SNIC Newsletter

Spring 2022

Published by the Swedish National Infrastructure for Computing

Contents

View from the SNIC director's desk

Welcome to the premiere issue of the SNIC newsletter!

For any readers who are not already familiar with the Swedish National Infrastructure for Computing, SNIC provides cost-efficient resources for large-scale computation and data storage that may be used by academic researchers from universities, university colleges and research institutes in Sweden. The resources are made available through open application procedures which ensure that high-quality Swedish academic research is supported.

SNIC has existed in various forms since 2003, and since 2018 it has operated as a consortium comprising ten universities: Chalmers University of Technology, the University of Gothenburg, the Karolinska Institute, the KTH Royal Institute of Technology, Linköping University, Lund University, the Swedish University of Agricultural Sciences, Stockholm University, Umeå University, and Uppsala University (which is the host for the infrastructure). SNIC is funded by the Council for Research Infrastructures (RFI), which is part of the Swedish Research Council (VR), with co-funding from the consortium. The infrastructure's compute and storage resources are hosted at and operated by consortium partners. SNIC also provides a range of services that assist researchers to utilise those resources efficiently. For example, all the partner universities provide expertise via a national user support service.

As most of you are probably aware, there is a new development happening this year and into next year. The responsibility for providing computational and data storage resources for academic research in Sweden is transitioning from SNIC to a new national infrastructure that is in the process of being determined and then established. This has naturally raised questions amongst the many academic researchers who are currently using the SNIC infrastructure for their research computations, simulations and data handling; people are wondering how these organisational changes may impact their research. To answer such questions and keep everyone informed about the transition process, SNIC will be publishing several newsletters over the course of this year.

To give a little historical background to the situation: in June last year, the Swedish Research Council decided to form a new infrastructure to provide computational and data handling capacity for academic research in Sweden, rather than extending the SNIC consortium in its current form (which had been set up to run to the end of this year). In December last year, the VR opened a call for proposals to host a new national infrastructure to take over the duties currently performed by SNIC. That call closed recently at the end of March.

The major Swedish universities all agreed to submit one joint application for an organisation to be known as the National Academic Infrastructure for Supercomputing (NAIS) which would be hosted by Linköping University. No other academic research organisations submitted applications in response to the call, so the NAIS application is currently being evaluated by the Council of Research Infrastructures, with the expertise of an international panel being utilised to complement the RFI's advisory committees. The plan is that a decision will be finalised in late June.

SNIC will continue to operate the existing infrastructure for the remainder of this year, while also playing a vital role in facilitating the smooth transition to the new organisation. At present, work is being undertaken to identify the key aspects of the transition, with a view to using those as a basis to formulate an optimal plan for implementing the necessary changes. To facilitate this process, SNIC is in regular dialogue with a working group from RFI and is also ready to initiate contact with NAIS as soon as the new organisation is able to do so. This should enable the transition process to commence immediately after the VR makes the final decision in June.

NAIS will take over the existing SNIC resources and operate them throughout the remainder of their lifetimes. In addition, the NAIS application includes plans for procuring and inaugurating a new general computational resource in 2025. Hopefully, this will eventuate on schedule, although SNIC's experience has been that various mishaps can delay such processes (as happened with the pandemic causing delays with the installation of the first phase of the Dardel system last year). To minimise potential discontinuity, SNIC is now reviewing, adapting and investing in existing resources to mitigate the risks of a future lack of adequate capacity to meet the needs of Swedish researchers. At their latest meeting, the SNIC board made two decisions to ensure that resources which are currently being managed by SNIC will continue to have sufficient capability and capacity through to 2025: the general compute resource Dardel will have its CPU partition strengthened, and the Bianca system (which is dedicated for working with sensitive data) will be upgraded to cope with more data and to improve its capacity for data analysis.

In addition, to facilitate the start-up process of NAIS, SNIC is presently performing a needs analysis for the future national infrastructure in regard to both computational resources and data management (including the handling of sensitive data). NAIS will be able to use the results of this analysis to help determine which infrastructure investments would be suitable for fulfilling the future needs of researchers from the broad range of research areas that all depend on this infrastructure.

