A Generic Scalable Method for Scheduling Distributed Energy Resources using Parallelized Population-based Metaheuristics

Hatem Khalloof
Karlsruhe Institute of Technology, IAI
Karlsruhe, Germany
hatem.khalloof@kit.edu

Wilfried Jakob
Karlsruhe Institute of Technology, IAI
Karlsruhe, Germany
wlfried.jakob@kit.edu

Shadi Shahoud
Karlsruhe Institute of Technology, IAI
Karlsruhe, Germany
shadi.shahoud@kit.edu

Clemens Duepmeier
Karlsruhe Institute of Technology, IAI
Karlsruhe, Germany
clemens.duepmeier@kit.edu

Veit Hagenmeyer
Karlsruhe Institute of Technology, IAI
Karlsruhe, Germany

ABSTRACT
Recent years have seen an increasing integration of distributed renewable energy resources into existing electric power grids. Due to the uncertain nature of renewable energy resources, network operators are faced with new challenges in balancing load and generation. In order to meet the new requirements, intelligent distributed energy resource plants can be used which provide as virtual power plants e.g. demand side management or flexible generation. However, the calculation of an adequate schedule for the unit commitment of such distributed energy resources is a complex optimization problem which is typically too complex for standard optimization algorithms if large numbers of distributed energy resources are considered. For solving such complex optimization tasks, population-based metaheuristics -as e.g. evolutionary algorithms- represent powerful alternatives. Admittedly, evolutionary algorithms do require lots of computational power for solving such problems in a timely manner. One promising solution for this performance problem is the parallelization of the usually time-consuming evaluation of alternative solutions. In the present paper, a new generic and highly scalable parallel method for unit commitment of distributed energy resources using metaheuristic algorithms is presented. It is based on microservices, container virtualization and the publish/subscribe messaging paradigm for scheduling distributed energy resources. Scalability and applicability of the proposed solution are evaluated by performing parallelized optimizations in a big data environment for three distinct distributed energy resource scheduling scenarios. Thereby, unlike all other optimization methods in the literature -to the best knowledge of the authors-, the new method provides cluster or cloud parallelizability and is able to deal with a comparably large number of distributed energy resources. The application of the new proposed method results in very good performance for scaling up optimization speed.

CCS CONCEPTS
• Computing methodologies → Massively parallel algorithms;
Planning under uncertainty; Genetic algorithms;
• Theory of computation → Evolutionary algorithms;
• Software and its engineering → Software as a service orchestration system;
• Applied computing → Service-oriented architectures; Event-driven architectures.

KEYWORDS
Evolutionary Algorithms, microservices, virtualization, container, metaheuristic optimization, scheduling distributed energy resources, microgrid, cluster computing

ACM Reference Format:
Hatem Khalloof, Wilfried Jakob, Shadi Shahoud, Clemens Duepmeier, and Veit Hagenmeyer. 2020. A Generic Scalable Method for Scheduling Distributed Energy Resources using Parallelized Population-based Metaheuristics. In Proceedings of ACM Conference (Conference’17). ACM, New York, NY, USA, 11 pages. https://doi.org/10.1145/nnnnnn.nnnnnn

1 INTRODUCTION
Renewable Energy Resources (RERs) are recently widely integrated into the grid paving the road for more clean and environment-friendly energy. To facilitate the adoption and management of such RERs, the transition from a traditional centralized grid (macrogrid) to more decentralized grids (microgrids) is required [19, 33]. Microgrids encompass respectively a localized group of Distributed Energy Resources (DERs) as e.g. batteries, fuel generators, solar, wind and gas turbines and distributed electrical loads with demand-response. The flexibility of microgrids provides a significant potential to promote and integrate more DERs to reduce greenhouse gases, meet the increasing energy demand and improve the overall system reliability, efficiency and resilience. A microgrid can be operated either in a grid-connected or in an isolated mode (island mode) as an autonomous power system in parallel with a main grid [26]. Despite being highly effective, microgrids have some limitations such as lack of system protection and customer privacy. Moreover, by increasing the number of DERs in the grid and due to the uncertainties of DERs and load, the efficient control and optimal usage of DERs represents a big challenge [48]. In general, scheduling e.g. scheduling DERs is an NP-hard optimization problem and therefore is typically too complex to be solved by exact optimization methods, especially if large size optimization problems are considered [7, 42].
Metaheuristics such as Evolutionary Algorithms (EAs) became one of the most robust methods to solve such complex problems by finding good local optima or even the global one.

Hence, developing new concepts, optimization approaches and software services based on modern technologies and artificial intelligence optimization techniques e.g. EAs are promising approaches for solving such complex optimization problems as scheduling of DERs.

The central concept of EAs is a population consisting of individuals representing tentative solutions. The individuals encode possible solutions and they are used to explore many areas of the solution space in parallel. Each individual is evaluated by a fitness function to identify its suitability as a solution for the problem. Genetic operators, namely recombination and mutation, are iteratively applied to individuals to generate a new offspring for each generation until a termination criterion has been reached [17, 47]. This approach of having a population of solutions and evaluating them over and over again takes a lot of computational resources for large problem sizes. Therefore, applying EAs for large scale optimization and NP-hard problems such as the problem of scheduling a large number of DERs can be time-consuming and computationally expensive. Moreover, DERs can contain several heterogeneous energy resources resulting in a more complex and heterogeneous search space. To speed up EAs, three different parallelization models, namely the Global Model (Master-Slave Model), the Fine-Grained Model and the Coarse-Grained Model have been introduced and investigated in [9]. In the Global Model, the evaluation step is parallelized over several computing units (called slaves). In the Fine and Coarse-Grained Models, the population is structured to apply the genetic operators in parallel.

