COMPETENCE CENTER HIGH PERFORMANCE COMPUTING

DR. FRANZ-JOSEF PFREUNDT
HEAD OF DEPARTMENT

INNOVATION, DISRUPTION, AND A HOLISTIC APPROACH TO THE WORLD OF PARALLEL COMPUTING

The department has developed innovative world-class technologies for solving large data problems, specifically BeeGFS, Pre-Stack PRO and the Global Address Space Programming model (GPI) in addition to the Big Data framework called GPI-Space. In recent years, we have gained international attention by successfully combining these technologies with deep learning methods. At its core, it's always about the scalable automatic parallelization of big data problems. This is based on the concept of "Memory Driven Computing" which combines scalability and performance. In our machine learning projects, we are further developing industry-specific solutions based on this technology.

The aim of our involvement in EU-sponsored HPC research is to strengthen European technologies and improve the marketability of European HPC software products. In addition, our goal in co-design projects is to bring together microelectronics development and application development. We see a way to improve Europe's position in the fast-growing HPC/Big Data market in application-specific development of computer hardware.

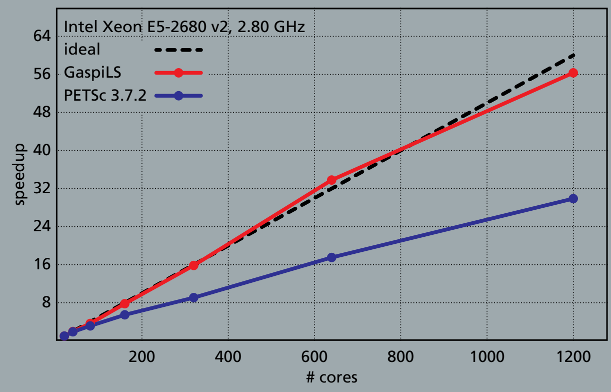
The energy systems of the future will consist of millions of distributed Internet-of-Things (IoT) computers. These optimize the self-consumption of PV power, regulate the creation of local grids, control large and small power storage systems, and coordinate the energy flow in power grids. Our technologies and project solutions are developed to master this world of distributed computing. Our commitment is to create intelligent solutions that work to advance the energy transition.

Contact
franz-josef.pfreundt@itwm.fraunhofer.de
www.itwm.fraunhofer.de/en/hpc

MAIN TOPICS

- Scalable parallel programming
- Deep learning tools and applications
- Memory driven computing (GPI-Space)
- BeeGFS – parallel cluster file system
- Green by IT
- HPC in seismic

1



2

GaspiLS AND THE GPI-2 ECOSYSTEM: GPI-2 SCALABILITY AND PERFORMANCE MADE EASY

1 *GaspiLS provides scalability for FEM and CFD simulations*

2 *Performance plots: Scalability advantage of GaspiLS compared to PetSC. Jacobi preconditioned Richardson Method; 3-D Poisson equation (2d order FD discretization), cubic matrix (359^3)*

The distributed systems used in high performance computing require highly efficient and scalable applications. Scalability is a measure of the efficiency of a parallel implementation and ultimately indicates whether the available resources – for example, CPUs – are being efficiently used. The Competence Center High Performance Computing develops GPI-2, a parallel programming model that is ideal for implementing such applications.

GaspiLS is a library of scalable, iterative linear solvers developed to easily exploit the benefits generated by GPI-2 and make them available for immediate practical use in a multitude of applications. GaspiLS is ready for direct use with a variety of new or existing simulation programs ultimately solving linear systems.