Overall the message that you should take away from this is that SNIC, VR and NAIS will be doing everything in their power to continue to provide all the researchers who use resources currently provided by SNIC with the same level of service during the period while those resources are transitioned to NAIS management. We are all working together to make the transition as seamless as possible, and look forward to sharing more news with you as further details of the transition process become available.

...Lars Nordström, SNIC Director

SNIC resources to help your research

SNIC centres

The SNIC supercomputer systems and storage resources are hosted at six centres: C3SE (Chalmers Centre for Computational Science and Engineering) at Chalmers University of Technology in Gothenburg, HPC2N (High Performance Computing Center North) at Umeå University, LUNARC (LUNARC Center for Scientific and Technical Computing) at Lund University, NSC (National Supercomputer Centre) at Linköping University, PDC (PDC Center for High Performance Computing) at the KTH Royal Institute of Technology in Stockholm, and UPPMAX (Uppsala Multidisciplinary Center for Advanced Computational Science) at Uppsala University. Swedish researchers also have access to the LUMI pre-exascale system, which is hosted at the CSC – IT Center for Science in Finland.

Compute resources

SNIC provides a variety of high-performance computing systems for different types of research. Eligible researchers can apply for a time allocation to run a project on the appropriate system. Dardel at PDC and Tetralith at NSC are two general-purpose computational systems, while Alvis at C3SE is an accelerator-based resource dedicated to AI research and research using AI techniques. The Rackham system at UPPMAX is focused on life sciences, and the venerable Kebnekaise system at HPC2N will be decommissioned after 2022.

SNIC SENS is a SNIC resource dedicated for research involving sensitive data (in other words, where aspects of the data need to remain private). SNIC SENS is based at UPPMAX and primarily consists of the Bianca computational system, along with the Castor storage system which will soon be replaced with a new storage solution, Cygnus.

The SNIC Science Cloud (SSC) is a community cloud based on OpenStack which is intended as a complement to SNIC HPC clusters. The SSC provides advanced functionality for researchers who need more flexible access to resources (for example, if more control over the operating systems and software environments is needed). Some typical uses of the SSC are for developing research software where added flexibility (such as having administrative privileges) is an advantage, or for exploring recent technology (such as for "Big Data") or building workflows combined from multiple services.

LUMI is a general-purpose pre-exascale compute system that has several partitions: one featuring CPUs, one that is GPU-based, and one for analysis, as well as a quantum simulator. Sweden owns 3.455% of LUMI, which is allocated to researchers through SNIC. SNIC also encourages Swedish researchers to apply for time allocations on the EuroHPC JU part of LUMI via PRACE.

Storage resources

SNIC storage is intended exclusively for active research data. Typically, a research group will obtain original data to use in their research (for example, by collecting it). That data is then copied to a SNIC resource, possibly for preprocessing, after which it is analysed or used in simulations or other forms of computation. The resulting data may undergo post-processing before being migrated back to the research group's own storage (for example, at their home institution). It is important to be aware that the Swedish higher education institutions are responsible for the long-term preservation of research data produced by their researchers.

All the above SNIC compute resources have an accompanying storage system, often in the form of a high-performance parallel file system, and researchers can apply for storage allocations through SNIC. With some of the SNIC systems, a default storage allocation may be assigned automatically when a research project is awarded a time allocation.

Researchers who need to work with vast quantities of data that exceed the capacity of the storage associated with specific SNIC systems, or who require access to data through multiple SNIC services, can apply for an allocation on SNIC's Swestore system. Although this data storage system appears as a single large system to users, it is deployed across various SNIC centres. SNIC also manages a tape storage system, primarily for backup purposes and to meet infrastructure needs.

National user support

SNIC provides practical support and training for Swedish researchers to help them make efficient and effective use of HPC in their research. Application experts (AEs), whose combined expertise covers most areas of research conducted on SNIC systems, are located at several of the SNIC centres. Assistance from these AEs is available to ongoing research projects that are using SNIC or EuroHPC resources. AEs can also assist research groups to prepare applications for potential future projects. In addition, researchers needing further assistance for a particular project may apply for dedicated user support from SNIC.