Over the last decades, various approaches and frameworks e.g. [1, 3, 8, 11–14, 18, 23, 25, 29, 30, 38–41, 44] have been introduced to enable the distributed/parallel processing of EAs following the above three parallelization models. For most of these frameworks e.g. [1, 3, 8, 11, 12, 18, 23, 29, 40, 41, 44], a monolithic software architecture was the classical approach for the implementation which decreases the modularity, usability and maintainability of the application. Recently, Big Data technologies such as Hadoop and Spark have been applied to speed up EAs e.g. [5, 11, 18, 23, 40, 41, 44]. However, most of these approaches also have a monolithic architecture which lacks hard boundaries and tends to become, with added functionality, complex and tightly coupled. This, in fact, limits the ability to provide simple and practical methods to plug in problem-specific functionality e.g. simulators and even to integrate existing EAs. By emerging modern software technologies, namely microservices, container virtualization and the publish/subscribe messaging paradigm, the parallelization of EAs in cluster and cloud environments to speed up EAs has become even more relevant, see e.g. [13, 14, 25, 39]. Unlike monolithic applications, a microservices-based application contains several small, autonomous, highly cohesive and decoupled services that work together to perform a specific task. Since all services are independent from each other, each microservice is able to utilize its own technology stack allowing great flexibility. The independence of the services allows each service to scale on demand. Microservice applications comprise two main features, namely modularity and technology heterogeneity which allow the microservices to be developed by different teams based on different technologies. These advantages combined with container runtime automation unlock the full potential of a parallelized EA by executing it on large scale computing clusters [25].

In the present work, a new highly scalable, generic and distributed approach to schedule DERs is introduced. The microservice and container virtualization-based framework presented in [25] is adapted to carry out the required tasks. As the simulation based evaluation is by far the most time consuming part, the proposed framework distributes EAs according to the Global Model (Master-Slave model) [9] where the evaluation is distributed over several computing units. On-demand deployment of services on a high performance distributed computing infrastructure, namely a computing cluster, is supported.

For evaluation purposes, the EA GLEAM (General Learning Evolutionary Algorithm and Method) [6] is integrated into the framework. As a test task, the creation of an hourly day-ahead schedule plan for a simulated microgrid in island mode is chosen. In this isolated microgrid, three use case scenarios are defined. In the first and second scenario, 50 DERs are considered to cover the required power for a simple load profile. The difference between the first and second scenario is mainly in the type of DERs used in each one. More precisely, each DER in the first scenario can generate power in 24 hours, however in the second one, 25 DERs can supply power in 24 hours and the other only in 10 hours. In the third scenario 100 DERs are utilized to supply the requested power from a more complex load profile. In these three scenarios, each DER offers the power with its own price. The main goal of an EA (e.g. GLEAM) for scheduling DERs is to minimize the electricity bill for the load by choosing the DERs that can cover the requested power at each time interval with the lowest price. For evaluation of the scalability and performance of the new solution, the framework is deployed on a cluster with four nodes, each one has 32 Intel cores (2.4 GHz) resulting in 128 independent computing units, an SSD disk and a total of 128 GB of RAM. The nodes are connected to each other by a LAN with 10 GBit/s bandwidth. Then, optimizations runs based on the mentioned scenarios are performed and evaluated.

The rest of the present paper is structured as follows. The next section reviews some related work for scheduling DERs based on EAs. Section 3 introduces the adapted architecture of the proposed approach. Section 4 starts with a short introduction of the EA GLEAM serving as metaheuristic, and continues with a description of the defined use case scenarios, and finally obtained results are presented. Section 5 concludes with a summary and planned future work.

2 RELATED WORK

EAs have attracted the attention of researchers to solve several optimization problems in energy systems, namely expansion planning, e.g. [34], maintenance scheduling, e.g. [45], scheduling energy resources (unit commitment) and economic dispatch [4, 24, 27, 28, 31, 32, 36, 43, 46], to name a few. In recent extensive overviews, Zia et al. [49] and Alvarado-Barrios et al. [2] presented comprehensive studies about different methods and techniques used in energy management systems to optimize and schedule the operations. In the following sections, we summarize some of these work studied in
For the problem of scheduling DERs, the authors of [4, 27, 28, 31, 32, 36] implemented GAs to schedule the power generation in microgrids. Several microgrids with sizes ranging from six to 12 DERs and a wide variety of generators e.g. PVs, wind turbines, microturbines and diesel engines and energy storage systems (batteries) are considered in these studies. While in [27, 31, 32, 36] standard GA implementations were used, in [4] a memory-based GA algorithm and in [28] an improved GA combined with simulated annealing technique were utilized to accelerate GAs for finding the optimal schedule. Minimizing the operation cost was the objective function for all these works. However, in [31] the eco-pollutant treatment costs were additionally considered as objective function. Quan et al. in [36] defined five deterministic and four stochastic case studies solved by GA. They concluded that GA can introduce robust solutions for stochastic optimization problems.

