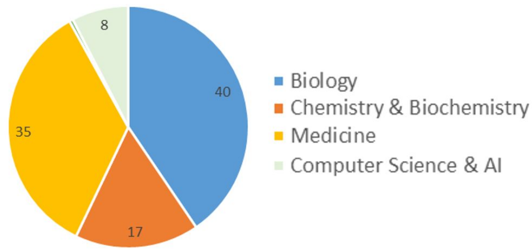

Figure 4. Distribution (%) of 210 scientific publications that used Computerome HPC from 2015 to 2017.

A. Research Areas that used national HPC

A 5-fold increase in the number of HPC publications was detected since the establishment of National HPC in 2014/2015 (Fig. 1; 2017, n=174; 2015, n=35). The majority of the HPC publications was detected in research areas such as Biology (34%), Chemistry & Biochemistry (21%) and Medicine (27%), respectively (Fig. 2). The distribution of scientific HPC publications among the National HPC facilities match the strategic intentions behind diverse HPC architecture for different scientific areas. ABACUS 2.0 was designed to aim for a broad range of scientific disciplines (Fig. 3), whereas Computerome addresses Life Science (Fig. 4). The National Cultural Heritage Cluster is intended to be used within Humanities and Social Sciences in the near future.

B. Scientific output measured by JIF

The Thomson Reuters Journal Impact Factor (JIF), is a common bibliometric index used to describe the impact of scientific publications [4]. The JIF is a proxy for the relative importance of a journal within its field. A total of 85% of the DeiC National HPC publications from 2015 to 2017 was JIF indexed above 2.5 (n=247), and 45% of the publications scored a JIF average of 13 (n=130) (Table II). As an example, the HPC publications from 2017 accounted for 17 publications (10% of total 174 publications) in journals related to the "Nature"-series [5-21], that is ranked with a JIF higher than 12 (e.g. Nature Communications (JIF 12.124) and Nature Structural and Molecular Biology (JIF 12.595)). Publications that included the use of DeiC National HPC were also published in Nature Cell Biology (JIF 20.060), Nature Methods (JIF 25.062), Nature Genetics (JIF 27.959), Nature (JIF 40.137), Nature Review Genetics (JIF 40.282), Nature Biotechnology (JIF 41.667) and Nature Reviews Drug Discovery (JIF 57), respectively, for the period 2015 to 2017 [5-35]. In addition, one paper was published in the journal Science (JIF 37.205) [36] and two papers in Nature Microbiology (No JIF, journal less than two years old) [28,29]. Denmark was recently rated as number 13 compared to other OECD-countries regarding the number of publications in Nature from 2012-2016 [37]. If the number of inhabitants was taken into consideration then Denmark was placed as number 3 on the list. This indicates that Danish research is of a high international standard. A full list of the 290 publications using national HPC in the present study is available on the DeiC National eScience Portal (Appendix 1).

In addition to counting scientific HPC publications and using a citation-based bibliographic metrics (JIF) in the analysis, one could in the future use a Field-Weighted Citation Impact (FWCI) index.

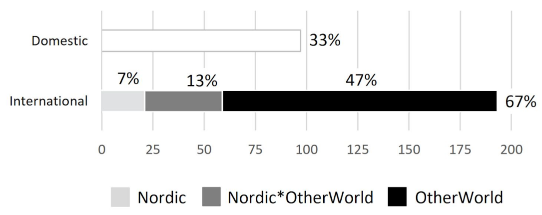

Figure 5. Distribution of 290 scientific publications that included international collaboration using DeiC National HPC from 2015-2017.

C. International Collaboration

An indicator for high scientific impact in the international community is publications made together with international research groups. Furthermore, it may be expected that international collaboration within different research areas leads to development of new competences of mutual benefit for the universities. Therefore, it is interesting to look at the rate of publications that actually involve international collaboration. A total of 58% of Denmark's scientific publications was recently rated to involve international collaboration [38]. This places Denmark as number 7 compared to other OECD-countries, which is the same level as Sweden and Norway. However, in the present study, international collaboration with other universities was detected for the majority of the scientific publications that included DeiC National HPC (Fig. 5). A total of 67% (n=193) of the publications involved International collaboration of which 20% was in collaboration with the Nordic countries (Fig. 5; Nordic and Nordic*OtherWorld).

VI. SUMMARY & CONCLUSIONS

Denmark mirrors a diverse and dynamic eScience landscape building upon a variety of organisational concepts among its eight universities. Each eScience organisation (and those without) develop and integrate their eScience competences at own pace and in line with the individual university strategies and resources. Through the national eScience Competence Center at DeiC, it is possible to share expertise and inspire across this diverse eScience landscape. At the same time, establish new partnerships, hence, slowly joining forces around a national eScience vision.