Training

SNIC organises a range of training events that enable researchers in Sweden to make the most of the resources that SNIC provides; training is an important and integral part of how SNIC supports Swedish research. In addition to online and onsite training events, online documentation is provided. For some training, SNIC collaborates with other organisations, such as ENCCS, PRACE, CodeRefinery and the LUMI consortium. Regular email newsletters advertise the training available not only from SNIC, but also from our partners. For more details, see the SNIC training website: https://snic.se/support/snic-training.

Note: Due to the conditions set forth regarding the financing of SNIC, SNIC resources may not be used either for undergraduate education or industry research at SNIC's expense.

Latest SNAC allocations

The SNIC Allocations Committee (SNAC) is tasked with distributing the available SNIC resources to Swedish research projects.

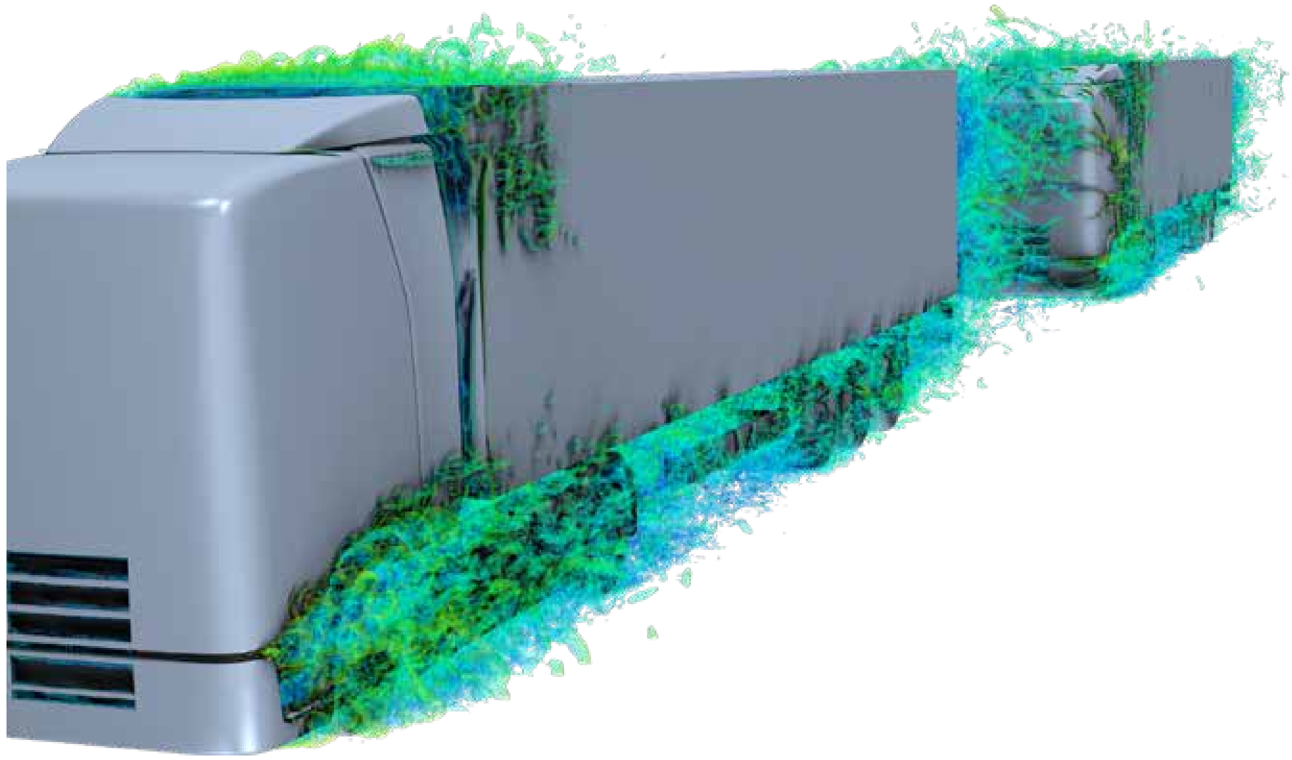
There were about a hundred million core hours per month available on SNIC resources in 2021. This time was distributed between 60 large-scale projects (which used about 75% of the total time), plus 800 medium-sized projects and about 1500 small projects (which shared about 25% of the time). Applications came in from all the universities and research institutions in Sweden. Typically the total amount of time that is requested is about twice as much as is available, and about three quarters of the large-scale projects are awarded allocations – these projects represent the traditional high-performance computing disciplines (such as computational chemistry, materials science and fluid dynamics) and also social sciences (such as studies on poverty in Africa using advanced machine-learning methods).