All the previous works focused on developing a respective new optimization algorithm based on EAs for achieving better solution quality. They tested their proposed solutions with microgrids consisting of small numbers of energy resources and deployed them using a monolithic software architecture. This limits the scalability and modularity of the proposed system which in turn restricts the possibility to handle scalable number of DERs. Despite the principally satisfactory performance of using EAs for scheduling DERs, there is no generic and scalable solution to facilitate the usage of the parallel EAs for scheduling a scalable energy system on a scalable runtime environment such as a cluster and to work efficiently with other components e.g. forecasting frameworks and simulators, to the best of our knowledge.

Therefore, the present work introduces a highly parallel and scalable approach using a proven and established software environment based on microservices and container virtualization with full runtime automation on big computing clusters and an easy-to-use web-based management for scheduling DERs. It provides a highly flexible environment for solving the problem of scheduling DERs for external applications e.g. EMS systems, and allows easy communication with other needed tools such as forecasting tools and external simulators.

3 SCHEDULING HIERARCHICAL DISTRIBUTED ENERGY RESOURCES USING EVOLUTIONARY ALGORITHMS

In the following, the concept of hierarchical Distributed Energy Resources (DER) used as example setting in this work are briefly introduced. Then the conceptual architecture of the proposed generic distributed approach for scheduling DERs based on EAs are detailed. The last subsection introduces GLEAM which is used as concrete EA for evaluating the approach.

3.1 Hierarchical Distributed Energy Resources

A DER represents a small or larger scale and self-autonomous subsystem connected to an electricity network for providing renewable energy generation and/or improving the overall power system reliability by balancing the energy supply and demand in a specific part of a power network by providing flexible load options or storage.

Typically, a DER encompasses a group of small generation units such as PVs, wind turbines and diesel generators, electrical loads (demand-response) e.g. electric vehicles or flexible heating systems, and maybe storage. A DER may also contain uncontrollable loads which complicate the DER coordination and management. DERs interconnect bidirectionally to the grid through one or more Point(s) of Common Coupling (PCC) [15]. By the time, the usage of DERs in smart grids will dramatically increase providing more clean energy generated from RERs and additionally also maintaining and increasing power quality and system reliability. However, the coordination of many DERs on large scale level e.g. in a distribution network or industrial area represents a challenge for the traditional Energy Management Systems (EMS). For facilitating the management process, namely the scheduling of DERs, a hierarchical approach proposed in [10, 15] is feasible. Using this approach, an energy subsystem can be structured into a hierarchical group of DERs, whereby each one can contain smaller DERs as shown in Figure 1. More specifically, each DER consists of components that are either elementary generation units or DERs with PCC and local EMS; see Figure 1 for an example. The blue DER1 in level 2 has an EMS which manages one local load located in the same level and n local DERs, namely the black DER1, 2, 4, ..., n located in a lower level. The black DERs e.g. the DER1, have an EMS that manages three elementary generation units, namely 2 PVs and CHP, one storage...
unit (ESS) and one controllable load. All elementary components and DERs have a communication interface (EMS-IF) enabling the EMS to communicate with each other and the outside to collect required data as well as for reacting to external control commands. DERs in the higher level consider the DERs located in lower ones as blackboxes offering required data and an abstract service interface through their EMS-IFs.

3.2 Microservice and Container Virtualization Approach

The conceptual architecture of the proposed highly scalable meta-heuristic optimization solution is derived from [25]. As shown in Figure 2, the architecture has three main tiers, namely the User Interface (UI) Tier, the Cluster Tier and the Distributed Energy Resources (DERs) Tier. On the front-end, the UI Tier is dedicated to user interaction, e.g. input for defining optimization tasks, uploading optimization models, starting and stopping optimization tasks and presenting the obtained results. The UI Tier introduces a simple web-based UI to manage the interaction with the back-end tier. On the back-end, the Cluster Tier contains two sub-layers, namely the Container Layer and the Data Layer. The Container Layer contains all the services necessary to execute a parallel EA in the framework. This includes not only the services that actually execute the EA, but also services for coordinating the execution and distributing the data. Each service is realized as a microservice running in a containerized environment. The Data Layer stores all data and acts as an intermediate for message exchange. To reflect the different properties of the data, the Data Layer is subdivided into a Persistent and Temporary Storage (in-memory database). On the one hand, the Persistent Storage is responsible for storing the data needed for each optimization task such as forecasting data and the final results for further usage. On the other hand, the Temporary Storage stores the intermediate data that are exchanged between services when performing an optimization job. The Temporary Storage also realizes a publish/subscribe message exchange pattern to improve the decoupling among the services. The DERs Tier contains abstractions of the DERs that have to be scheduled. Each DER abstraction provides required data, e.g. the predicted generation, consumption and market price for the period considered, and the necessary technical properties about the nonrenewable sources e.g. diesel generators. This data is needed by the services within the Container Layer for creating an optimized scheduling plan. In the following sections, the container layer with the implemented services will be described in greater detail.

