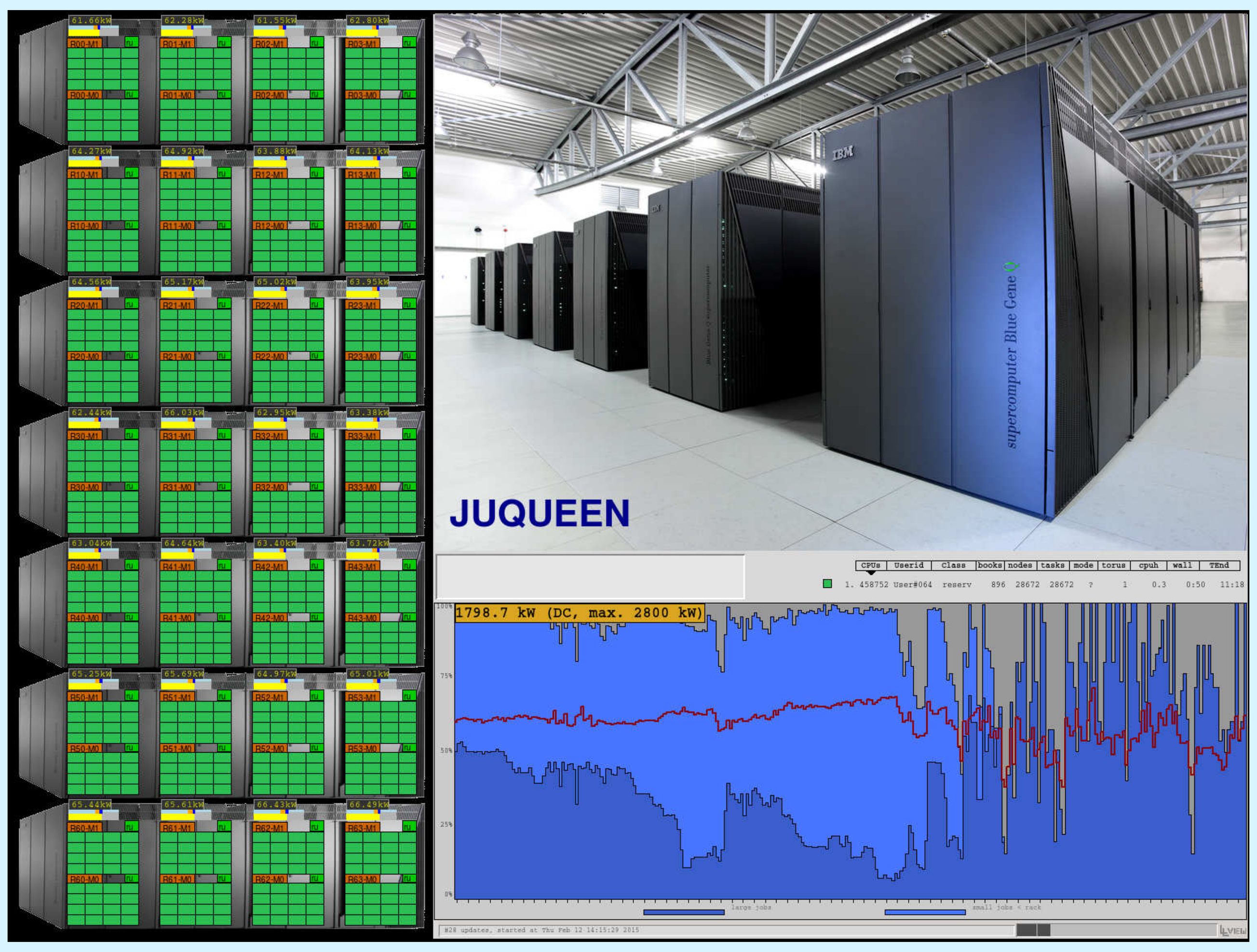
High-Q en route to exascale

Scaling applications to the full JUQUEEN system with 458 752 cores and over 1.8M threads

