CSCS
Centro Svizzero di Calcolo Scientifico
Swiss National Supercomputing Centre

ETH zürich

# Accelerated Data Analytics and Computing (ADAC) Workshop 6

*June 20 - 21, 2018*

ETH Zurich
Rämistrasse 101
8092 Zurich
Switzerland

Via Trevano 131
6900 Lugano
Switzerland

T +41 91 610 82 11
F +41 91 610 82 82
www.cscs.ch

CSCS

# Wednesday, June 20, 9:00 – 18:00

All plenary meetings will take place in room HG D 3.2

| | | |
|---|---|---|
| 8:30 | **Registration** (next to HG D 3.2) | |
| 9:00 | **Welcome & Opening** | Thomas Schulthess, ETH Zurich/CSCS |
| 9:10 | **Demystifying Parallel and Distributed Deep Learning ~ An In-Depth Concurrency Analysis** | Torsten Hoefler, ETH Zurich |
| 10:00 | **Emerging Technologies: Selected results from European Projects** | Dirk Pleiter, Jülich Supercomputing Centre |
| 10:30 | **Break** | |
| 11:00 | **"Computational Memory": A Stepping Stone to Non-von Neumann Computing** | Evangelos Eleftheriou, IBM Zurich Research Laboratory |
| 11:30 | **Mixed-signal neuromorphic VLSI devices for spiking neural network** | Ning Qiao, Institute of Neuroinformatics, University of Zurich & ETH Zurich |
| 12:00 | **Lunch** | |
| 13:00 | **OpenMP Loop Scheduling Revisited: Making a Case for More Schedules** | Florina Ciorba, University of Basel |
| 13:30 | **FPGAs and Beyond: Specialized Hardware Architectures for HPC with High-Level Synthesis** | Johannes de Fine Licht, ETH Zurich |
| 14:00 | **Learning how to forget** | Oliver Fuhrer, MeteoSwiss |
| 14:30 | **Non-orthogonal Configuration Interaction for the Calculation of Electronic Couplings** | Remco W.A. Havenith, University of Groningen |
| 15:00 | **Break** | |
| 15:30 | **The Fast Multipole Method in molecular dynamics** | Berk Hess, Royal Institute of Technology, Stockholm |
| 16:00 | **Appropriate Use of Containers in HPC** | Bill Sparks, Cray |
| 16:30 | **High Performance Containers and the Convergence of Big Data and HPC** | Christian Kniep, Docker |
| 17:00 | **Characterizing Faults, Errors, and Failures in Extreme-Scale Systems** | Christian Engelmann, Oak Ridge National Laboratory |
| 17:30 | **Current Status of TSUBAME3.0 Operation** | Toshio Endo, Tokyo Tech |
| 18:30 | **Banquet – Sorell Hotel Züricichberg** | |

# Thursday, June 21, 9:00 – 17:00

All plenary meetings will take place in room HG D 3.2

| | | |
|---|---|---|
| 9:00 | **Day 2 Welcome and Logistics** | Thomas Schulthess, ETH/CSCS |
| 9:05 | **Update on Working Group Activities** – Room HG D 3.2<br>• Applications<br>• Emerging Technologies<br>• Resource Management | |
| 9:30 | **Working Group Goals and Expectations** | Satoshi Matsuoka, Tokyo Tech<br>Jeff Nichols, ORNL<br>Thomas Schulthess, ETH/CSCS |
| 10:00 | **Break** | |
| 10:30 | **Breakout Groups #1**<br>• Applications – Room HG D 3.1<br>• Emerging Technologies & Resource Management – Room HG D 3.3<br>• Governance – Room HG E 33.5 | |
| 12:00 | **Working Lunch** | |
| 13:30 | **Breakout Groups #2**<br>• Applications – Room HG D 3.1<br>• Emerging Technologies & Resource Management – Room HG D 3.3<br>• Governance – Room HG E 33.5 | |
| 15:00 | **Break** | |
| 15:15 | **Breakout Groups #3**<br>• Applications – Room HG D 3.1<br>• Emerging Technologies & Resource Management – Room HG D 3.3<br>• Governance – Room HG E 33.5 | |
| 16:15 | **Final Report Out from Working Groups** – Room HG D 3.2 | |
| 17:00 | **Adjourn Day 2** | |
| 17:30 | **Tour at ETH Zurich** | |
| 18:30 | **Banquet – Restaurant Mère Catherine** | |