3.2.1 Container Layer. For finding the optimal scheduling plan for a group of DERs using parallel EAs, the software solution needs to perform several tasks, namely coordination of the execution of tasks, i.e. managing the containers, starting and managing external simulators and executing the parallel EAs for generating, splitting, distributing, and evaluating the chromosome lists, and collecting and joining the subresults to form the final results and applying the genetic operators. These tasks are performed by ten decoupled and cohesive microservices as shown in Figure 3. The presented microservices are adapted from [25] for scheduling DERs using parallel EAs based on the global parallelization model. Three new services, namely Supporting and Learning Service, DERs Service and Interpretation Service are added. Some of the existing microservices are modified and renamed to reflect their extended functionalities and new tasks. The Distribution & Synchronization service is split into two services depending on its functionalities, namely the Optimization Task Coordination Service and the Splitting & Joining Service. The framework is designed with this hierarchical structure for facilitating manageability and allowing extensibility. For instance, in this hierarchy, the Container Management Service is employed to add resources for the Interpretation Service and Calculation Service instances, namely containers as many as required. Adapted and newly added microservices are described in detail below.

3.2.2 Coordination Service. The Coordination Service (formerly named Optimization Job Management Service [25]) is one of the core parts of the framework. It acts not only as a coordinator for multiple jobs, but also for the whole framework. After receiving the configuration, the Coordination Service asks the Container Management Service to start the required number of instances of the Interpretation Service and the Calculation Service. As soon as the
required services are booted up, the Coordination Service calls the Evolutionary Operators Service to create the requested number of chromosomes of the initial population. At the end of an optimization job, the Coordination Service receives the aggregated result and sends it to the visualization component to be visualized. The Coordination Service does not act as a master in the global model, rather it coordinates the services by initialization and termination.

3.2.3 Evolutionary Operators Service. This service performs the task of the master in the global model. At first, it generates the initial population when called by the Coordination Service. Then, it calculates the fitness function to identify the individuals surviving for the next generation. Furthermore, it applies the genetic operators, namely crossover and mutation as well as the selection operation to generate the offspring.

3.2.4 Optimization Task Coordination Service. Optimization Task Coordination Service (formerly named as Distribution & Synchronization Service) coordinates one optimization task by e.g. assigning a Task ID, selecting one of the available simulation models that is available and starting and stopping an optimizing task. Indeed, it acts as a coordinator between the Evolutionary Operators Service and other services.

3.2.5 Splitting & Joining Service. The Splitting & Joining Service (formerly named as Distribution & Synchronization Service) receives the offspring, i.e. the chromosome list from the Evolutionary Operators Service. Afterwards, it evenly splits and distributes the population to the Interpretation Service instances. By finishing the distribution of the subpopulations successfully, the Interpretation Services start the interpretation processes by receiving a start signal from the Splitting & Joining Service. As soon as the optimization task is finished, the Splitting & Joining Service creates the overall result list matching the original list format supported by Evolutionary Operators Service by joining the partial results. Finally, the overall result is sent back to the Optimization Task Coordination Service which in turn sends it back to the Evolutionary Operators Service for applying the genetic operators, namely selection, crossover and mutation.

3.2.6 DERs Service. The DERs Service provides other services dynamic and static data about the DER components. Examples of dynamic data are the actual state of batteries, forecasting data for the generation of RERs, consumption and market prices which are continuously changed according to different factors such as the weather. Static data encompasses the number and type of DER components, technical constraints for the conventional energy resources e.g. minimum and maximum capacity, ramping limits and minimum up and down times, to name a few. Both types of data are stored in a database where each DER can insert and update its related data automatically, if it has an Energy Management System Interface (EMS-IF). Otherwise, a manual insertion and update is required. The Evolutionary Operators Service and the Interpretation Service instances need such data for the generation of the initial population and for the chromosome interpretation process as described later.

3.2.7 Interpretation Service. As its name implies, it is responsible for interpreting the chromosomes in the context of the optimization problem solved. For controlling DERs, the Evolutionary Operators Service generates scheduling operations represented by genes with relative values (e.g. in percent of the maximum provable power within a given time interval) representing the requested power share from each DER at specific time interval. These values must be interpreted by converting them to absolute values for evaluation (simulation) purposes. For example, for RERs, the relative generation values are multiplied by the corresponding forecasting data of the RERs to obtain the absolute values of a certain schedule. Since the interpretation process can require much computing time according to the size of chromosomes, the framework can deploy as many Interpretation Service instances as required allowing a parallel interpretation for scalability.

3.2.8 Calculation Service. The Calculation Service (or simulator) performs the calculations required to evaluate the individuals of the distributed population. It is called by the Interpretation Service for evaluating the offspring with respect to the given evaluation criteria. It takes a list of unevaluated individuals as the input and outputs the related evaluation results for each individual.

3.2.9 Container Management Service. The Container Management Service creates as many Interpretation Service and Calculation Service instances as needed allowing runtime scalability. After creating and initializing the required instances successfully, the Container Management Service publishes a ready signal in order to start the processing of the optimization job.

3.2.10 Supporting and Learning Service. Typically, EAs start to generate the initial population randomly which ensures the necessary diversity of the start population and allows for an initial breadth search. On the other hand, using a given solution of a similar task can speed up the search at the risk of pushing the search into the direction of these solutions. Thus, only a few prior solutions should be taken as a part of the initial population. This can significantly accelerate an EA [22]. This service supports the Evolutionary Operators Service by generating the initial population and can use some already-found solutions (i.e. scheduling plans in case of DERs scheduling) for this based on predefined selection criteria.