Making road vehicles more efficient

Improving the energy efficiency of road vehicles and reducing their greenhouse gas emissions are vital aspects of transforming modern society so it becomes fully sustainable. To do this, we need better designs for vehicles.

Simone Sebben and the members of the Road Vehicle Aerodynamics & Thermal Management Group (RVAD) at Chalmers University of Technology in Gothenburg are using the SNIC system Tetralith (which is based at the NSC SNIC centre in Linköping) to find ways to minimise the aerodynamic drag for road vehicles and improve their thermal efficiency. Optimising the aerodynamic properties of road vehicles can help to reduce the levels of greenhouse gases that are emitted and to increase the driving range of Battery Electric Vehicles (BEVs).

When there is a layer of air flowing over the outside surface of a vehicle, that "surface layer" of air can become separated from the vehicle surface and become turbulent (as shown in the image below) which then increases the drag on the vehicle. To optimise the efficiency of vehicles, different designs and solutions need to be tested. It is both time-consuming and extremely expensive to build multiple physical models and test them, so vehicle research and development is relying more and more on mathematical models that simulate the air flow around vehicles. Data from the simulations (which are run on high-performance computing systems) can then be used to design more efficient vehicles.



This image from RVAD shows the results from a simulation of the air flow around a two-truck platoon (that is, two trucks being driven at a fixed distance apart from each other). A lot of turbulence can be seen around the wheels of the trucks. The computations for this type of simulation typically take about 150,000 CPU hours.

Modelling the flow of air around a road vehicle, such as a car or truck, is very complex due to the presence of three-dimensional, unsteady air flow structures that are linked to regions on the vehicle surface where separation is likely to occur. These separation-prone regions typically have a very large negative effect on the drag and thermal behaviour of vehicles. The RVAD group has extensive experience with these types of flows and works in close collaboration with the Swedish automotive industry and a range of software companies on various research and development projects. The group also does collaborative research with the KTH Royal Institute of Technology, Linköping University, the Polytechnic University of Milan, and with other divisions and departments at Chalmers.

Access to large computational resources, such as the SNIC systems, is critical for the research and development being performed by RVAD and its many partners. The simulations conducted by the group are based on the use of the Delayed Detached Eddy Simulation (DDES) numerical approach, which is applied to detailed geometries that represent the vehicle surfaces, often at full-scale, and using realistic values for the Reynolds numbers. The mathematical equations for modelling flows include values, known as Reynolds numbers, which are used to indicate how turbulent a flow is. As the movement of particles in a turbulent flow is much more complex and chaotic than in a non-turbulent flow, the calculations for modelling turbulent flows are likewise far more complex and time-consuming. Consequently, when modelling turbulence around moving vehicles, it is important to use realistic Reynolds numbers to obtain accurate results about drag and vehicle efficiency. These myriad calculations make the whole process very computationally demanding. As an example, in recent years, the RVAD research projects have used approximately 700,000 CPU hours each month on SNIC systems and on resources based at RVAD's main industrial partners.