# ABSTRACTS OF TALKS JUNE 20

**9:10 Demystifying Parallel and Distributed Deep Learning: An In-Depth Concurrency Analysis**
Torsten Hoefler, ETH Zurich

Deep Neural Networks (DNNs) are becoming an important tool in modern computing applications. Accelerating their training is a major challenge and techniques range from distributed algorithms to low-level circuit design. In this talk, we describe the problem from a theoretical perspective, followed by approaches for its parallelization. Specifically, we present trends in DNN architectures and the resulting implications on parallelization strategies. We discuss the different types of concurrency in DNNs, synchronous and asynchronous stochastic gradient descent, distributed system architectures, communication schemes, and performance modeling. Based on these approaches, we extrapolate potential directions for parallelism in deep learning.

**10:00 Emerging Technologies: Selected results from European Projects**
Dirk Pleiter, Jülich Supercomputing Centre

The European Union has in recent years provided significant funding to realize a R&D agenda, which is regularly updated by the European Technology Platform for HPC (ETP4HPC). In this talk we will, after a short presentation of the current R&D agenda, present details from a selected set of ongoing as well as emerging projects. We will cover projects focusing on processor technologies, I/O architectures, and system hardware and system architectures.

**11:00 "Computational Memory": A Stepping Stone to Non-von Neumann Computing**
Evangelos Eleftheriou, IBM Research Laboratory – Zurich

In today's computing systems based on the conventional von Neumann architecture, there are distinct memory and processing units. Performing computations results in a significant amount of data being moved back and forth between the physically separated memory and processing units. This costs time and energy, and constitutes an inherent performance bottleneck. Thus, it is becoming increasingly clear that for application areas such as cognitive computing, we need to transition to computing architectures in which memory and logic coexist in some form.
Most recently, post-silicon nanoelectronic devices with memristive properties are also finding applications beyond the realm of memory. A critical requirement in these novel computing paradigms is a very-high-density, low-power, variable-state, programmable and non-volatile nanoscale memory device. Phase-change-memory (PCM) devices based on chalcogenide phase-change materials, such as Ge2Sb2Te5, are well suited to address this need owing to their multi-level storage capability, state dynamics and potential scalability.
The computational memory or in-memory computing is a non-von Neumann approach, in which the physics of nanoscale PCM devices as well as their organization in crossbar arrays are exploited to perform certain computational tasks within the memory unit. I will present large-scale experimental demonstrations using about one million PCM devices organized to perform high-level computational primitives, such as compressed sensing, linear solvers and temporal correlation detection. Moreover, I will discuss the efficacy of this approach to efficiently address the training and inference of deep neural networks. The results show that this co-existence of computation and storage at the nanometer scale could be the enabler for new, ultra-dense, low-power, and massively parallel computing systems.

**11:30 Mixed-signal neuromorphic VLSI devices for spiking neural network**
Ning Qiao, Institute of Neuroinformatics, University of Zurich and ETH Zurich

Neuromorphic computing platforms represent a new generation of non von Neumann massively parallel architectures that are ideally suited to implementing spiking neural networks for real- time sensory processing. In this presentation, I am going to show neuromorphic architectures and mixed-mode analog/digital VLSI circuits/systems we have been developing in recent years. These hardware devices promise to reduce power consumption by several orders of magnitude and have the potential to solve the von Neumann memory bottleneck problem thanks to their co-localized memory and computing features.

**13:00 OpenMP Loop Scheduling Revisited: Making a Case for More Schedules**
Florina M. Ciorba, University of Basel

In light of continued advances in loop scheduling, this talk revisits OpenMP loop scheduling by reviewing the current state-of-the-art in loop scheduling and presenting evidence that the existing OpenMP schedules are insufficient for all combinations of applications, systems, and their characteristics. The talk reviews the state-of-the-art which reveals that no single loop scheduling technique can be a one-fits-all solution and effectively optimize the performance of all parallel applications in all situations. The irregularity in computational workloads and hardware systems, including operating system noise, results in performance loss and their impact has often been neglected in loop scheduling research, in particular the context of OpenMP schedulers. Existing dynamic loop self-scheduling techniques, such as trapezoid self-scheduling, factoring and weighted factoring, offer an unexplored potential to this situation due to the fact that they explicitly target the minimization of load imbalance and scheduling overhead. In this talk we show, through theoretical and experimental evaluation, that these loop self-scheduling methods provide a benefit in the context of OpenMP. We conclude that OpenMP must include more schedules to offer a broader performance coverage of applications executing on an increasing variety of heterogeneous shared memory computing platforms, including manycore and accelerator architectures.