3.2.11 Local Search Service. The Local Search Service is an extension of a deployed EA to support Memetic Algorithms (MAs). This service provides the ability for using appropriate local search methods or heuristics to accelerate the evolutionary search of an EA by local improvement of the offspring. While other solutions exist, for complex and multidimensional optimization problems MAs can be especially useful for finding an optimal solution quickly [20].

RESTful service APIs are useful for enabling the communication among the services specifically the microservices that are non-scalable in runtime, namely the Coordination Service, the Evolutionary Operators Service, the Optimization Task Coordination Service, the Splitting & Joining Service, the Container Management Service and the DERs Service. However, the publish/subscribe pattern is used to realize the communications between the scalable microservices i.e. the Interpretation Service and the Calculation Service as well as between them and the other (non-scalable) microservices. The use of the publish/subscribe messaging paradigm ensures a seamless deployment and full decoupling among the services.
4 EA GLEAM FOR SCHEDULING DERS

The process of scheduling DERs consists of a set of scheduling operations that determine which DERs are involved in the power generation process and to what extent, in order to supply the required energy per time interval. The concrete EA GLEAM [6] is integrated into the Evolutionary Operators Service for scheduling DERs, as it has proven its suitability for general scheduling problems in several different applications e.g. [22]. GLEAM is acting as a master of the Global Model and generates the initial population, it applies the genetic operators and calculates the fitness value for each chromosome. The main feature that distinguishes GLEAM from other EAs is its flexible coding used to optimize not only time-dependent processes but also any other optimization problems such as scheduling and design optimization. The coding in GLEAM is based on a set of genes that are linked together forming a linear chain which represents a chromosome. The length of the chromosomes can either be fixed or altered dynamically by evolution. In the following section, the GLEAM based solution for chromosome representation and interpretation for scheduling DERs is described.

4.1 Solution Representation and Interpretation

Typically, a scheduling problem is broken down into several scheduling operations (e.g. one or more for each DER) which are represented by genes. In GLEAM, the structure of a gene is flexible and the number and types of its decision parameters are defined related to the nature of the optimization problem. The genes are moved as a whole by the respective genetic operators, which corresponds directly to the change in the sequence of the planning operations. Each scheduling operation is coded by one gene that consists of a fixed gene ID, which corresponds to the unit ID of the related DER, and the following decision variables: start time, duration and the power fraction as shown in Figure 4. While the start time is used to determine the start time of taking energy from this DER and the duration refers to the number of time intervals to which this setting applies, the power fraction variable determines the amount of energy in relation to the forecasted maximum that can be obtained from a DER. Since the number of required scheduling operations is not known a priori, the length of each chromosome is changed dynamically by the evolution. Mutation operators such as the duplication, deletion or insertion of individual genes or gene segments are used to alter the length of chromosomes (cf. [6, 21] for a detailed discussion).

4.1.1 Chromosome Interpretation. For the construction of an allocation matrix, the genes of a chromosome are successively treated so that a later gene overwrites matrix entries of the previous ones with the same Unit ID. This is considered as the first step of chromosome interpretation by the Interpretation Service. For each chromosome list, the first task is generating an allocation matrix where the number of rows m is equal to the number of resources, i.e. DERs in this chromosome list, and the number of columns n represents the time intervals. Figure 5 shows an example of the creation of an allocation matrix for a chromosome with three genes where the first and the third genes have the same Unit ID as shown in Figure 4. The third gene overwrites the first one at the time intervals from 10 to 15 as shown in Figure 5. When the building of the allocation matrix is finished, the Interpretation Service starts the second step of interpretation, namely converting the relative values of power fraction to absolute values by multiplying each value in the allocation matrix by the corresponding values of the actual maximum power generation supplied by the DERs Service for the corresponding time interval as shown in Figure 5. As a result, a new matrix with absolute values is produced and prepared for evaluation (simulation) by the Calculation Service.
5.1 Use Case Scenarios

For evaluating the scalability and generality of the proposed approach, the concept of hierarchical distributed energy resources introduced in section ?? is used to create three DER scheduling scenarios instrumenting a different number of DERs and DER mixes (only PV, PV with other generation sources or storage) with predefined generation behaviour, and two different load profiles as depicted in 6. For defining renewable generation behaviour, the hourly real power generation data for 50 and 100 PVs provided by Ausgrid [37] is used. Each DER has an EMS which manages and coordinates this DER. The EMS has a communication interface (EMS-IF) which provides flexibility of the DER in terms of the amount of energy that can be sold at a specific time interval with a specific price to consumers, e.g. the aggregated more or less controllable load (house symbols) located in blue DERs as shown in Figure 6. In the first scenario depicted in the upper left part of 

![Figure 6: Use Case Scenarios used for evaluation](image)

Figure 6, 50 DERs can offer power for 24 hours (not only PVs) to cover a simple load profile (load profile A) as shown in Figure 7. For the period between 8 and 17 o’clock, the EMSs offer the power to be sold from PVs and outside this period from other resources such as batteries or wind turbines. In the second scenario depicted in the upper right part of Figure 6, the same load profile (load profile A) as in the first scenario is used. However, only a part of DERs, namely 25 DERs can offer power for the consumer 24 hours from the PVs combined with other resources. The other 25 DERs have only PVs and therefore can offer power only for 10 hours between 8 and 17 o’clock. In the third use case scenario depicted in the lower part of Figure 6, 100 DERs provide the requested power for 24 hours to cover a more complex load profile (load profile B) as shown by the blue line in Figure 7. The main task of the distributed GEAM solution is to minimize the daily bill costs of the customer by generating the optimal hourly scheduling plan for one day ahead. Additionally, there are some constraints which have to be fulfilled.

