Enabling energy systems research on HPC

Thomas Breuer
Jülich Supercomputing Centre
Forschungszentrum Jülich GmbH
Jülich, Germany
t.breuer@fz-juelich.de

Karl-Kien Cao
Institute of Networked Energy Systems
German Aerospace Center
Stuttgart, Germany
Karl-Kien.Cao@dlr.de

Manuel Wetzel
Institute of Networked Energy Systems
German Aerospace Center
Stuttgart, Germany
Manuel.Wetzel@dlr.de

Ulrich Frey
Institute of Networked Energy Systems
German Aerospace Center
Stuttgart, Germany
Ulrich.Frey@dlr.de

Shima Sasanpour
Institute of Networked Energy Systems
German Aerospace Center
Stuttgart, Germany
Shima.Sasanpour@dlr.de

Jan Buschmann
Institute of Networked Energy Systems
German Aerospace Center
Oldenburg, Germany
jan.buschmann@dlr.de

Aileen Böhme
GAMS Software GmbH
Braunschweig, Germany
aboehme@gams.com

Charlie Vanaret
Applied Algorithmic Intelligence Methods
Technische Universität Berlin
Berlin, Germany
vanaret@math.tu-berlin.de

***Abstract*—Energy systems research strongly relies on large modeling frameworks. Many of them use linear optimization approaches to calculate blueprints for ideal future energy systems, which become increasingly complex, as do the models. The state of the art is to compute them with shared-memory computers combined with approaches to reduce the model size. We overcome this and implement a fully automated workflow on HPC using a newly developed solver for distributed memory architectures. Moreover, we address the challenge of uncertainty in scenario analysis by performing sophisticated parameter variations for large-scale power system models, which cannot be solved in the conventional way. Preliminary results show that we are able to identify clusters of future energy system designs, which perform well from different perspectives of energy system research and also consider disruptive events. Furthermore, we also observe that our approach provides the most insights when being applied to complex rather than simple models.**

Keywords*—*energy system modeling, mathematical optimization, distributed solver, model coupling, HPC workflow

I. Motivation

High quality models are essential to study scenarios of the future. In energy systems research the interaction of multiple technologies is modelled to find scenarios that are in line with climate targets and which provide insights for decision making, e.g., in energy policy. However, for policy advice, the question of the reliability of such scenarios is of great importance, as they can be subject to large uncertainties. Nevertheless, the state of the art in energy systems research is to study only small subsets of all possible scenarios. This has however proven to be inadequate as the models are highly sensitive to certain input data. Instead, broad parameter variations are required which automatically generate, calculate, and evaluate a hitherto unattained number of energy scenarios including disruptive events. Another major challenge is the reduction of model computing times for solving the optimization problems because the widely-used commercial solvers show poor scalability and are limited to single shared-memory compute nodes. Thus, models are downscaled to lower temporal and spatial resolutions or a lower technological diversity which limits the applicability to real-world decision support. In this context, it needs to be stressed that single models, such as optimization models cannot provide insights to all decision-relevant aspects. Therefore, different models are often coupled. However, the resulting data processing workflows usually require a lot of human interaction.

To summarize: To perform best-practice analyses in energy systems research, one needs to first perform a multitude of model runs each representing one scenario within the funnel of countless possible futures. Second, it requires methods to make complex large-scale models solvable in manageable time. Third, only assessments from different perspectives are able to provide comprehensive results which add value to decision-making processes. This calls for coupled modeling workflows.

Our contribution to these challenges is an automated workflow which provides both the investigation of many different scenarios by farming, and a solver for energy system optimization models suited for HPC. Due to the complexity of the envisaged workflow, a specialized tool is applied to fully automate the management of the execution on an HPC system. Efficiently leveraging of the capabilities of HPC in both compute and storage, could be a game changer for the energy systems research community.

II. 2 Methodological Approach

A. *New parallel solver – PIPS-IPM++*

Optimization problems that result from energy system optimization models have a special structure. To balance energy from renewable sources those models interlink different components of the energy system on different dimensions. For example, battery storage and power grids cause interrelations of discrete time steps and network nodes, respectively. Accordingly, the mathematical structure of the corresponding optimization models shows linkages that prevent a straight-forward decomposition, and thus parallelization of solvers.