**13:30 FPGAs and Beyond: Specialized Hardware Architectures for HPC with High-Level Synthesis**
Johannes de Fine Licht, ETH Zurich

After the end of Dennard scaling, improving energy efficiency of computations has had to rely on tweaks to existing architectures. While GPUs have improved the situation by increasing the ratio of computations to control overhead, all fixed architectures suffer from spending the vast majority of their energy in increasingly large register files, caches, and instruction decoding. All these overheads can be avoided with specialized hardware architectures. FPGAs, CGRAs, and ASICs are three approaches that have been employed with varying amounts of success to achieve this. In this talk, we will discuss the impact of high-level synthesis (HLS) from languages such as C/C++ and OpenCL on the feasibility of implementing efficient custom hardware architectures for HPC. How do we program for performance in this paradigm? Is a high-level language enough to capture the benefits of specialized architectures? We will show optimization principles for hardware, and results of designing high-throughput FPGA kernels with HLS, providing some insight to these questions and give a glimpse into what could be the future of HPC.

**14:00 Learning how to forget**
Oliver Fuhrer, MeteoSwiss

The weather and climate community has set ambitious goals to reach global km-scale modeling capability on future exascale high-performance computing (HPC) systems. But currently, state-of-the-art models are executed using much coarser grid spacing and almost none of the productive weather and climate models is capable of exploiting modern HPC architecture with hybrid node designs. In this talk we present our experience from an investment into the COSMO model, to render it performance portable and capable of targeting multiple hardware architectures. Two approaches have been followed. On the one hand, a complete rewrite of the code using a domain-specific language has been used for the dynamical core. On the other hand, a more traditional porting approach has been used adding compiler directives to the Fortran base code. While the former approach hides many of the hardware dependent details from the domain-scientist, the latter adds a significant amount of hardware dependent code to an already large code base. The advantages and disadvantages of these two approaches both in the development and the maintenance phase of COSMO are discussed. Finally, we present a baseline simulation of what can be achieved using COSMO on Piz Daint when scaling it to the full system. The results can serve as a baseline of how far we are from conducting global, km-scale resolution simulations for weather and climate. We conclude by highlighting some of the remaining challenges and potential solutions on the way to global km-scale climate simulations.

**14:30 Non-orthogonal Configuration Interaction for the Calculation of Electronic Couplings**
Remco W.A. Havenith, University of Groningen

Co-Authors: Remco W.A. Havenith, University of Groningen and Ghent University, Tjerk P. Straatsma, Oak Ridge National Laboratory, Shirin Faraji, University of Groningen, and Ria Broer, University of Groningen

In singlet fission, the energy of a molecule in one of its excited states is split over two molecules, each then being in their first triplet state (Figure 1) [1]. This process can adequately be described in ensembles of molecules in terms of molecular states. For this purpose, we have developed a non-orthogonal configuration interaction method that allows the description of a molecular crystal, using the (embedded) cluster approach, in terms of many-electron basis functions (MEBFs), each describing a particular electronic state of a molecule in the ensemble. The MEBFs can be constructed as combinations of antisymmetrised products of molecular wavefunctions of the MCSCF-type [2,3]. The advantage of this approach is that we can calculate the diabatic excited states in the ensemble, and the coupling between excited states localised on different molecules, such as the electronic coupling matrix element relevant for singet fission, [1]. In this presentation, an overview of the method will be given, and we show different applications of this method for the calculation of the relevant matrix elements for singlet fission in solids like tetracene and cibalackrot.

[1] M.B. Smith, and J. Michl, Chem. Rev. 2010, 110, 6891. [2] M. Wibowo, R. Broer, and R.W.A. Havenith, Comput. Theor. Chem. 2017, 1116, 190. [3] R.W.A. Havenith, H.D. de Gier, and R. Broer, Mol. Phys. 2012, 110, 2445.