5.2 Objective Functions and Constraints

For scheduling energy resources, there are three well-known operation modes, namely Cost-Effective Operation Mode, Eco-Mode and Robust Operation Mode as introduced in [2]. For the present evaluation, only the Cost-Effective Operation Mode is considered. Equation (1) defines the linear cost function to be minimized for the above three use cases:

\[\text{Cost} = \sum_{i=1}^{N} \sum_{t=1}^{24} C_{i,t} \cdot P_{i,t}\]  

(1)

where \(N\) is the number of DERs varied between 1 and 50 in the first and second use cases and between 1 and 100 for the third one, \(t\) is the time interval, \(C_{i,t}\) is the price in (EUR) for each kWh taken from resource \(i\) in time interval \(t\) and \(P_{i,t}\) is the scheduled power in kWh taken from resource \(i\) in time interval \(t\). DERs should work as much as possible by only using locally supplied power, the power balance within the blue DER of Figure 6 is considered an important optimization objective. Therefore, an additional objective function, namely the Daily Total Deviation (DTD) shown in Equation (2) is defined as the sum of absolute differences between the required power and the scheduled one at every time interval \(t\). For arriving at a local balance DTD should be as low as possible:

\[\text{DTD} = \left| \sum_{i=1}^{N} \sum_{t=1}^{24} P_{i,t} - D_{t} \right|\]  

(2)

where \(D_{t}\) is the requested power by the load in time interval \(t\) in kWh. To guarantee that the evolutionary search process preferably finds solutions without undersupply at each hour, the Hours of Undersupply (HU) shown in Equation (5), representing the number of hours of undersupply are defined:

\[\text{HU} = \begin{cases} 0 & \text{if } D_{t} > \sum_{i=1}^{N} P_{i,t} : t \in (1, ..., 24) \\ \text{otherwise} \end{cases}\]  

(3)

Due to the nonlinearity of the DTD function defined in Equation (2), the optimization problem is formulated as a nonconvex mixed-integer nonlinear optimization problem. Moreover, it is a multi-objective problem:

Minimize\([\text{Cost}, \text{DTD}]\]  

subject to \(\text{HU} = 0\)  

(4)

(5)

Since the scheduling of DERs is NP-hard optimization problem [42] and formulated as nonconvex mixed-integer nonlinear optimization problem based on Equation (2), its numerical solution is typically too
complex for exact optimization methods [7]. Hence, EAs represent a robust and powerful alternatives [16]. The presented solution using GLEAM should minimize the cost and DTD objective functions as far as possible while holding the constraint HU.

The Calculation Service is responsible for computing the values of the above objective functions (criteria) and constraint for each individual i.e. chromosome. For giving GLEAM a hint into what direction the populations should evolve, GLEAM uses (as other multiobjective solutions also) a weighting function called fitness curve which needs to be defined in the GLEAM configuration. The weighted sum defined in Equation (6) is used to combine the results of the criteria into a fitness value. For this purpose, the numerical value provided by the Calculation Services for each criterion must be mapped to a uniform fitness scale using one of standard normalization functions of GLEAM, namely linear, exponential and mixed linear-exponential. The fitness scale is set to a range between 0 and 100,000. The fitness value determines in GLEAM the likelihood of an individual reproducing and passing on its genetic information. This happens especially when choosing a partner and deciding whether to accept or reject the offspring when forming the next generation. For the cost and DTD criteria, the inversely proportional exponential function is used as shown in Figures 8a and 8b, respectively. In order to handle the equality constraint HU, a Penalty Function PF, which yields a value between zero and one (no undersupply) is defined. The fitness determined from the other two criteria is multiplied by this, so that an undersupply of 5 hours already reduces the fitness value to a third, see Figure 8c.

\[
\text{Fitness} = (0.4 \times \text{Cost} + 0.6 \times \text{DTD}) \times PF(HU) \tag{6}
\]

5.3 Deployment on a Cluster

For instrumenting the solution on a computer cluster, it is deployed on a cluster with four computing nodes where each node has 32 Intel cores (2.4 GHz), 128 GB RAM and an SSD disk. The nodes are connected to each other by a LAN with 10 GBit/s bandwidth. A modern software environment based on container automation technology guarantees a seamless deployment of the microservices on the cluster. For enabling containerization, the most popular open source software, namely Docker\(^1\) is used. Docker performs operating-system-level virtualization to isolate the applications. This is achieved by running containers on the Docker engine that separates the applications from the underlying host operating system. For container orchestration, Kubernetes\(^2\) is chosen as container orchestration system. It is used in many production environments due to its flexibility and reliability. Kubernetes defines several building blocks which are called Pods to separate “computing loads” from each other and provide mechanisms to deploy, maintain and scale applications. A Pod is the smallest building block in the Kubernetes object model and represents one or more running processes. Redis\(^3\) is deployed as an in-memory database serving as the Temporary Storage. Figure 9 shows the technological layers and an example of how the services can be mapped to the CPUs on the four nodes. While the Coordination Service and the Optimization Task Coordination Service are deployed on the first node, the Splitting & Joining Service and Container Management Service are deployed on the second node. The Evolutionary Operators Service is running on the third node and the DERs Service is running on the last node. The required instances from the Interpretation Service and the Calculation Service are distributed over all nodes dynamically.

