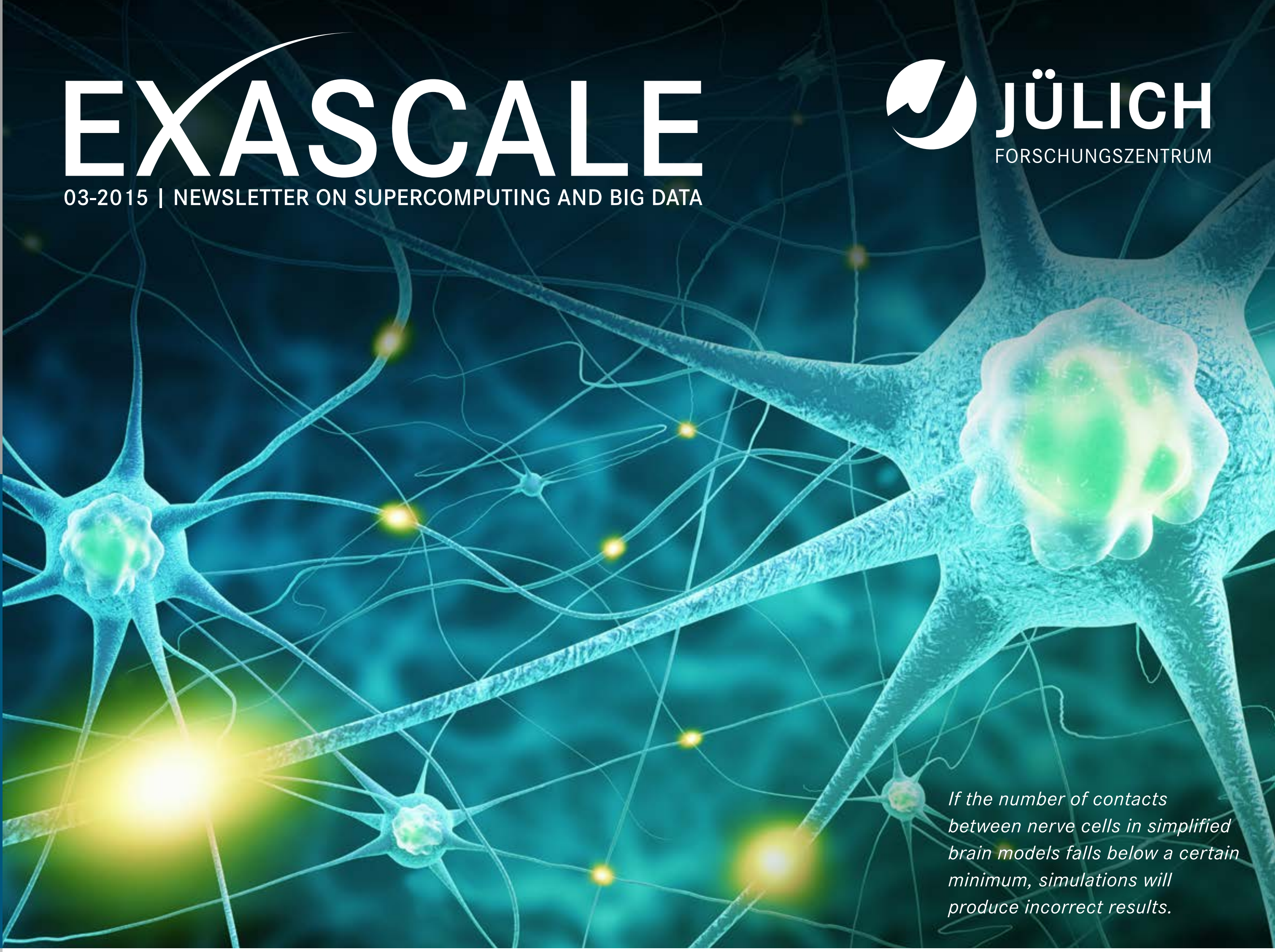
EXASCALE

03-2015 | NEWSLETTER ON SUPERCOMPUTING AND BIG DATA

JÜLICH FORSCHUNGSZENTRUM



If the number of contacts between nerve cells in simplified brain models falls below a certain minimum, simulations will produce incorrect results.