**15:30 The Fast Multipole Method in molecular dynamics**
Berk Hess, Royal Institute of Technology, Stockholm

Parallel scaling of molecular dynamics simulations is limited by the communication in the 3D-FFT used by particle-mesh methods for electrostatics in all molecular simulation packages.

All other communication is local, so use of a better scaling electrostatics method can improve scaling by more than an order of magnitude. Two candidates are the Fast Multipole Method (FMM) and multi-level summation, both are order(#particles) in computation and communication. Here I will discuss how FMM can be integrated in molecular dynamics and what computational benefits this brings. To enable efficient use of FMM, the well known energy conservation issue must be resolved. I will show that energy conservation can be achieved by using regularization of the FMM.

**16:00 Appropriate Use of Containers in HPC**
Jonathan "Bill" Sparks, Cray

Linux containers in the commercial world are changing the landscape for application development and deployments. Container technologies are also making inroads into HPC environments, as exemplified by NERSC's Shifter and LBL's Singularity. While the first generation of HPC containers offers some of the same benefits as the existing open container frameworks, like rkt or Docker, they do not address the cloud/commercial feature sets such as virtualized networks, full isolation, and orchestration. This talk will discuss container technologies and the appropriate use of containers in HPC, along with best practices and guiding principles.

**16:30 High Performance Containers and the Convergence of Big Data and HPC**
Christian Kniep, Docker

This session will briefly introduce container technology and outline the challenges when trying to fit those workloads into containers. Afterwards the community solutions are touched on before an approach based on proper Docker is shown. The talk will wrap-up with an outlook how containers can foster scientific discoveries by allowing HPC to be used by everyone and how Big Data and HPC converge.

**17:00 Characterizing Faults, Errors, and Failures in Extreme-Scale Systems**
Christian Engelmann, Oak Ridge National Laboratory

Building a reliable supercomputer that achieves the expected performance within a given cost budget and providing efficiency and correctness during operation in the presence of faults, errors, and failures requires a full understanding of the resilience problem. The Catalog project develops a fault taxonomy, catalog and models that capture the observed and inferred conditions in current supercomputers and extrapolates this knowledge to future-generation systems. To date, the Catalog project has analyzed billions of node hours of system logs from supercomputers at Oak Ridge National Laboratory and Argonne National Laboratory. This talk provides an overview of our findings and lessons learned.

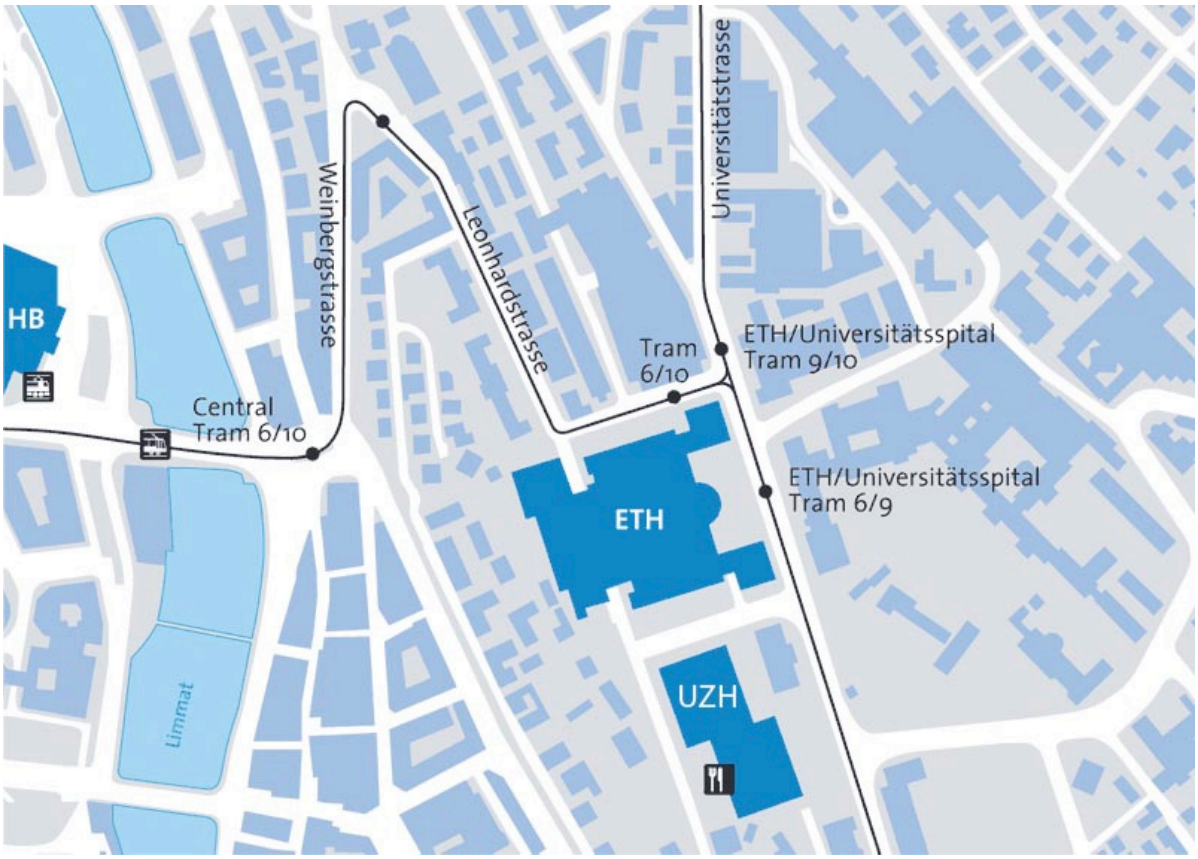
**17:30 Current Status of TSUBAME3.0 Operation**
Toshio Endo, Tokyo Tech