---

\(^{1}\)https://www.docker.com/

\(^{2}\)https://www.kubernetes.io/

\(^{3}\)https://www.redis.io/
5.4 Results and Discussion
In the following, the efficiency of the proposed microservice and container-based approach in terms of the achieved quality of the schedules and the scalability for scheduling varying numbers of DERs in cluster environments are introduced and discussed.

5.4.1 Resulting Schedules. For achieving a good trade-off between exploration and exploitation, appropriate strategy parameters of the EA, namely the size of the population and the number of generations must be determined. For this, we perform several tests with 120 slaves and varying the population size as follows: 120, 180, 240, 300 and 420 individuals, so that each slave at minimum can process one individual. The number of offsprings per pairing is set to eight.

Table 1 shows the best obtained results. To limit the effort, the number of generations is set to 420. For the first use case an optimal schedule with 21 DERs—which is depicted in Figure 10a—, can be obtained with a population size of 180 individuals. For the second use case, GLEAM needs more individuals, namely 300 to explore the search space sufficiently to find an optimal scheduling plan—which is shown in Figure 10b—using more DERs namely 31. In comparison with the first use case, the number of scheduling operations (genes) and the corresponding number of evaluations are increased significantly. This is due to the fact that one half of the used 50 DERs are restricted to supply power only for 10 hours per day resulting in a more heterogeneous search space and a further constraint to GLEAM. Figure 11 shows how the required energy...
is covered by the two types of DERs considered in the second use case. As shown, the pure PV-DERs contribute with a generation portion between 13% (at 8 o’clock) and 54% (at 13 o’clock). For the third use case with 100 DERs and a more complex load profile, a scheduling plan –shown in Figure 10c– is found with a population of size 240 individuals. It is noteworthy that in the use cases 1 and 3, the number of evaluations increases by only 69% if the number of considered DERs are doubled.

5.4.2 Framework Scalability. In order to assess the performance of the proposed software architecture, we analyze the scalability of the framework for the above three use case scenarios introduced in section 5.1. The number of computing units (slaves), namely the Interpretation Service instances as well as the number of the corresponding Calculation Service instances is varied between 1 and 120 so that the minimum of two cores is left on each node for the OS.

Table 2 shows the scalability results of the three use cases where the total time for each optimization job is measured. It can be concluded that by increasing the difficulty of the optimization problem, the total time needed to find an optimal solution is increased. Therefore, the scheduling process for the second and third use cases takes more time as the first one since GLEAM performs more evaluations. By using more computing units, the framework is able to reduce the total time from 780 to 38 minutes in the first use case, from 1290 to 66 minutes in the second use case and from 4175 to 200 minutes in the last use case. For each use case, the computation time of the parallel implementation decreases more slowly at a certain point, since the communication overhead of the increased number of computing units (slaves) exceeds the increased performance of the parallelization.

6 CONCLUSION AND FUTURE WORK

In this paper, a new highly modular, flexible and scalable method for scheduling DERs based on EAs is presented. In contrast to other optimization methods, the new solution enables an efficient parallelization of EAs, full runtime automation and an easy deployment on high performance computing environments such as clusters or cloud environments. Furthermore, it provides the ability to deal with a comparably large number of distributed energy resources. Modern software technologies, namely microservices, container virtualization and the publish/subscribe messaging paradigm are exploited to develop the desired method. The architecture clearly separates functionalities related to EAs and the ones related to scheduling DERs. For each functionality, a microservice is designed and implemented. Furthermore, container virtualization is utilized to automatically deploy the microservices on nodes of an underlying cluster to perform their tasks. The combination of microservices and container virtualization enables an easy integration of an existing EA like GLEAM into the framework and facilitates the communication with other required services like generation, consumption, market price and weather forecasting tools and simulators. Furthermore, using the publish/subscribe messaging paradigm guarantees a seamless data exchange between the scalable services which are deployed on-demand.

In order to evaluate the functionalities of the proposed solution, three use case scenarios with different types and numbers of DERs are defined and studied. The scalability of the framework is demonstrated by varying the number of computing units between 1 and 120. The results show that the new distributed solution is an applicable approach for scheduling a scalable number of DERs using EAs based on the mentioned three lightweight technologies in a scalable runtime environment.

As part of future work, more detailed evaluations related to the communication overhead of the solution will be undertaken. Other parallelization models for EA such as Coarse-Grained Model can also be applied and compared with the current presented approach. Furthermore, the Supporting and Learning Service mentioned could be implemented to accelerate the EA based on learned knowledge about prior solutions.
REFERENCES

[1] Enrique Alba, Francisco Almeida, Maria Blesa, Carlos Cotta, Manuel Diaz, Isabel Doria, Joaquim Gabarró, Coromoto León, Gabriel Luque, Jordi Petit, et al. 2006. Efficient parallel LAN-WAN algorithms for optimization. The MALLBA project. Parallel Comput. 32, 7-8 (2006), 415–440.

[2] Lázaro Alvarado-Barrios, A Rodríguez del Nozal, A Tapia, José Luis Martinez-Ramos, and D.G Reina. 2019. An evolutionary computational approach for the problem of unit commitment and economic dispatch in microgrids under several operation modes. Energies 12, 11 (2019), 2143.