Dirk Brömmel, Wolfgang Frings, and Brian J. N. Wylie
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH

JUQUEEN Porting & Tuning and Extreme Scaling Workshops, 2–6 February 2015

- **7 teams invited to (im)prove their application scalability with 30 hours of dedicated access to the entire JUQUEEN system**
- **Expert assistance provided by JSC Simulation Laboratories (SimLabs) and Cross-sectional teams, with IBM and JUQUEEN technical support**
- **All 7 teams within 24 hours successfully ran their codes using the full system, preparing for entry to the High-Q Club**
- **Additional experience gained about effectively exploiting extremely scalable computer systems**



CoreNEURON (F. Delalondre & A. Ovcharenko, EPFL Blue Brain Project) simulation of electrical activity of neuronal networks including morphologically detailed neurons

FE2TI (M. Lanser, Universität Köln & O. Rheinbach, TUB Freiberg) scale-bridging incorporating micro-mechanics in macroscopic simulations of multi-phase steels

FEMPAR (A. Martín, UPC-CIMNE) massively-parallel finite-element simulation of multi-physics problems governed by PDEs

ICON (T. Jahns, DKRZ & C. Meyer, JSC SimLab Climate Science) icosahedral non-hydrostatic atmospheric model

MPAS-A (D. Heinzeller, KIT) multi-scale non-hydrostatic atmospheric model for global, convection-resolving climate simulations

psOpen (M. Gauding, RWTH-ITV Inst. for Combustion Technology & J.-H. Göbbert, JARA) direct numerical simulation of fine-scale turbulence

SHOCK (M. Gageik, RWTH Shock Wave Laboratory) structured high-order finite-difference kernel for compressible flows

JUQUEEN

IBM BlueGene/Q with **28 racks**, 28 672 nodes.
Each 1.6 GHz PowerPC A2 compute **node** has:

- 16 cores with 64 hardware threads
- 16 GB memory, 32 MB L2 Cache

In total 458 752 cores with 1 835 008 threads.
Aggregate I/O bandwidth 200 GB/s to GPFS.

Impediments to large-scale executions

File I/O Large machine runs often require parallel I/O, but which is no guarantee of success: SHOCK and CoreNEURON had to limit or disable I/O. ICON, MPAS-A, and psOpen employ PIO/NetCDF, HDF5 or pNetCDF to read in up to 1.2 TB of data. An exception are codes making use of SIONlib, which is developed by JSC to address I/O scalability. Examples using SIONlib are High-Q members KKRnano, MP2C, and muPhi.