The current status of Tokyo Tech TSUBAME3.0 operation, which has been started in August 2017, is presented. The system consists of 540 computing nodes, each of which has 2 Broadwell Xeon CPUs and 4 Tesla P100 GPUs. Each of these "fat" nodes are dynamically partitioned corresponding to users requests. We also present improved advanced node reservation and accounting with TSUBAME points. Testing status with Docker containers is also mentioned.

# Logistics

**ETH Zurich, Main Building (Hauptgebäude, HG)**
Rämistrasse 101, 8092 Zürich T +41 44 632 11 11

From the airport (journey time about 30 minutes) take the train to Zurich Central Station. From the Central Station take tram no. 10 (direction Airport) or tram no. 6 (direction Zoo) as far as the ETH/University Hospital stop. Alternatively take the Polybahn from Central to ETH Polyterrasse.




**Wireless Access**
Eduroam is available throughout the entire campus.

You may also use the following credentials:

SSID: public or public-5
Login: adac6
Password: ADAC@cscs

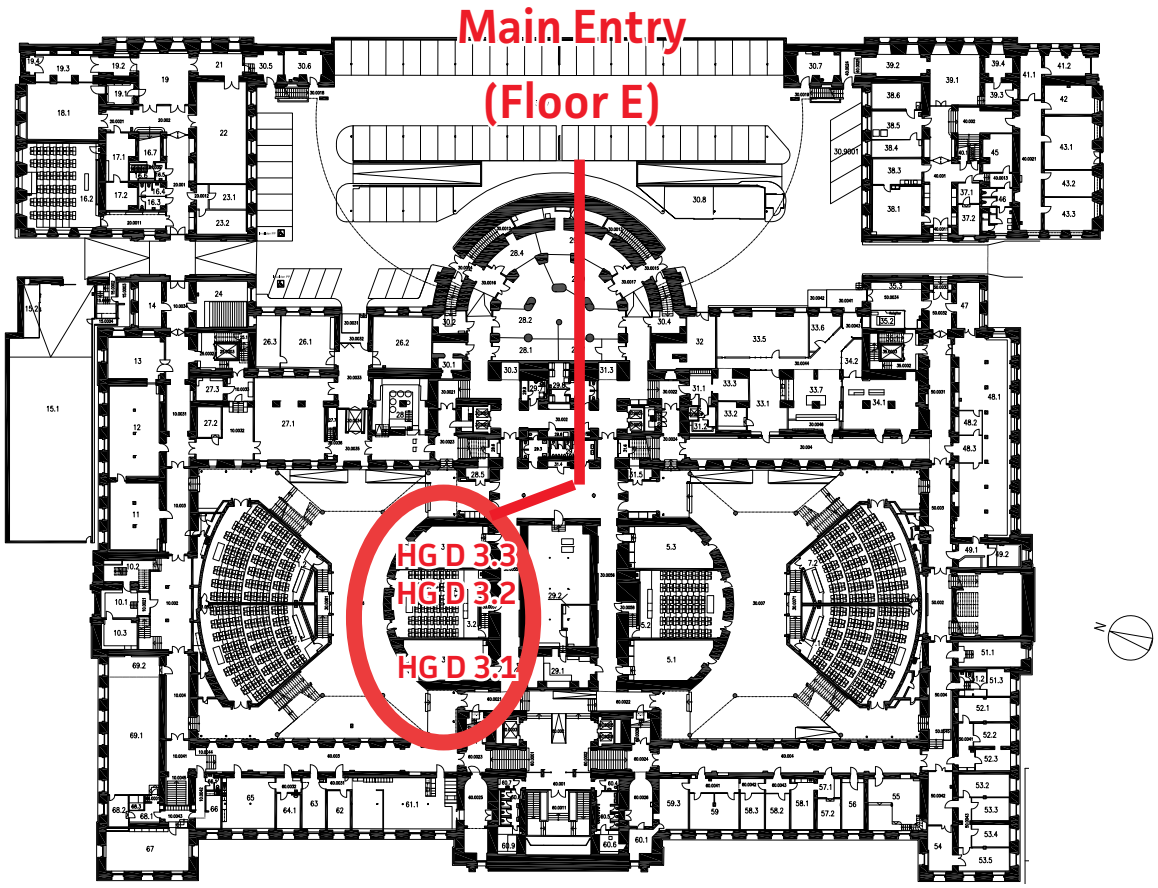
If no login page is shown automatically upon connection, please go to:

https://enter.ethz.ch/welcome

and authenticate.

**Location of Rooms**
From the main entry (Floor E) you have to go down by one floor (Floor D).

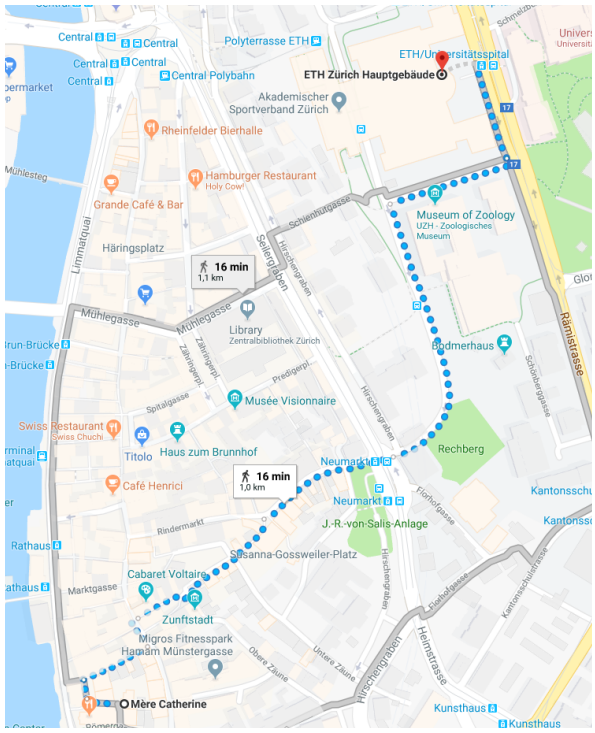
The workshop will take place in HG D 3.2 (plenary presentations) and in HG D 3.1, HG D 3.3, HG E 33.5 (breakout groups).



**Banquet – Sorell Hotel Zürichberg — Wednesday, June 20, 18:30**
Orellistrasse 21, 8044 Zürich T +41 44 268 35 35

In front of the main entry of ETH Zurich, take the number 6 tram (in the direction of Zoo). Get off at the final stop, Zoo, and walk the last 100 metres back to the Sorell Hotel Zürichberg.

**Banquet – Restaurant Mère Catherine — Thursday, June 21, 18:30**
Nägelihof 3, 8001 Zurich, T +41 44 250 59 40


Getting to the Restaurant Mère Cathering is a nice 15 minutes walk in the old city of Zurich from the main building of ETH Zurich.