HySCALA explores new areas of application and new markets for GaspiLS

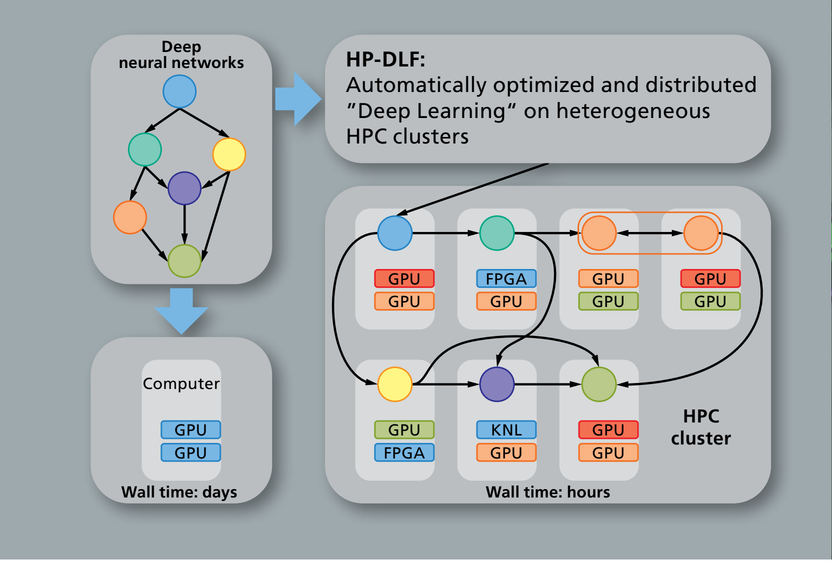
GaspiLS has already proven itself in several industry projects. Its further distribution is currently being promoted as part of the EU project HySCALA (Hybrid Scalable sparse matrix linear algebra for industrial applications).

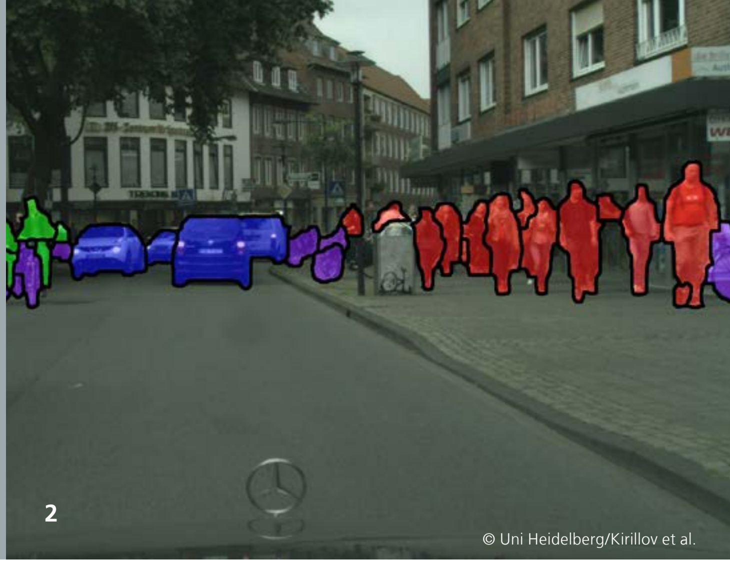
The aim is to analyze various, potential new market segments and fields of application for GaspiLS and to identify specific requirements for a competitive linear solver library. We are looking primarily for generic yet efficient preconditioners that allow us to reduce the number of iterations required for convergence of the iterative process and minimize the total run times. Presently, the focus is on the scalable implementation of efficient preconditioners that can be applied to a broad class of problems.

GaspiCxx for increased productivity

Within GaspiLS, we have factorized the implementation for the explicit management of communication resources required by the GPI-2 data transfer and used GaspiCxx to supply it to other applications. GaspiCxx defines an easy to use C++ interface. It delivers the full native GPI-2 performance. At the same time, the management of GPI-2 communication resources is fully transparent to the application.

This eliminates a large part of the implementation work normally required to develop GPI-2 applications. Development of GPI-2 applications and the exploitation of the advantages – like the good scalability – has never been so easy.





HPC FOR MACHINE LEARNING: HIGH PERFORMANCE DEEP LEARNING FRAMEWORK

Artificial neural networks have become established in many areas of machine learning in recent years. For example, they are at the leading edge of computer vision, speech and character recognition as well as machine translation. One reason for their success is the ability to create highly complex interrelationships between the raw data input and the classification (the labels) of the output data.

This often requires several million free parameters that have to be changed (i. e., learned) while training the network. Because of the large number of these so called weights, training a single neural network often takes several days or even weeks. Clearly, making these algorithms highly scalable through the use of supercomputers is highly desirable. In the ideal case, doubling the number of computers connected in parallel would halve the running time of the algorithm.

Small neural networks or fewer files?