The first science in Denmark using national HPC matches the strategic intentions behind the three national HPC architectures.We have not seen the scientific outcome yet from the Cultural Heritage Cluster, however, we foresee a clear dedication from the humanities and social sciences.

These first results of the scientific output from Denmark's national HPC facilities have shown a strong scientific impact. A total of 45% of the peer reviewed scientific publications that have used national HPC in Denmark, within these last three years, lies within a Journal Impact Factor range of 5 and up to 57, holding an average of 13. Furthermore, 67% of these publications involved international collaboration - an indicator for high scientific standing in the international community for Danish research.

It is important to state that the conclusions drawn from the initial peer-reviewed scientific publications that have used national HPC, is based solely on what was possible to identify through the bibliometric analysis described. Hence, there is a potential for an even higher amount of scientific publications present, than those identified through the analysis.

ACKNOWLEDGMENT

We acknowledge the comprehensive collection and verification of the scientific publications that have used the national HPC, carried out by Executive Officer Wendy Engelberts, SDU eScience Center; Senior Executive Officer Myhanh Nguyen and Administrative Officer Spela Zajec, National Life Science Supercomputing Center, DTU; Senior Information Officer Jeannette Ekstrøm, DTU Library. We thank Director of SDU eScience Center, Professor Claudio Pica for his excellent comments and reflections and special advisor Per Møldrup-Dalum for good discussions and his engagement in the overall national collaboration, Cultural Heritage Cluster, The Royal Library, Aarhus. Last but not least, we thank Communications Officer Sanne Holm at DeiC, for her editorial work on this paper.

REFERENCES

[1] Danish Ministry of Higher Education and Science (2011) [Dansk roadmap for forskningsinfrastruktur 2011]. ISBN: 978-87-92776-00-6. https://ufm.dk/publikationer/2011/dansk-roadmap-for-forskningsinfrastruktur-2011

[2] Danish e-Infrastructure Coorporation (2017) [Analyserapport: Digital infrastruktur til forskning i verdensklasse 2025]. Page 10-22. https://www.deic.dk/da/analyserapport-digital-infrastruktur-til-forskning-i-verdensklasse-2025

[3] Danish Ministry of Higher Education and Science(2017) Table 2.4 in [Forskningsbarometeret 2017.Årlig statistik og analyse om forskning og innovation]. Page 43. ISBN:

978-87-93468-80-1. https://ufm.dk/publikationer/2018/filer/forskningsbarometer-2017.pdf

[4] McVeigh, M. E. et al. (2009) The Journal Impact Factor denominator defining citable (counted) items. JAMA. doi:10.1001/jama.2009.1301.

[5] Doll, S. et al. (2017) Region and cell-type resolved quantitative proteomic map of the human heart. Nature Communications. DOI: 10.1038/s41467-017-01747-2.

[6] Hein, J. B. et al. (2017) Distinct kinetics of serine and threonine dephosphorylation are essential for mitosis. Nature Cell Biology. DOI: 10.1038/ncb3634.

[7] Hendriks, I. A. et al. (2017) Site-specific mapping of the human SUMO proteome reveals co-modification with phosphorylation. Nature Structural & Molecular Biology. DOI: 10.1038/nsmb.3366.

[8] Hopkins, J. F. et al. (2017) Mitochondrial mutations drive prostate cancer aggression. Nature Communications. DOI: 10.1038/s41467-017-00377-y.

[9] Jiang, X. et al. (2017) Dissemination of antibiotic resistance genes from antibiotic producers to pathogens. Nature Communications. DOI: 10.1038/ncomms15784.

[10] Lasko, L. M. et al. (2017) Discovery of a selective catalytic p300/CBP inhibitor that targets lineage-specific tumours. Nature. DOI: 10.1038/nature24028.

[11] Li, T. et al. (2017) A scored human protein-protein interaction network to catalyze genomic interpretation. Nature Methods. DOI: 10.1038/nmeth.4083.

[12] Maretty, L. et al. (2017) Sequencing and de novo assembly of 150 genomes from Denmark as a population reference. Nature. DOI: 10.1038/nature23264.

[13] Northcott, P. A. et al. (2017) The whole-genome landscape of medulloblastoma subtypes. Nature. DOI: 10.1038/nature22973.

[14] Sanz, F. et al. (2017) Legacy data sharing to improve drug safety assessment: the eTOX project. Nature Reviews Drug Discovery. DOI: 10.1038/nrd.2017.177.

[15] Sapkota, Y. et al. (2017) Meta-analysis identifies five novel loci associated with endometriosis highlighting key genes involved in hormone metabolism. Nature Communications. DOI: 10.1038/ncomms15539.

[16] Stella, S. et al. (2017) Class 2 CRISPR–Cas RNA-guided endonucleases: Swiss Army knives of genome editing. Nature Structural & Molecular Biology. DOI: 10.1038/nsmb.3486.