The new solver PIPS-IPM++ that we have developed is able to treat these so-called linking variables and linking constraints. It is an advancement of the parallel interior-point solver (PIPS) originally developed at Argonne National Laboratory to solve a different problem class [1]. Involving code changes in more than

We thank our colleagues from the research project UNSEEN, a project funded by the German Federal Ministry for Economic Affairs and Climate Action, grant number FKZ 03EI1004. The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. (www.gauss-centre.eu) for funding this project by providing computing time through the John von Neumann Institute for Computing (NIC) on the GCS Supercomputer JUWELS at Jülich Supercomputing Centre (JSC).

60k lines, PIPS-IPM++ can be considered as new solver. Its implementation uses MPI and OpenMP. We modified and extended the code to exploit the block-structure of the coefficient matrix of large-scale linear optimizations problems despite the presence of linking constraints. Further enhancements include distributed pre- and post-solve and regularization of augmented linear system to reduce numerical instability. A detailed description of the algorithm that allows the treatment of the so-called doubly-bordered arrow-ahead block structure of the matrix can be found in [2]. We applied the latest and most advanced version of PIPS-IPM++ which treats the matrix by a hierarchical approach [3]. The advantage of this approach is that it enables solving of very large model instances using significantly less memory.

B. *Parameter sampling and model coupling*

For performing a broad range parameter variation in scenario analyses, we use a new tool for sampling consistent parameter sets as input to the energy system model REMix [4]. For that purpose, we define intercorrelations between parameters and randomly select from parameter value distributions which we defined based on a literature review of more than 50 sources (e.g. [5] and [6]) and assuming truncated normal distribution functions. This represents the first step of our HPC workflow pipeline. It enables us to automatically generate a huge number of different scenarios to map the possibility space. Hence, we have drastically increased the number of scenarios to tackle the inherently large uncertainties. Next, each scenario is pre-processed so that it can be solved in parallel by PIPS-IPM++. Thereafter, the calculated solution points need an additional postprocessing step which converts them back into the parameter space of the REMix model. Finally, the solved scenario is used as input to a further energy system model, AMIRIS [7]. This model provides insights into each scenario from a more realistic market perspective where - contrary to the central planning perspective of REMIX - a variety of stakeholders interact. Together with further ex-post indicator calculations this enables comprehensive statistical evaluations as the final step of our approach.

C. *Workflow environment – JUBE*

A basic prerequisite of our approach is an HPC workflow environment that can execute the described pipeline flow for a large amount of scenarios in an automated manner. Therefore, we have used the flexible and generic Python package JUBE [8] for the technical implementation of our approach. Since the individual scenarios can be processed independently of each other, we have parallelized JUBE so that work packages can be processed in parallel. In combination with the distributed solver PIPS-IPM++ we have achieved a 2-level parallelism designed for HPC: PIPS-IPM++ reduces the calculation time for each scenario and JUBE speeds up the total execution time of the overall workflow.

In addition, we have extended JUBE by a database output so that the processing of results and the workflow overview is simplified and user-friendly. Beside these extensions, JUBE organizes the automated execution, manages the tailor-made naming scheme, and prepares the hierarchical data structure that we have designed for the data exchange between the components of the workflow. We have either extended or newly developed each of these components such that all of them fulfill our requirements and guarantee data consistency throughout the whole pipeline.

III. PRELIMINARY RESULTS

As a first step, we have executed several tests with PIPS-IPM++. Two of them are shown on the poster. During the development phase we have used a simplified and flexible benchmark instance, called SIMPLE-PIPS, to validate our solver and compare the scaling with four widely used state of the art commercial solvers. With 5.1M rows and 5.6M columns this is a relatively small instance compared to the much larger PyPSA-Eur setup with a problem size of 234M rows and 213M Columns [9]. For both test cases, our open-source solver PIPS-IPM++ outperforms state of the art commercial solvers on massively parallel architectures such as the JUWELS cluster [10].