Member of the Helmholtz Association

The Limits of Simplification

Small models of the brain distort the relation between neurons

Our brain is a gigantic network with around 100 billion nerve cells. Each one of them possesses on average 10,000 contact points, known as synapses. It is through these that nerve cells exchange information in the form of electrical signals. These networks are so extensive that even the most powerful computers in the world can only simulate as little as one percent of the human brain. In order to draw conclusions on the entire brain in spite of this, neuroscientists work with simplified brain models with a reduced number of neurons and synapses, which in turn decreases memory usage for these simulations. But this downscaling has its drawbacks, as Jülich experts Dr. Sacha van Albada, Prof. Moritz Helias, and Prof. Markus Diesmann from the Institute of Neuroscience and Medicine have found out.

Minimum number of contacts

A person's nerve cells establish temporary relationships with each other for each action performed. These relationships, or correlations, can be of varying intensity – depending on how strongly involved an area of the brain is in laughing, walking, or thinking. "We have discovered that in simulations, the number of contacts between nerve cells may not be below a certain limit. Otherwise, the model will not correctly reflect the intensity of the correlations," says Diesmann. His team succeeded in mathematically determining the discrepancies and partially compensating for them. The accuracy of the calculated correction can, however, only be shown in a simulation of the relevant brain circuit in its full size. The technological prerequisites for this are to be created by the European Human Brain Project in collaboration with the Neural Simulation Technology Initiative (NEST) by 2022. Markus Diesmann and his team are involved in this endeavour together with the Jülich Supercomputing Centre.

PLOS Computational Biology,
DOI: 10.1371/journal.pcbi.1004490

STATEMENT

Prof. Markus Diesmann
Head of Computational and Systems Neuroscience

In order to examine the activity in brain circuits in detail, there is no alternative to simulating them in their full size. We are developing the simulation software for the necessary exascale computers and the evaluation software for big data analytics.

Getting to Grips with Big Data

Saving large volumes of data and making it available to users is one challenge. Generating knowledge rapidly and efficiently from these data is another. Scientists at the Jülich Supercomputing Centre (JSC) are seeking solutions to both.

In the field of data management, for example, they are concerned with a joint European data infrastructure. “There are hundreds of research infrastructures of various disciplines in Europe. They collect increasing amounts of data, and they all require the same basic services,” says Daniel Mallmann, head of the Federated Systems and Data division at JSC. In order to offer a series of joint services, the pan-European initiative EUDAT was brought into being. It enables scientists to save their data in infrastructures which can simultaneously be used to carry out compute-intensive data analyses. Researchers also have the option of giving other scientists access to their data. A metadata search simplifies locating data. JSC was involved in developing these services and provides one of the generic data centres of this collaborative pan-European infrastructure.

Optimize and learn

“Our aim is fast, efficient, and partly automated big data analysis,” explains Prof. Morris Riedel, head of a research unit on data analysis. He is utilizing parallel computing to adapt the analysis algorithms for use on supercomputers – and also benefits from the infrastructure services mentioned. For this purpose, Riedel and his team are optimizing parallel approaches to data analysis, and have adapted some of the algorithms for the first time to parallel computing. The JSC researchers are working on data-mining methods and algorithms to better determine decisive relations between data. At the same time, the data volume must be reduced to the actually relevant portion and deviations and irregularities corrected. In addition to multiple methods from statistics, an important role is played by scalable machine learning, an analysis tool capable of learning from examples from the relevant area of application. Protecting users’ privacy is also an important aspect for the scientists, for example in the case of patient data: “Specific methods of pseudonymization and anonymization could prevent certain data from being analysed,” says Morris Riedel.

JSC research topic
“Distributed Computing”

EU project EUDAT
(European Data Infrastructure)

More Data Does Not Mean Deeper Knowledge

Interview with Prof. Thomas Lippert, head of the Jülich Supercomputing Centre and spokesperson of the Helmholtz programme “Supercomputing & Big Data”

Prof. Lippert, “big data” is a much-debated buzzword. What’s it all about?

It is estimated that the data volume generated globally doubles every two years. Today’s databases and management and analysis tools are increasingly overwhelmed by this. We require new methods for processing and evaluating data.

This is where the focus is often on industry. What’s the situation in science?

Every day, more than 20,000 publications on various scientific topics are produced. Researchers create gigantic volumes of data, thanks to constantly improving experimental methods and increasingly powerful supercomputers. In order to gain new insights, science in particular relies more and more on the analysis of data sets.