In aerodynamics, the group's focus has been on looking at two major areas where drag can be reduced: the wake behind vehicles and the air flow around wheels. For example, a recent PhD project extensively studied how the wake behaves when a vehicle is being driven in crosswinds. (The wake refers to the disturbed, and usually turbulent, flow of air behind a vehicle that results from the air flowing around the vehicle as it is driven along.) Vehicle aerodynamics are typically assessed in idealised conditions (that is, with no crosswinds) and with low turbulence levels. However, on the road, vehicles are often exposed to unsteady effects from crosswinds, which are known to increase the drag. The conclusion from the project was that optimising vehicle geometry without considering yaw (how a vehicle could swing to the side due to a crosswind) can reduce the performance over the entire operating range. In recent years, the RVAD group has extended its areas of research into heat transfer in vehicles, which is often referred to as thermal management (TM). The majority of that work has been dedicated to understanding underhood flow (that is, air flow involving the systems under the hood or bonnet of a vehicle, such as the engine and cooling fans) and to proposing solutions to guarantee that the necessary amount of cooling is provided without influencing the drag negatively. Cooling of other vehicle systems, such as the brakes, has also been a large part of the group's work. With the growing demand for vehicles that do not produce greenhouse gases, RVAD's TM research is now focusing mostly on BEVs. Controlling the temperature of batteries and electric components so that they operate under optimal conditions will increase their life cycle and increase the driving range of vehicles.

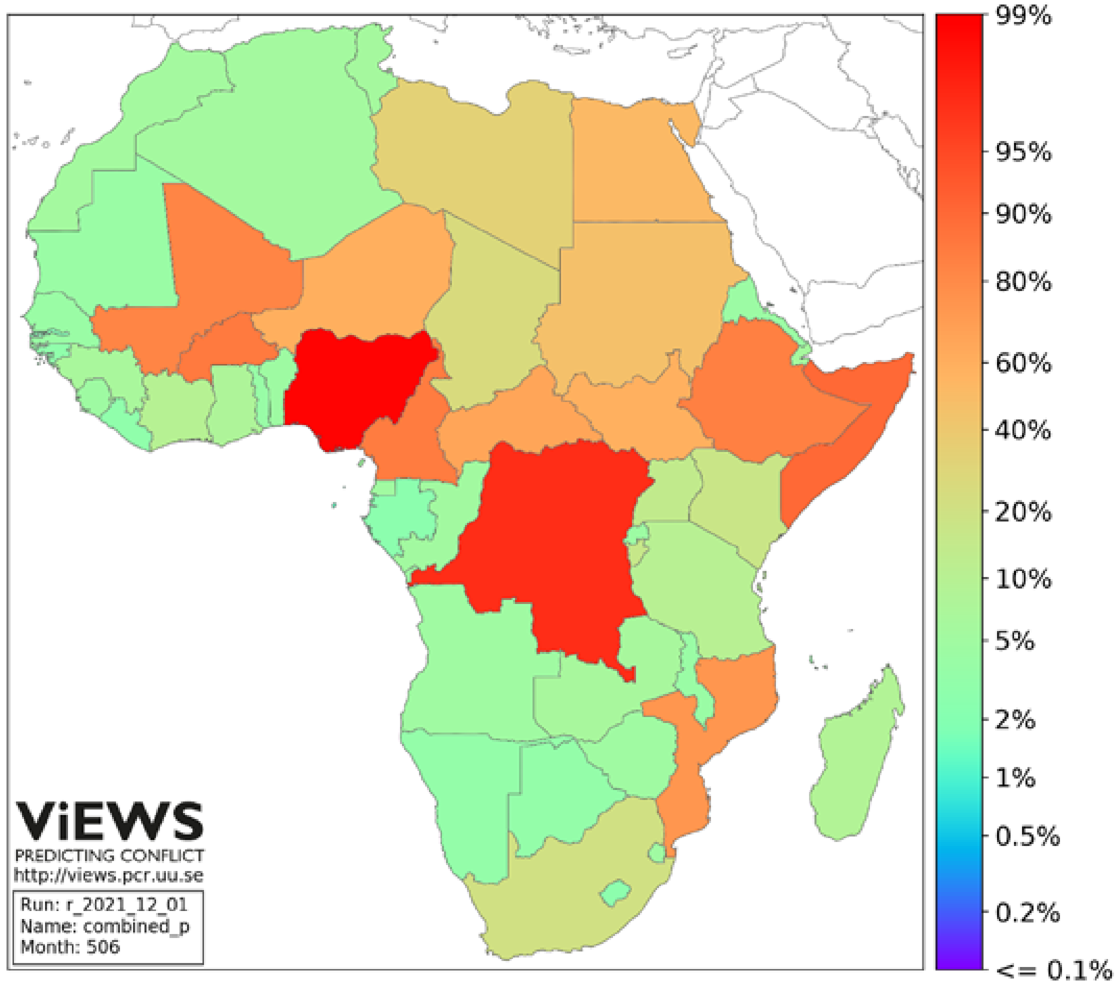
RVAD also works on a broader perspective and studies solutions for future transportation systems, such as platooning (that is, linking two or more trucks using automated driving support systems so they maintain a set distance between each other) and high-speed and crosswind driving stabilities. For more information about RVAD's work, see https://www.chalmers.se/en/departments/m2/research/veas/Pages/Road-Vehicle.aspx.