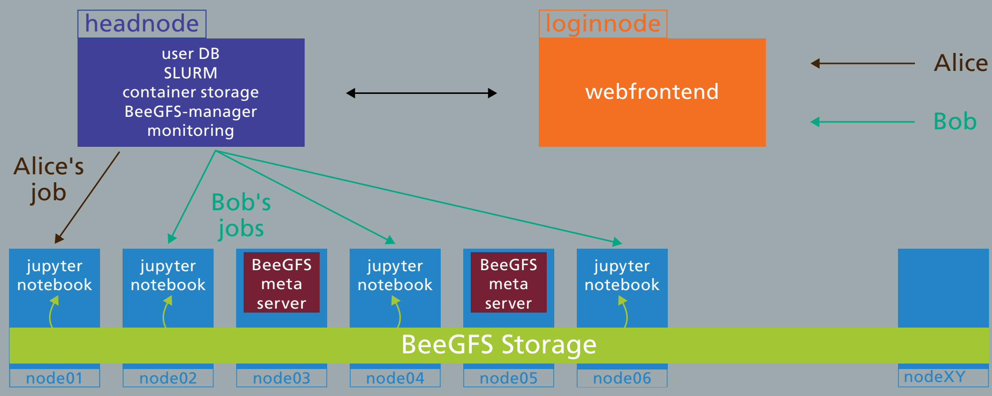
Neural networks encounter an additional problem: they require a very large main memory. As a result, only relatively small neural networks can be trained on a single computer, or even the amount of data used for learning must be limited. Neither of these options is desirable because they reduce the capacity, i. e., the learning ability of the network. Rather, it is more desirable to train networks of twice the size with twice the number of computers. This is called "weak scalability" in the jargon of parallel computing.

High degree of scalability with GPI Space

Enabling both weak as well as strong scalability in the training of neural networks is the subject of the BMBF project "High Performance Deep Learning Framework, (HP-DLF)." A particular focus is placed on enabling the construction of neural networks of any size and ensuring easy access to existing and future high performance computing systems. No prior knowledge of parallel computing is required on the part of the user. Our in-house runtime system GPI-Space manages everything. When represented in the form of a special graph, a so-called Petri net, algorithms can be automatically and dynamically parallelized.

1 *HPC enables deep learning without storage limits.*

2 *Large amounts of data play a special role in autonomous driving.*



1

HPC FOR MACHINE LEARNING: CARME

1 *Simplified scheme of the most important system components and their connections*

Machine learning has an increasingly higher priority in both scientific and industrial enterprises. This is evident from the investment in new, above all, GPU-based hardware – from simple desktop computers to high performance computing clusters. Computing clusters are used in Data Analysis (DA) and highly complex Machine Learning (ML) systems to process and simulate very large amounts of data – to include even the human brain.

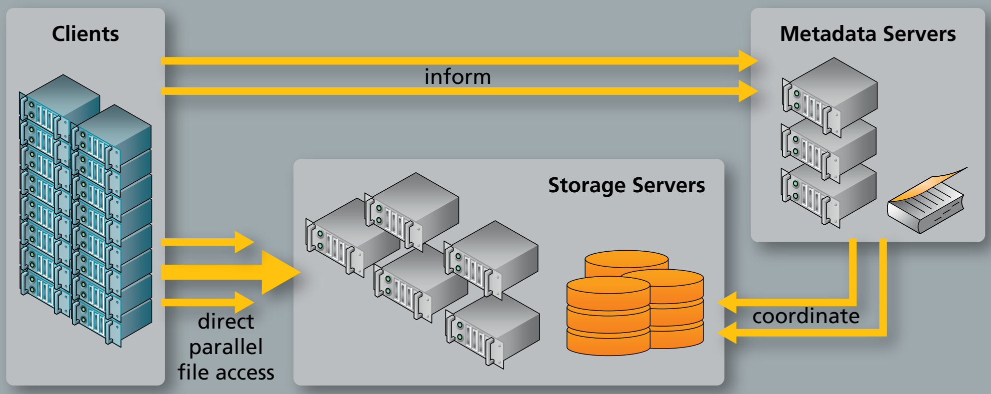
Machine learning in HPC clusters presents certain challenges. The procurement of the individual hardware components is the least of these challenges. The biggest questions arise subsequent to that acquisition:

- How to manage existing resources?
- How to make an application scalable to several GPUs?
- How to solve the challenge of data storage and continuous upload to the program?
- How to train users to effectively use the hardware?

The answers to these questions begin with our open-source software stack Carme. The basic concept is to combine the world of machine learning and data analysis with the world of HPC systems. We achieve this using established ML and DA tools with HPC back ends. Specifically, we use a variety of HPC and ML technologies. Some of these technologies are developed in this department, for example, the highly reliable parallel file system BeeGFS for fast data links.