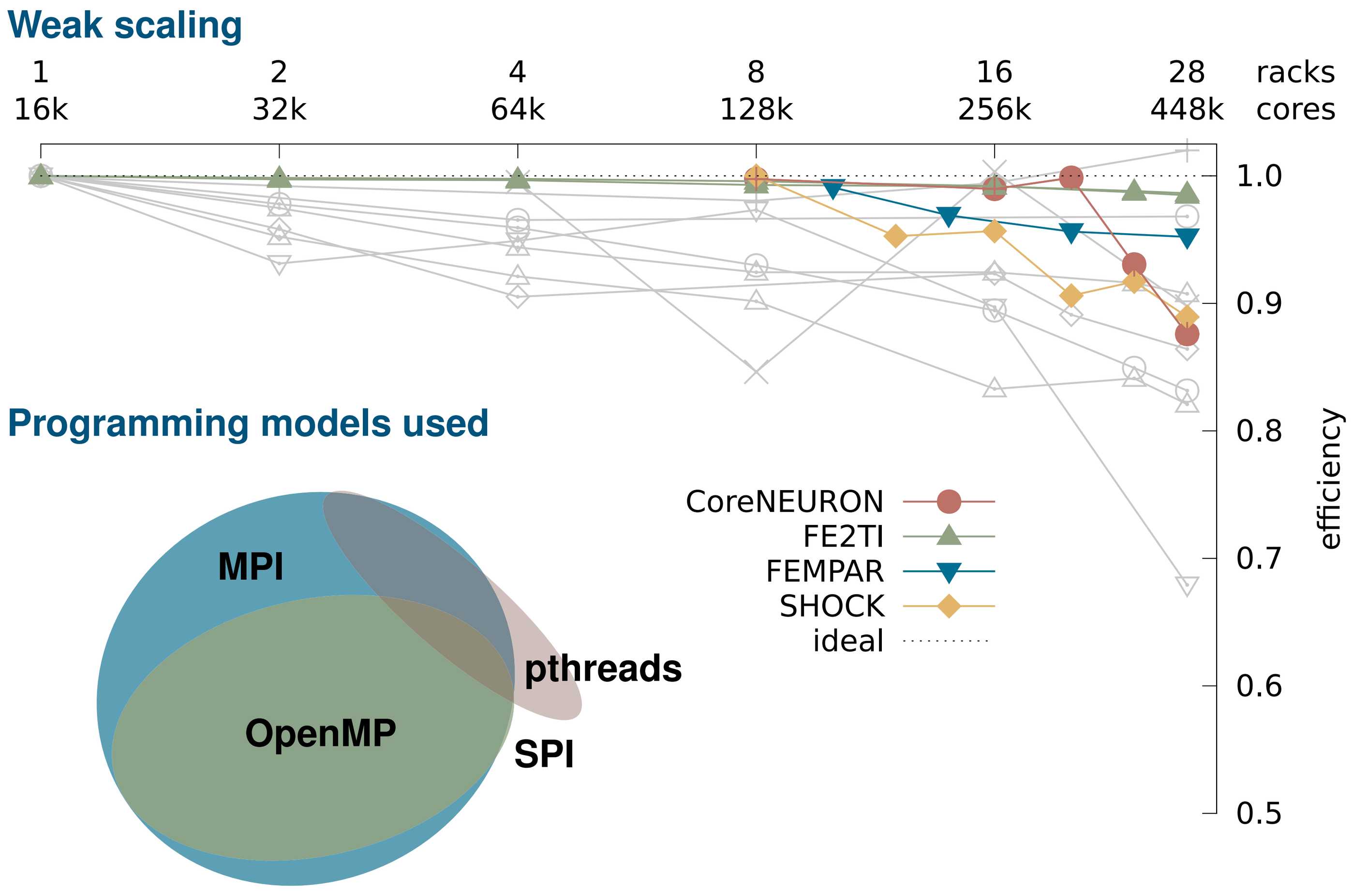
Memory The limited amount of memory available per node provides constraints not only for strong scaling runs: After tuning the problem size CoreNEURON was able to use 15.9 GB (99%) of the available RAM, FEMPAR had to trade memory for MPI against prohibitive MPI execution times, and MPAS-A was limited to 16 MPI ranks per node to fit in 1 GB per process.