Developing a system to predict violence

Håvard Hegre and his colleagues at the Department of Peace and Conflict Research at Uppsala University have been using the SNIC system, Rackham, for developing a system to provide early warnings about the likelihood of violence escalating.

The Violence Early Warning System (ViEWS) project started in 2017 when it began working on a pilot study focusing on the African continent. The project started by developing an early warning system for predicting armed conflict based on the statistical analysis of data and then moved on to testing and improving the system.

The ViEWS system uses a variety of different machine learning models which "learn" from data about armed conflict that is being stored at the Uppsala University campus as well as from information about a large set of structural factors. The system has been producing forecasts that have been made available to the public each month since 2018. (If you are curious, have a look at: https://viewsforecasting.org.) Last year, a version of the system was expanded to include the Middle East and will be used by the UN Economic and Social Commission for West Asia in their work with member governments and local civil-society actors; other UN organisations are also interested in using versions of the system.



The ViEWS team is in the process of significantly expanding and refining the machine learning system to predict not only how likely it is that conflict will break out but also the probable severity of such violence. This extension will incorporate more complicated and more computationally intensive models (such as neural networks) into the system and also work with text-as-data machine learning models based on news.

The ViEWS project is in the relatively early stages considering how much work is involved in extending the system to reliably predict outbreaks of violence anywhere in the world. Recent events have highlighted how useful it would be to have the early warnings about likely forced population displacement that the project is working towards producing.

"The challenges of preventing, mitigating, and adapting to large-scale political violence are daunting, particularly when violence escalates where it is not expected. Early warning systems of sufficient quality and transparency do not exist, limiting the ability of the international community to effectively assist affected populations."

The project uses the SNIC cluster Rackham which is based at the UPPMAX SNIC centre. Rackham contains 9720 cores in the form of 486 nodes, each of which is made up of two 10-core Intel Xeon V4 CPUs. Computational time on Rackham is used to update the ViEWS monthly forecasts, as well as researching improvements to the model system. The project uses forecasting software written by the ViEWS team, mainly in the Python language.

The process or pipeline for producing the monthly predictions consists of three main steps: preparing the data, model fitting and finally forecasting. For each forecast, the system does calculations to produce five to ten sets of data. After that, about a hundred features are created for each of these sets and then merged back into the main data sets. Next, hundreds of models are fitted to each dataset. When it comes to the forecasting, two methods are used: the first produces ten thousand distinct forecasts in a single task which are then averaged, and the second produces a single forecast for a given model.

Many aspects of these three steps can be computed in parallel. Being able to run the computations in parallel on the SNIC supercomputers makes it feasible to produce a monthly prediction within a month. To give an idea of the computational scale, the monthly run for ViEWS uses about 30,000 core hours if everything goes perfectly on the first try. This is rarely the case because, as the system is improved, errors are often introduced, which can necessitate re-running large parts of the whole process/pipeline. That can add up to over 50,000 hours in total. By the time additional development of models and software is factored in, the project can use towards 80,000 core hours each month.