Carme combines the worlds of machine learning and HPC clusters. ML is a steady and fast growing field of technology. This new agility challenges data centers to provide very different applications for single users. It is not enough to have one user interface for the user; rather there must also be a guarantee of a seamless integration of this surface in existing and emerging clusters. To make clusters attractive to ML and DA users, an intuitive software environment must be provided to the clusters. Interactive management of the cluster is essential in the development of ML applications. Users must have the opportunity to use tools they are familiar with on a complex HPC cluster, making it easier for them to migrate to and use the cluster.

www.open-carme.org



1

BeeGFS – THE FILE SYSTEM FOR BIG DATA AND AI

The success of current AI technologies such as neural networks is based on the increased power of today's processors – mostly GPU´s – but, above all, on the availability of very large amounts of data. For example, new medical devices, autonomous vehicles, and genome analyses supply ever more fine resolution data in quick succession forming the basis for future AI solutions. Developed at ITWM and distributed by ThinkparQ, the parallel file system BeeGFS (also known as Fraunhofer Parallel File System – FhGFS) helps in mastering the large data volumes with a very flexible software solution.

1 *BeeGFS architekture*

BeeGFS is a parallel file system where storage capacity as well as read and write speeds grow linearly with the number of linked storage units. As a pure software solution, it can be flexibly installed both on existing hardware and on the latest, superfast flash-memory systems. In addition to very good scalability, our system development team attaches great importance on easy handling and a high degree of flexibility for a variety of potential use cases.

BeeGFS on NVMe

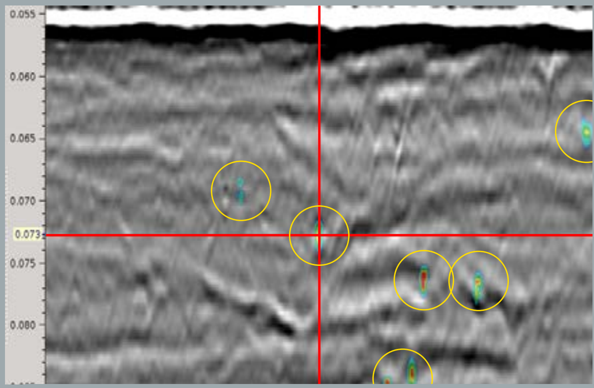
Training deep neural networks (Deep Learning) demands that existing data be provided several times very quickly to the computing units. Most external storage systems are hardly suitable for this task, so the data is cached directly to the computer servers on fast local systems (NVMe). Since these have relatively small capacities, the need arises for data to be distributed on several units in parallel.

The BeeGFS software system is specially optimized for high speed requirements even with a large number of files and this ability is its biggest strength. BeeGFS can be installed directly on the computer servers and is scalable to high I/O rate of 1 TByte/sec and more. Japanese AI researchers were convinced: BeeGFS is now successfully deployed on the two major Japanese AI systems TSUBAME 3.0 (HPE) and AI Bridging Cloud Infrastructure (ABCI, Fujitsu).

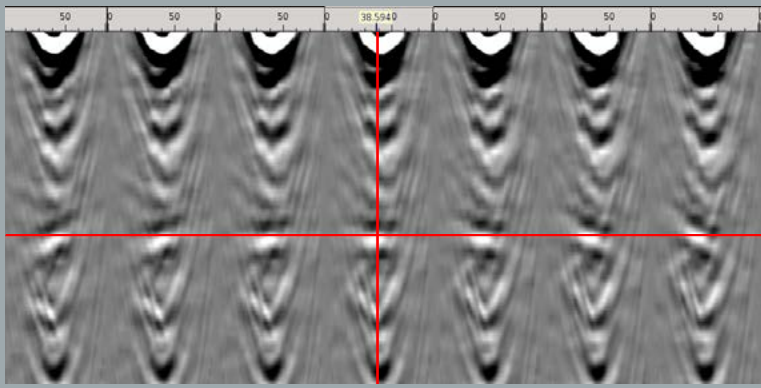
Open-source license

The software is distributed with an open-source license and source files are provided on the BeeGFS website. A spin-off of Fraunhofer ITWM, ThinkparQ, supplies worldwide commercial support for BeeGFS and manages further development from a customer perspective. The joint development team also successfully applies its extensive knowledge in several EU funded projects that focus on the use of BeeGFS on future Exascale computing systems.