Results

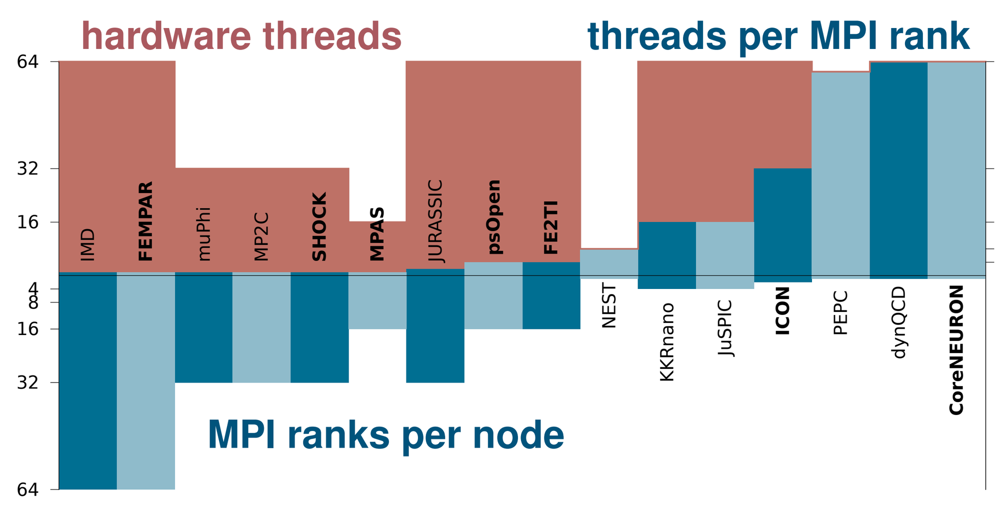
On the right is a selection of results from the High-Q Club and Extreme Scaling Workshop:

- Weak and strong scaling graphs summarise some of the submitted performance data. (For reference, High-Q club codes grey and unlabelled.)
- Venn set diagrams with statistics about the programming models and languages used by the codes provide the colour key to the code listing. While all 3 programming languages are roughly equally popular, mixed-mode programming combining MPI and multi-threading dominates.
- A bar chart shows the breadth of MPI process / thread combinations used and the extent of oversubscription of cores to fully exploit the hardware capabilities.
- More results can be found in the technical report **FZJ-JSC-IB-2015-01**