However, not all aspects of the steps can be broken down into tasks that can easily be performed in parallel. So, in addition to working on the models related to conflict, the team is also addressing a problem that is relevant to many researchers who are adapting or developing systems to run on HPC resources, namely working on parallelising calculations that are difficult to process in parallel by handling them asynchronously. Anyone who is interested in learning more about the team's approach is welcome to contact ViEWS (see https://viewsforecasting.org).

The map on the previous page shows an example of the VIEWS predictions for countries in Africa in February 2022. The predicted risk can range from 0 to 100% and indicates the likelihood that there will be 25 fatalities (arising from the three main types of political violence) in any given country during that month.

SNIC SENS: Bianca is getting a facelift

SNIC SENS is a SNIC resource that provides secure handling of sensitive research data in an HPC environment. The Bianca cluster at UPPMAX has been a core part of SNIC SENS from the start, along with its storage system Castor. These provide the ability to analyse sensitive data in an isolated environment.

Originally, Bianca and Castor were dedicated to the analysis of massive molecular biology data. Through various collaborations that have arisen over time, Bianca now handles other types of sensitive research data as well. This includes population registry data maintained in collaboration with the national infrastructure SIMPLER.

> SIMPLER, the Swedish Infrastructure for Medical Population-Based Life-Course and Environmental Research, is a research infrastructure for making life in Sweden healthier.
>
> The infrastructure's mission is to provide national and international researchers with data that can be used in studies about how dietary and lifestyle factors, as well as genetics, affect our health, especially during the latter part of life. The knowledge can, for example, be used to develop recommendations on diet and lifestyle, to develop new markers for easier and more accurate diagnosis of chronic diseases and their precursors, and to develop new individualised treatments.

The popularity of Bianca has been clear, with over five hundred projects using the system since its inception. However, such extensive use has also put stress on the scalability of the system and some of the design assumptions. Each project that uses Bianca gets its own virtual HPC cluster which is set up and changed in size dynamically depending on usage. Sometimes, this process has been slower and less reliable than intended. UPPMAX is currently remedying this by upgrading the underlying virtualisation and cloud middleware. We are also tweaking the multifactor login process to be more in line with new standards and the login approaches used for accessing other SNIC resources.

Finally, Bianca has seen important hardware additions in the last year, and more are coming. The primary storage system, Castor, will soon be transitioned to the new system, Cygnus, with a larger total volume. UPPMAX has added four GPU nodes (two of which are financed by users) that are open to all projects on Bianca. We are also adding four new nodes, each with 768 GB of memory, to handle extremely demanding jobs. Finally, over fifty nodes with 256 GB of RAM are being added. This will also enable UPPMAX to make the login nodes in each virtual cluster larger, going from 8 GB to 16 GB in each of them. Having access to beefier login nodes has been a common request from our users. However, please note that even after these changes, heavy processing should still ideally be done in separate (possibly interactive) jobs.

With these developments, Bianca will stay relevant until new national resources for sensitive data are in place.

SNIC centre round-up

C3SE

C3SE runs the Alvis GPU-based system (available to Swedish researchers through SNIC) and a PC-cluster, Vera, which is for local users at Chalmers. C3SE is also part of the SNIC Science Cloud and Swestore.

HPC2N

HPC2N's main system, Kebnekaise, contains Intel Broadwell processors, NVIDIA GPUs and nodes with up to 3 TB of memory, all coupled with a fast InfiniBand network. HPC2N also operates large-scale disk and tape storage systems.

LUNARC

LUNARC is collaborating with SNIC on developing an interactive HPC environment for the new Dardel system. The initial focus will be on the interactive use of development environments, as well as access to numeric environments such as MATLAB and Jupyter Notebooks. In addition, the collaboration with the MAX IV laboratory is continuing; the aim is to improve workflows for other resources at both LUNARC and SNIC.

Together with SNIC, LUNARC recently held a workshop on Swestore, which was appreciated very much. A number of training videos were developed in connection with the workshop.