BeeGFS®
developed by Fraunhofer



1



2

BOULDER-DETECTION MIT MACHINE LEARNING

1 *Subsurface image computed from seismic data; colour spots show locations of high probablity for the presence of boulders*

2 *Illumination representation of part of the subsurface at which a boulder was identified; the cross-hair points into the symmetry center of the identified pattern and indicates the lateral position of the boulder at 73 m depth*

Piles of windmills of offshore windparks must be firmly coupled into the subsurface layers deep below the sea-floor. Large boulders, that would be obstacles for the planting process of such piles, must be identified in the phase of defining the exact locations of windmills. Seismic data sets acquired for site-planning of windparks for the purpose of assessing subsurface stability conditions have limited frequency content. We developed a new methodology for object identification below seismic wavelength (here: 1 m) from such data by means of machine learning methods.

Key to this methodology is the preprocessing of the data by a prestack migration method that highlights the weak-amplitude diffractions contained in the seismic data.

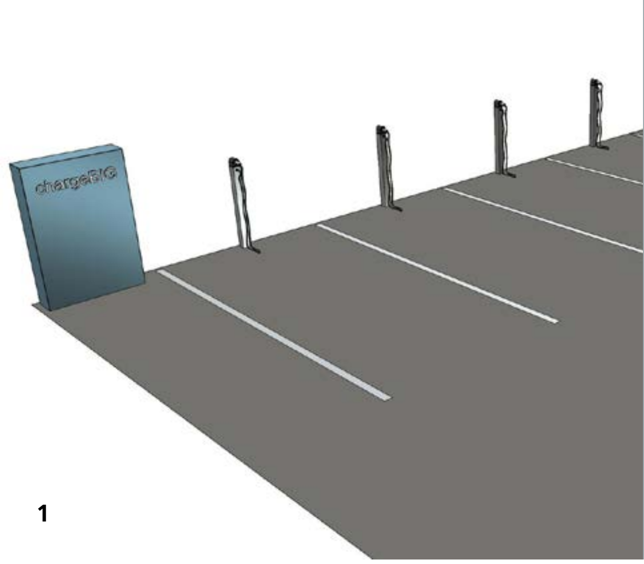
Pattern recognition in seismic sections

Together with colleagues from Fraunhofer IWES and sponsored by BMWI we developed a process that maps the seismic data into a domain in which diffraction responses show a typical pattern. The task of finding such patterns, thus to localize the associated diffracting objects, is similar to the task of assigning pixels of photos to object classes, which constitutes a problem that can successfully be solved with the help of deep neural networks (DNNs).

Here, however, we are dealing with a high-dimensional problem, as even for 2D seismic, i. e. 2D subsurface images, illumination-direction and velocity-variation contribute two additional dimensions. Further, the Earth's subsurface is not accessible so that the networks cannot be trained with the ground truth for real data; rather, we have to rely on training solely based on synthetic data generated and perturbed towards the appearance of real data.

Benefit for the user: reduced amount of data

Our results demonstrate that transfer learning from synthetic to real data works and that our DNNs that consist of a large number of convolutional layers offer the necessary complexity for computing the probability for the existence of diffracting objects of 1m scale length even in noisy data. After application of this automated process, the user is left with the task to further interpret the seismic data only in those areas that are marked by high probablility values.



1

2

chargeBIG – A NEW CHARGING INFRASTRUCTURE FOR ELECTRIC DRIVE VEHICLES

The chargeBIG project is a joint development effort with MAHLE Group and the Eliso company to create a new kind of charging infrastructure for electric drive vehicles. The resulting system is to be cost efficient, highly scalable, and will contribute to stable grid operations. The goal is the large scale electrification of parking garages by placing a charging option at every parking space of a parking garage as cheaply as possible.

Conventional technologies are very expensive and car park operators often decide against complete electrification. Instead, they install charging stations for electric cars only at a few designated spaces; with the disadvantage of finding that the charging station is often blocked by already fully charged vehicles.

Charging infrastructure for all parking spaces

The vision of chargeBIG: In place of expensive components at a few parking spaces, the necessary technical components are combined at a central location and just one tower with a charging cable is needed for each parking space. In effect, the requirements at each parking space are reduced to a minimum. The centralized concept provides substantial savings, both in manufacturing costs as well as current maintenance expenses, with the benefit of low-cost electrification of as many parking spaces as possible.