What are the problems?

More data by no means automatically entails deeper knowledge. It is a matter of appropriately evaluating the data and interpreting the results. Recognizing connections, such as correlations, is crucial, but the main focus is on causal links between the data. Only then can we obtain useful and reliable information. But first, we need to create the appropriate basis.

Which is what?

In addition to the methods mentioned, we have to agree on standards by which we can usefully structure, use, and share data. In addition, we require a suitable infrastructure for more and more fields of science so that we can reconcile diverse data. Such structures have continuously been further developed in climate research, for example. We also depend on this for elaborate simulation calculations in brain research.

What role does the Helmholtz programme “Supercomputing & Big Data” play in this context?

As part of the programme, we want to provide national and European science with big-data infrastructures and instruments such as memory systems together with supercomputers of the highest performance class for data analysis at Jülich. We are also setting up simulation and data laboratories within the Helmholtz Association, where algorithms for data analysis and data-mining methods will be developed and optimized.

Prof. Lippert, thank you for the interview.

The All-Rounder

The supercomputer family at Forschungszentrum Jülich has a new member: JURECA

The new petaflop computer JURECA is highly versatile with regard to potential applications.

The petaflop-class computer achieves up to 2.2 quadrillion arithmetic operations per second and replaces JUROPA. Over the course of six years, JUROPA had provided numerous scientific results which were published in journals such as *Science* and *Nature*. JURECA will take it up from here. It serves more than 200 research groups as a working computer and is used for projects from various fields, including astrophysics, particle physics, life sciences, geoscience, and materials research.

JURECA is short for Jülich Research on Exascale Cluster Architectures and refers to the next generation of supercomputers, the exascale class. These will be a hundred to a thousand times more powerful than today's systems. But it is not only about performance: the new computer also uses only one megawatt of electric power – one third less than JUROPA – in spite of its tenfold computing power.

Experts from the Jülich Supercomputing Centre (JSC) were heavily involved in developing the system and designed it for a particularly wide range of tasks. "We do not want to break any records with JURECA. We are more concerned with providing a new instrument for those research fields that depend on the increased power for pioneering results. This is about big data, i.e. large volumes of data, which play an increasingly important role in many areas," says JSC director Prof. Thomas Lippert. JURECA complements Jülich's other supercomputer, JUQUEEN, one of the fastest in the world. While JUQUEEN was especially designed for parallel calculations on large numbers of processor cores, JURECA can be used for more general purposes.

JURECA is a cluster computer and comprises around 1900 individual computers connected to each other. Particularly for memory- or compute-intensive simulation calculations, there is additional equipment available. The system, which was delivered and installed in two phases by the company T-Platforms, required no special hardware developments. "As a consequence, we are able to provide computation time very cost-effectively and users can use free software without extensive customizations," explains Dr. Dorian Krause, who is responsible for the installation at JSC.

JURECA

Prototype for Next-Generation Supercomputers

JSC researcher Dr. Estela Suarez coordinated the EU project DEEP.

The EU research project DEEP (Dynamical Exascale Entry Platform) has developed an innovative computer architecture setting the course for future supercomputers. The system is based on the Cluster-Booster concept which involves accelerating a cluster of conventional processors using a Booster formed by massively parallel, interconnected multicore processors. The Cluster is responsible for handling parts of the applications which are not highly parallelized, while the highly parallelized code parts make use of the Booster. This task-sharing method increases the performance of the whole system.

The DEEP team has developed an entire software stack for this complex hardware system, which will simplify the programming process for users. They have also already optimized six applications from science and industry for their system. "We were able to demonstrate the advantages of this architecture very well: its high flexibility and efficient use of system resources," says Jülich scientist Dr. Estela Suarez, DEEP project coordinator.

The DEEP computer with its peak performance of 500 teraflop/s is installed at JSC as a prototype. In the future, not only DEEP partners will make use of the system: external developers will also have access to it for their applications. The EU project DEEP, which has now come to a close after four years, was carried out by 16 European institutions from research and industry.