# Participants List

Nora Abi Akar, ETH Zurich / CSCS
Sadaf Alam, ETH Zurich / CSCS
Muhammad Atif, Australian National University
Paul Baumeister, Jülich Supercomputing Centre
Lucas Benedicic, ETH Zurich / CSCS
Nicholas Cardo, ETH Zurich / CSCS
Florina Ciorba, University of Basel
Johannes de Fine Licht, ETH Zurich / CSCS
Michele De Lorenzi, ETH Zurich / CSCS
Jens Domke, Tokyo Institute of Technology
Evangelos Eleftheriou, IBM Research - Zurich
Toshio Endo, Tokyo Institute of Technology
Christian Engelmann, Oak Ridge National Laboratory
Benjamin Evans, Australian National University
Oliver Fuhrer, MeteoSwiss
Tom Gibbs, NVIDIA
Berk Hess, Royal Institute of Tehcnology
Geetika Gupta, NVIDIA
Remco Havenith, University of Groningen
Torsten Hoefler, ETH Zurich / CSCS
Hideyuki Jitsumoto, Tokyo Institute of Technology
Prashanth Kanduri, ETH Zurich / CSCS
Mark Klein, ETH Zurich / CSCS
Christian Knief, Docker
Doug Kothe, Oak Ridge National Laboratory
Maxime Martinasso, ETH Zurich / CSCS
Thomas Metzger, Intel
Jeff Nichols, Oak Ridge National Laboratory
Akihiro Nomura, Tokyo Institute of Technology
Hans Pabst, Intel
Dirk Pleiter, Forschungszentrum Juelich
Duncan Poole, NVIDIA
Ning Qiao, University of Zurich and ETHZ
Dale Roberts, Australian National University
Daniel Rodwell, Australian National University
Thomas Schulthess, ETH Zurich / CSCS
Sean Smith, Australian National University
Bill Sparks, Cray, Inc.
T.P. Straatsma, Oak Ridge National Laboratory
Frederick Streitz, Lawrence Livermore National Laboratory
Matthieu Taillefumier, ETH Zurich / CSCS
Valentina Tamburello, ETH Zurich / CSCS
Felix Thaler, ETH Zurich / CSCS
Joost VandeVondele, ETH Zurich / CSCS
Becky Verastegui, Oak Ridge National Laboratory
Jeffrey Vetter, Oak Ridge National Laboratory
Allan Williams, Australian National University
Rio Yokota, Tokyo Institute of Technology

CSCS

# About the Accelerated Data Analytics and Computing Institute (ADAC)

The Accelerated Data Analytics and Computing Institute has been established to explore potential future collaboration among UT-Battelle, LLC (UT-Battelle), the Swiss Federal Institute of Technology, Zurich (Eidgenössische Technische Hochschule Zürich/ ETH Zurich), and Tokyo Institute of Technology (Tokyo Tech). Consistent with their respective missions, the Participants seek to collaborate and leverage their respective investments in application software readiness in order to expand the breadth of applications capable of running on accelerated architectures. All three organizations manage HPC centers that run large, GPU-accelerated supercomputers and provide key HPC capabilities to academia, government, and industry to solve many of the world's most complex and pressing scientific problems.

ADAC will focus on multiple objectives spanning performance, hardware, and applications, including:

- Adapting important scientific and engineering applications to hybrid accelerated architectures.
- Partnering with HPC vendors to evaluate architecture diversity.
- Enabling collaborative scientific efforts in hybrid accelerated data and compute.
- Ensuring sustainability and portability of critical applications.
- Sharing best practices regarding the operation, management, and procurement of HPC resources.

The institute lays the groundwork for more focused collaboration centered around three inaugural technical areas—applications, performance, and resource management. Designated representatives from each member institution serve as the leads in these areas.

# Workshop 6 Organization Commitee

**Program Chair**
Becky J. Verastegui, Oak Ridge National Laboratory

**Program Commitee**
Sadaf Alam, ETH Zurich / CSCS

**Local Organisation**
Michele De Lorenzi, ETH Zurich / CSCS
Tatjana Ruefli, ETH Zurich / CSCS
Valentina Tamburello, ETH Zurich / CSCS

During the workshop the organization commitee can be reached by phone as follows:
- Becky +1 865 316 8878
- Michele +41 79 623 25 83

**Conference Web Page**
iadac.github.io