The new computing resources COSMOS and COSMOS-SENS have now been ordered and will replace LUNARC's current Aurora system at the end of 2022.

NSC

NSC has two flagship supercomputers for research: Tetralith and Berzelius, along with five smaller systems dedicated for specific research purposes. Tetralith, as a well-established system, is in the operational phase, and NSC is also developing other resources and conducting analyses of future computational, storage, data-processing and software needs. In particular, there is a strong focus on GPU-based computing, particularly for artificial intelligence/machine learning (AI/ML) and also in other research areas.

NSC's new resource, Berzelius, is sponsored by the Knut and Alice Wallenberg foundation. The system is a GPU-dense resource for Swedish academic research with a focus on AI. It is attracting a lot of attention from several user communities and is also being used to assist NSC to develop new services for future usage of SNIC resources.

PDC

The first phase of PDC's new HPE Cray EX system, Dardel, was installed last year and has been generally available for research since the start of this year. All SNIC users have now been migrated to Dardel, and various teams are making extensive use of the full system. With Dardel taking over workloads from older and smaller SNIC systems, PDC is also working on improving utilisation of the system for jobs that require only a small number of CPU cores. The added support of authentication using secure shell key-pairs has proved to significantly simplify access to the system.

Various extensions of the system are now in preparation, most notably the second phase, which is a GPU-accelerated partition. Together with AMD and HPE, PDC is actively supporting the preparation of applications for the AMD MI250x GPUs.

UPPMAX

UPPMAX runs the SNIC systems, Rackham and Bianca (which is part of SNIC SENS), plus the SciLifeLab system, Miarka, that just took over from Irma, and the UPPMAX Cloud, which is part of the SNIC Science Cloud.

HPC opportunities & organisation links

HPC in Sweden and Scandinavia

ENCCS: http://enccs.se

eSSENCE: http://essenceofescience.se

NeIC: http://neic.no

SeRC: https://e-science.se

SeSE: http://sese.nu

SNIC: https://snic.se

SNIC centre project participation

BioExcel CoE: https://bioexcel.eu

TREX: https://trex-coe.eu

EBRAINS: https://ebrains.eu

PerMedCoE: https://permedcoe.eu

EXCELLERAT: https://www.excellerat.eu

EOSC-Nordic: https://eosc-nordic.eu

DICE: https://www.dice-eosc.eu

HPC-Europa3: http://www.hpc-europa.eu

European HPC ecosystem

HPC in Europe: https://hpc-portal.eu

EuroHPC: https://eurohpc-ju.europa.eu

PRACE: https://www.prace-ri.eu

LUMI: https://www.lumi-supercomputer.eu

ETP4HPC: https://www.etp4hpc.eu

EOSC: https://eosc-portal.eu

General HPC news sources

HPCwire: http://www.hpcwire.com

insideHPC: https://insidehpc.com

Upcoming SNIC training & HPC events

SNIC zoom-in (Online interactive support and discussion forum)

9 June (also 15 September &13 October) 2022, 14:00-15:00, online

Researchers who are using (or who want to use) SNIC resources are invited to attend the SNIC zoom-ins where SNIC experts are available to discuss computing and storage needs or other matters related to research using SNIC resources: https://snic.se/support/zoom-in.

PDC Summer School

15-26 August 2022, Stockholm: https://www.pdc.kth.se/summer-school/2022

Save the dates

Here are preliminary dates for some upcoming SNIC training events. For details, visit the SNIC website https://snic.se/support/snic-training, or sign up in the SNIC User and Project Repository (SUPR) where you can register for our newsletter: https://supr.snic.se.

- UPPMAX Introductory Course: 23-26 August 2022; onsite, Uppsala
- Awk workshop: 29-30 August 2022 organised by UPPMAX;
- Introduction to HPC2N: mid-September 2022, online
- An introduction to the UNIX/LINUX command-line: 14 September 2022, 2h, online seminar
- CodeRefinery workshop: 20-22 & 27-29 September 2022, online