EU project DEEP

Innovative: the DEEP system with Booster (right, with blue wiring) and Cluster (left, silver rack)

NEWS IN BRIEF

Building Blocks of Safety

As part of the BaSiGo project, researchers from the Jülich Supercomputing Centre (JSC) worked together with partners from industry, science, and the authorities for more than three years to investigate how safety at large-scale events can be improved. Their findings of the now concluded project contributed to the BaSiGo Guide, which contains numerous recommendations for planning and implementing such events. The large volumes of data collected by JSC's Civil Security and Traffic division during pedestrian experiments will now be made available for research.

more ...

Pilot Phase Begins

The realization of an interactive supercomputer for the Human Brain Project is not far off. In mid-2016, two consortia will put their pilot systems into operation at the Jülich Supercomputing Centre. The company Cray and a consortium comprising IBM and NVIDIA were both awarded contracts as part of pre-commercial procurement of research and development work. Interactive use is viewed as a key element for future supercomputers, designed specifically for applications in brain research.

more ...

Nanolattice to Replace Bone Parts

With the aid of Jülich's supercomputer JUQUEEN, scientists from the University of Siegen and TU Darmstadt are looking for efficient nanolattices. It is hoped that the lattices will replace conventional materials in automotive engineering and medicine. A minute structure made of metal nanorods could, for example, replace parts of a bone that a patient has lost due to an accident or a tumour. Both permanent and temporary applications are conceivable, the latter using the structure until bone and tissue have grown back. The initial objective is finding suitable structural materials, which must be light but strong. Researchers are using JUQUEEN for this purpose: they are simulating various lattices on the nanoscale with 50–200 million atoms. They are also hoping for new insights into how the metal nanorods must be organized geometrically.

more ...

Simulating Nanostructure Formation

The John von Neumann Institute for Computing (NIC) awarded the title of 2015 NIC Excellence Project to the research project entitled "Structure and dynamics of polymer and lipid systems". The team headed by Prof. Marcus Müller from Georg-August-Universität Göttingen is investigating how hundreds of molecules cooperate to self-organize into nanostructures. Such phenomena are crucial in technical and biological processes such as the manufacturing of microelectronic components or changes of membrane shapes during transport processes in cells. Computer simulations of coarse-grained, molecular models should make it possible to understand the mechanisms of structure formation as well as to improve materials properties and process conditions.

more ...

Now available for smartphones and tablets!

Exascale Newsletter: www.exascale-news.de
effzett – the crossmedia magazine: www.fz-juelich.de/effzett

PUBLICATION DETAILS

EXASCALE NEWSLETTER of Forschungszentrum Jülich **Published by:** Forschungszentrum Jülich GmbH | 52425 Jülich, Germany **Conception and editorial work:** Dr. Anne Rother (responsible according to German press law) Tobias Schlößer, Christian Hohlfeld **Graphics and layout:** Graphical Media, Forschungszentrum Jülich **Translation:** Language Services, Forschungszentrum Jülich **Photos:** p. 1 top: ©adimas/Fotolia.com; p. 2 left: ©Victoria/Fotolia.com; p. 4 Forschungszentrum Jülich/Ralf Eisenbach; Forschungszentrum Jülich **Contact:** Corporate Communications | Tel: +49 2461 61-4661 | Fax: +49 2461 61-4666 | Email: info@fz-juelich.de **Printed by:** Schloemer & Partner GmbH **Print run:** 300

UPCOMING EVENTS

Programming and Usage of the Supercomputer Resources at Jülich
26 – 27 November 2015
at Jülich Supercomputing Centre
Instructors: JSC employees and representatives of IBM, Intel, and ParTec
www.fz-juelich.de/ias/jsc/events/sc-nov

Advanced Parallel Programming with MPI and OpenMP
30 November – 1 December 2015
at Jülich Supercomputing Centre
Instructor: Dr. Rolf Rabenseifner, HLRS Stuttgart
www.fz-juelich.de/ias/jsc/events/mpi

4th JLESC Workshop
2 – 4 December 2015
at Gustav-Stresemann-Institut in Bonn
www.fz-juelich.de/ias/jsc/jlesc-4

Introduction to OpenGL
8 December 2015
at Jülich Supercomputing Centre
Instructor: Dr. Herwig Zilken, JSC
www.fz-juelich.de/ias/jsc/events/opengl

NIC-Symposium 2016
11 – 12 February 2016
at Forschungszentrum Jülich
www.john-von-neumann-institut.de/nic/nic-symposium-2016

You can find an overview of events at the Jülich Supercomputing Centre at:
www.fz-juelich.de/ias/jsc/events