[17] Stella, S. et al. (2017) Structure of the Cpf1 endonuclease R-loop complex after target DNA cleavage. Nature. DOI: 10.1038/nature22398.

[18] Weiner, D. J. et al. (2017) Polygenic transmission disequilibrium confirms that common and rare variation act additively to create risk for autism spectrum disorders. Nature Genetics. DOI: 10.1038/ng.3863.

[19] Weischenfeldt, J. L. et al. (2017) Pan-cancer analysis of so matic copy-number alterations implicates IRS4 and IGF2 in enhancer hijacking. Nature Genetics. DOI: 10.1038/ng.3722.

[20] Yang, B. et al. (2017) Protein-altering and regulatory gene-tic variants near GATA4 implicated in bicuspid aortic valve. Nature Communications. DOI: 10.1038/ncomms15481.

[21] Zhang, G. et al. (2017) Bub1 positions Mad1 close to KNL1 MELT repeats to promote checkpoint signaling. Nature Communications. DOI: 10.1038/ncomms15822.

[22] Ehret, G. B. et al. (2016) The genetics of blood pressure regulation and its target organs from association studies in 342,415 individuals. Nature Genetics. DOI: 10.1038/ng.3667.

[23] Ellinghaus, D. et al. (2016) Analysis of five chronic inflammatory diseases identifies 27 new associations and highlights disease-specific patterns at shared loci. Nature Genetics. DOI: 10.1038/ng.3528.

[24] Hu, J. X. et al. (2016) Network biology concepts in complex disease comorbidities. Nature Reviews Genetics. DOI: 10.1038/nrg.2016.87.

[25] Malaspinas, A. S. et al. (2016) A genomic history of Aboriginal Australia. Nature. DOI: 10.1038/nature18299.

[26] Menzel, P. et al. (2016) Fast and sensitive taxonomic classification for metagenomics with Kaiju. Nature Communications. DOI: 10.1038/ncomms11257.

[27] Pedersen, H. K. et al. (2016) Human gut microbes impact host serum metabolome and insulin sensitivity. Nature. DOI: 10.1038/nature18646.

[28] Plichta, D. R. et al. (2016) Transcriptional interactions suggest niche segregation among microorganisms in the human gut. Nature Microbiology. DOI: 10.1038/nmicrobiol.2016.152.

[29] Roager, H.M. et al. (2016) Colonic transit time is related to bacterial metabolism and mucosal turnover in the gut. Nature Microbiology. DOI: 10.1038/nmicrobiol.2016.93.

[30] Allentoft, M. E. et al. (2015) Population genomics of Bronze Age Eurasia. Nature. DOI:10.1038/nature14507.

[31] Besenbacher, S. et al. (2015) Novel variation and de novo mutation rates in population wide de novo assembled Danish trios. Nature Communications. DOI: 10.1038/ncomms6969.

[32] Forslund, K. et al. (2015) Disentangling type 2 diabetes and metformin treatment signatures in the human gut microbiota. Nature. DOI: 10.1038/nature15766.

[33] Munck, C. et al. (2015) Limited dissemination of the wastewater treatment plant core resistome. Nature Communications. DOI: 10.1038/ncomms9452.

[34] Schölz, C. et al. (2015) Acetylation site specificities of lysine deacetylase inhibitors in human cells. Nature Biotechnology. DOI: 10.1038/nbt.3130.

[35] Schölz, C. et al. (2015) Avoiding abundance bias in the functional annotation of posttranslationally modified proteins. Nature Methods. DOI: 10.1038/nmeth.3621.

[36] Somyajit, K. et al. (2017) Redox-sensitive alteration of replisome architecture safeguards genome integrity. Science. DOI: 10.1126/science.aao3172.

[37] Danish Ministry of Higher Education and Science (2017) Table 2.8 in [Forskningsbarometeret 2017.Årlig statistik og analyse om forskning og innovation]. Page 54. ISBN: 978-87-93468-80-1. https://ufm.dk/publikationer/2018/filer/forskningsbarometer-2017.pdf

[38] Danish Ministry of Higher Education and Science (2017) [Forskningsbarometeret 2017.Årlig statistik og analyse om forskning og innovation]. Page 57-59. ISBN: 978-87-93468-80-1. https://ufm.dk/publikationer/2018/filer/forskningsbarometer-2017.pdf

Appendix 1. List of www-material sources

DeiC National eScience Portal [Vidensportalen]:
https://vidensportal.deic.dk/en/publications

DeiC National HPC Centre, SDU (ABACUS 2.0):
http://www.abacus.deic.dk

DeiC National Life-Science Supercomputer (Computerome):
http://www.computerome.dk

DeiC National Cultural Heritage Cluster (KAC):
https://vidensportal.deic.dk/en/KAC