A parking garage as an actual lab

In joint development with ITWM, MAHLE has already developed a chargeBIG prototype with 18 charging points. This serves as a blueprint for a demonstrator with 108 charging points to be installed in a MAHLE parking garage in Stuttgart. The garage is a real laboratory for testing the effectiveness of grid operations. In addition to the charging infrastructure, we installed a storage battery system, a DC-DC fast charging station (i. e., a station for charging an e-vehicle directly with DC current from a stationary battery), and a dedicated photovoltaic system.

The Amperix energy management system developed by Green by IT Group optimizes component usage with a view towards increasing local self-sufficiency, reducing peak loads (peak shaving), and the adoption of flexible electricity prices. The chargeBIG project is funded by the Immediate Action Program "Saubere Luft" of the Federal Ministry for Economic Affairs and Energy (BMWi). As part of the accompanying scientific research, ITWM is also analyzing and evaluating the project's contribution to the reduction of nitric oxide (NOX) emissions in the city of Stuttgart.

1 *Rendering of the chargeBIG charging infrastructure; the charging towers are connected to the chargeBIG central.*

2 *Left: Charging towers with charging plugs, right: Prototype of the chargeBIG central for 18 charging points*

NEWS

EPEEC: EUROPEAN JOINT EFFORT TOWARDS A HIGHLY PRODUCTIVE PROGRAMMING ENVIRONMENT

The aim of the project launched in October is to develop a parallel programming environment for heterogeneous supercomputers. We are adding application-specific data compression algorithms to GPI, the parallel programming model, in order to better support machine learning and loosen the requirements for consistency. Today's programs rely on a consistent view of global attributes, which presents a limiting factor for scalability and requires synchronization between parallel computer nodes. GPI enables EPEEC to meet these challenges in achieving exascale computing.

AMPERIX® SUPPORTS ENERGY SELF-SUFFICIENCY

In the Schoonschip community, a water estate north of Amsterdam, 30 houses form an energy unit/microgrid that generates most of its electricity itself using solar energy and stores the energy itself using heat pumps and batteries. The houses are interconnected with each other, but also have a shared connection to the municipal power grid. The power supply inside of the microgrid is coordinated by our Amperix® energy management system. In addition to controlling the battery storages, we also implement the sector coupling with controlling the heat pumps.

AUTOMATED DATA ANALYSIS USING MACHINE LEARNING: DEEP TOPOLOGY LEARNING (DETOL)

Although Deep Learning methods are established in many sectors, intensive training is still required. Artificial neural networks demand huge amounts of data and enormous processing power to complete this task. The aim of the DeToL project, launched last June, is to substantially simplify and accelerate the design process of Deep-Learning solutions by means of automated, data-driven design algorithms.

One of the core aspects in DeToL is the development of the central software framework for distributed parallelization of the massive compute load. Our distributed runtime system GPI-Space forms the basis on which the individual modules are implemented. Here we focus on the algorithmic realization of the parallelization of the partial modules for reinforcement learning, genetic algorithms, graph embeddings, early stopping, meta learning and pruning. (See page 91)

Front, left to right: Sabine Müller, Kalun Ho, Dr. Somnath Madzumdar, Dr. Rui Machado, Dr. Alexandra Carpen-Amarie, Avraam Chatzimichailidis, Dr. Abel Amirbekyan, Dr. Dimitar Stoyanov, Peter Michael Habelitz, Dr. Tiberiu Rotaru, Valentin Tschannen Dr. Matthias Balzer, Sebastian Schumb, Kai Krüger, Frauke Santacruz, Dr. Norman Ettrich, Matthias Klein, Javad Fadaieghotbi, Bernd Lörwald, Dominik Loroch, Mikita Vedzeneyeu, Dr. Valeria Bartsch, Dr. Franz-Josef Pfreundt, Dr. Mirko Rahn, Delger Lhamsuren, Dr. Alexander Janot, Bernd Lietzow, Dr. Martin Kühn, Ricard Durall Lopez, Raju Ram, Lukas Ristau, Dr. Peter Labus, Christian Mohrbacher, Matthias Deller, Dr. Dirk Merten, Dr. Dominik Straßel, Julius Roob, Dr. Roman Iakymchuk, Dr. Alexander Klauer, Philipp Reusch, Dr. Janis Keuper, Patrick Reh