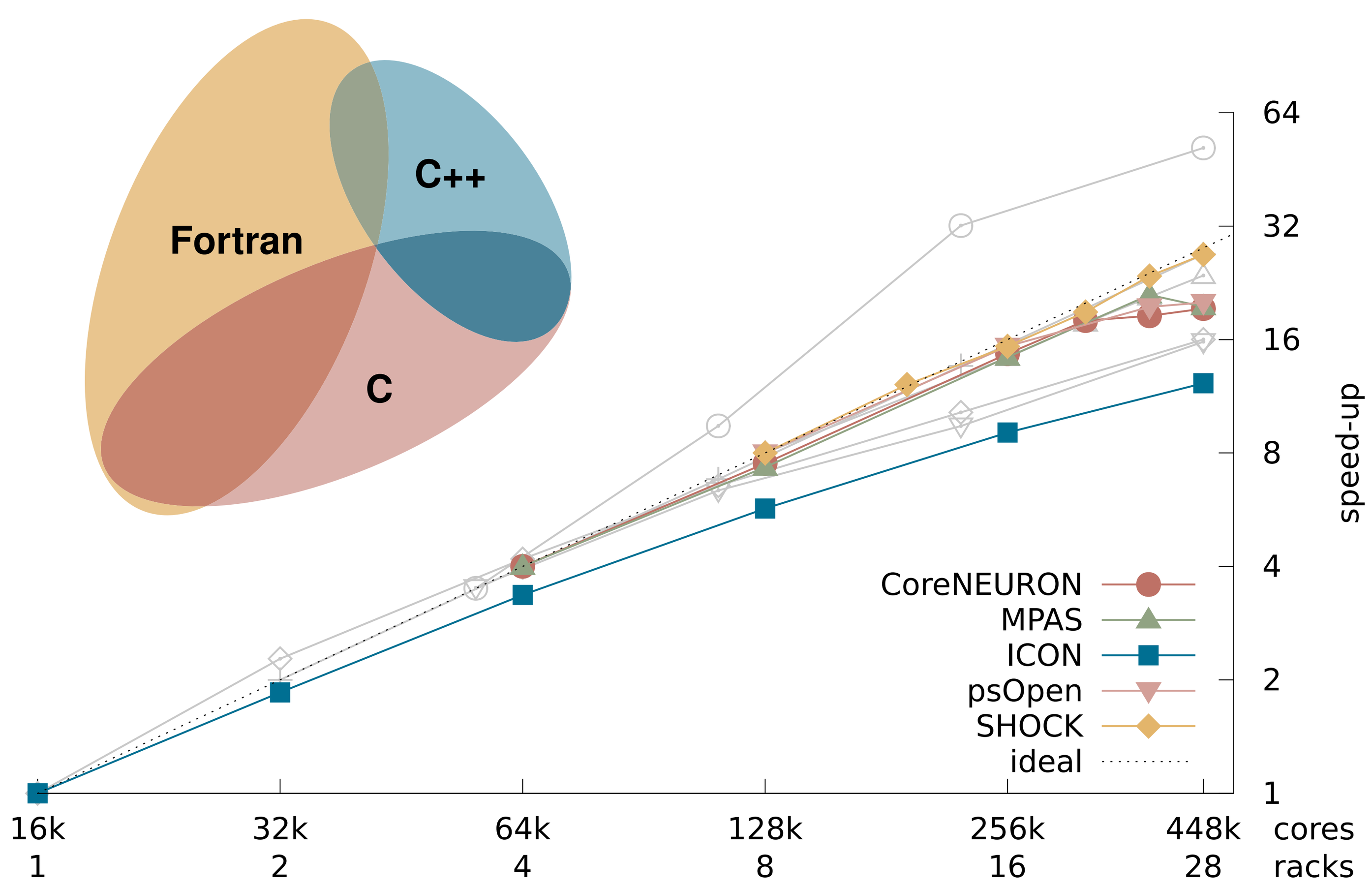
Weak scaling



Programming models used



Programming languages used



Strong scaling

High-Q Club Codes

dynQCD (JSC SimLab Nuclear and Particle Physics & Bergische Universität Wuppertal) lattice quantum chromodynamics with dynamical fermions

Gysela (CEA-IRFM Cadarache) gyrokinetic semi-Lagrangian code for plasma turbulence simulations

IMD (Ruhr-Universität Bochum & JSC SimLab Molecular Systems) classical molecular dynamics simulations

JURASSIC (JSC SimLab Climate Science) solver for infrared radiative transfer in the Earth's atmosphere

JuSPIC (JSC SimLab Plasma Physics) fully relativistic particle-in-cell code, plasma physics simulations, laser-plasma interaction

KKRnano (FZJ-IAS) Korringa-Kohn-Rostoker Green function code for quantum description of nano-materials in all-electron density-functional calculations

MP2C (JSC SimLab Molecular Systems) massively-parallel multi-particle collision dynamics for soft matter physics and mesoscopic hydrodynamics

$\mu\phi$ (muPhi) (Universität Heidelberg) modelling and simulation of water flow and solute transport in porous media, algebraic multi-grid solver

Musubi (Universität Siegen) multi-component Lattice Boltzmann solver for flow simulations

NEST (FZJ/INM-6 & IAS-6) large-scale simulations of biological neuronal networks

OpenTBL (Universidad Politécnica de Madrid) direct numerical simulation of turbulent flows

PEPC (JSC SimLab Plasma Physics) tree code for *N*-body simulations, beam-plasma interaction, vortex dynamics, gravitational interaction, MD simulations

PMG+PFASST (LBNL, Universität Wuppertal, USI & JSC) space-time parallel solver for systems of ODEs with linear stiff terms, e.g. from lines discretisations of PDEs

PP-Code (University of Copenhagen) simulations of relativistic and non-relativistic astrophysical plasmas

TERRA-NEO (Universität Erlangen-Nürnberg, LMU & TUM) modeling and simulation of earth mantle dynamics

waLBerla (Universität Erlangen-Nürnberg) Lattice-Boltzmann method for the simulation of fluid scenarios

Programming Model
and Language

Contact: {d.broemmel, w.frings, b.wylie}@fz-juelich.de - **Website:** www.fz-juelich.de/ias/jsc/high-q-club

Member of the Helmholtz Association