After successfully demonstrating the performance and usability we have integrated the solver in the center of the HPC workflow. In a proof of concept 1K small instances of German power system scenarios have been processed. The scientific analysis of the final dataset indicated that the instances were too small to derive any further conclusions. But the technical analysis has shown that we have developed a stable pipeline flow so that the first real-world models can be processed. At the time of writing around 40 large instances of a highly resolved German power system scenario have passed through the whole pipeline while further scenarios are still running or waiting to get started. So far around 1 TB of data including roughly 1K directories and 42K files have been generated. Even though the number of finalized scenarios is on the lower limit for a statistical analysis, a first indicator evaluation of more than 30 indicators shows that the scenarios can be grouped into 3 clusters. First, scenarios that reveal high gas consumption and accordingly higher CO2-emissions. Second, scenarios with high shares of renewable power generation where load-balancing is more challenging but CO2 can be significantly reduced. The last group of scenarios is in between these two worlds. Although these are preliminary results of an ongoing study, our approach shows that we are able to pave the way for a better and more robust energy system analysis in the future.

IV. OUTLOOK

Our next step is to cover the uncertainties in energy scenario modeling by performing more comprehensive parameter variations. Conceivable candidates are parameters that describe weather and energy demand projected into the future. Furthermore, the ongoing trend in energy systems analysis to increase model resolutions calls for technologically more representative modeling approaches. Therefore, solving mixed-integer linear programs on HPC becomes the next challenge in order to also sample the existence, expansion or outage of individual technology infrastructures, such as transmission lines or power plants. We have prepared our workflow and solver software to solve such problems. We can therefore build upon our LP-solver PIPS-IPM++. However, further research aims at heuristics for finding feasible integer solutions. In this context, one approach we are currently developing is the application of neural networks for this purpose.

REFERENCES

[1] C. G. Petra, O. Schenk, M. Anitescu, “Real-Time Stochastic Optimization of Complex Energy Systems on High-Performance Computers”, https://doi.org/10.1109/MCSE.2014.53

[2] D. Rehfeld, H. Hobbie, D. Schönheit, T. Koch, D. Möst, A. Gleixner, "A massively parallel interior-point solver for LPs with generalized arrow-head structure, and applications to energy system models”. Eur. J. Oper. Res. 296.1 (2022) https://doi.org/10.1016/j.ejor.2021.06.063

[3] Nils-Christian Kempke, Thorsten Koch, Charlie Vanaret, Karl-Kiên Cao, Manuel Wetzel (2021), “PIPS-IPM++ - A Parallel Interior-Point Method for Solving Large-Scale Doubly Bordered Block Diagonal LPs”, OR 2021, Bern

[4] H. C. Gils, Y. Scholz, T. Pregger, D. L. de Tena, D. Heide, “Integrated modelling of variable renewable energy-based power supply in Europe”. Energy 123 (2017), https://doi.org/10.1016/j.energy.2017.01.115

[5] Ruiz, P., et al. ENSPRESO-an open, EU-28 wide, transparent and coherent database of wind, solar and biomass energy potentials. Energy Strategy Rev. 26 (2019)

[6] Bernath, C., et al. Langfristszenarien für die Transformation des Energiesystems in Deutsch-land—Modul 3: Referenzszenario und Basisszenario. Karlsruhe, Germany (2017)

[7] M. Deissenroth, M. Klein, K. Nienhaus, M. Reeg, “Assessing the plurality of actors and policy interactions: agent-based modelling of renewable energy market integration”. Complexity 2017 (2017)

[8] S. Lührs, D. Rohe, A. Schnurpfeil, K. Thust, W. Frings, 2015, “Flexible and Generic Workflow Management”, In Proceedings of International Conference on Parallel Computing 2015, ParCo 2015, Edinburgh, United Kingdom, http://dx.doi.org/10.3233/978-1-61499-621-7-431

[9] Cao, Karl-Kien und Wetzel, Manuel und Kempke, Nils-Christian und Koch, Thorsten (2021) “Pushing computational boundaries: Solving integrated investment planning problems for large-scale energy systems with PIPS-IPM+”. Operations Research 2021, 1.-3. Sep. 2021, Bern

[10] Jülich Supercomputing Centre. (2021). “JUWELS Cluster and Booster: Exascale Pathfinder with Modular Supercomputing Architecture at Juelich Supercomputing Centre”. Journal of large-scale research facilities, 7, A183. http://dx.doi.org/10.17815/jlsrf-7-183