[3] Maribel García Arenas, Pierre Collet, Agoston E Eiben, Márk Jelasity, Juan J Merelo, Ben Faesche, Mike Preuß, and Marc Schoenauer. 2002. A framework for the reusable design of parallel and distributed metaheuristics. In International Conference on Parallel Problem Solving from Nature. Springer, 665–675.

[4] Alireza Askarzadeh. 2017. A memory-based genetic algorithm for optimization of power generation in a microgrid. IEEE transactions on sustainable energy, 9, 3 (2017), 1081–1089.

[5] Cristóbal Barba-González, José García-Neto, Antonio J Nehoro, José A Cordero, Juan J Durillo, Ismael Navas-Delgado, and José F Aldana-Montes. 2018. mTAPen: a framework for dynamic multi-objective big data optimization. Applied Soft Computing 69 (2018), 737–748.

[6] Christian Blume and Wilfried Jakob. 2002. GLEAM-An Evolutionary Algorithm for Planning and Control based on Evolution Strategy. In GECO Late Breaking Papers. 31–38.

[7] Vijay Madisetti. 2004. PARSEC: A framework for the reusable design of parallel and distributed metaheuristics. Journal of heuristics 10, 3 (2004), 357–380.

[8] Peter Brucker. 2007. Scheduling Algorithms, Springer.

[9] Sébastien Cahon, Nordine Melab, and E-G Talbi. 2004. Paradiso: A framework for the reusable design of parallel and distributed metaheuristics. Journal of heuristics 10, 2 (2004), 141–171.

[10] F. M. Cleveland. 2008. IEC 61850-7-420 communications standard for distributed energy resources (DER). In IEEE Transactions on Energy Conversion. 1202–1208.

[11] Juan J Merelo Guervós and J Mario García-Valdeví. 2018. Introducing an Event-based Architecture for Distributed Energy Resource Management. In International Conference on Parallel Problem Solving from Nature. Springer, 399–410.

[12] Yuansheng Guo, Yuguang Fang, and Pramod P Khargonekar. 2017. Hierarchical Architecture for Discrete Energy Resource Management. In Stochastic Optimization for Discrete Energy Resources in Smart Grids. Springer, 1–8.

[13] William E Hart, Natalio Krasnogor, and James E Smith. 2004. Handbook of heuristic optimization: in application, design, and prospects. IGI Global, 1031–1040.

[14] Elizabeth L Ratnam, Steven R Wellers, Christopher M Kellett, and Alan T Murray. 2009. A cost-effective dynamic pricing strategy optimization procedure to solve the nonconvex economic dispatch problem with generator constraints. IEEE Transactions on Power Systems, 24, 2 (2009), 939–946.

[15] Andrija Volkanovski, Borut Mavko, Tome Boševski, Anton Čauševski, and Marko Keržen. 2017. Towards coding strategies for forecasting-based scheduling in microgrid. In International Conference on Parallel Problem Solving from Nature. Springer, 1031–1040.

[16] Chao Jin, Christian Blume, and Rajkumar Buyya. 2008. MRPGA: an extension of MapReduce for parallelizing genetic algorithms. In 2008 IEEE Fourth International Conference on EScience. IEEE, 224–225.

[17] Spyros A Kazarlis, AG Bakatiras, and Vassilios Petridis. 1996. A genetic algorithm solution to the unit commitment problem. IEEE transactions on power systems 11, 1 (1996), 83–92.

[18] Hao Quan, Dipti Sriivasan, and Abbas Khosravi. 2014. Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals. IEEE transactions on neural networks and learning systems 26, 9 (2014), 2123–2135.

[19] Jorge Valenzuela and Alice E Smith. 2002. A seeded memetic algorithm for large enhanced genetic algorithms. In 2002 IEEE Eindhoven PowerTech. IEEE, 1–6.

[20] Mohsen Nemati, Karima Benrimar, Stefan Tenbohlen, Liang Tao, Holger Mueller, and Martin Braun. 2015. Optimization of microgrids short term operation based on distributed multiagent evolutionary algorithm. In 2015 IEEE Eindhoven PowerTech. IEEE, 1–6.

[21] Mohsen Nemati, Martin Braun, and Stefan Tenbohlen. 2018. Optimization of unit commitment and economic dispatch in microgrids based on genetic algorithm and mixed integer linear programming. Applied energy 210 (2018), 944–963.

[22] K Rajmahon Padvey and Anil M Kulkarni. 2019. Dynamics and Control of Electric Transmission and Microgrids. Wiley Online Library.

[23] Hao Quan, Dipti Sriivasan, and Abbas Khosravi. 2014. Incorporating wind power forecast uncertainties into stochastic unit commitment using neural network-based prediction intervals. IEEE transactions on neural networks and learning systems 26, 9 (2014), 2123–2135. https://doi.org/10.1109/TNNN.2014.2334537

[24] Elizabeth L Ratnam, Steven R Wellers, Christopher M Kellett, and Alan T Murray. 2009. A cost-effective dynamic pricing strategy optimization procedure to solve the nonconvex economic dispatch problem with generator constraints. IEEE Transactions on Power Systems, 24, 2 (2009), 939–946.

[25] Darrell Whitley. 1994. A genetic algorithm tutorial. Statistics and computing 4, 2 (1994), 65–85.

[26] Muhammad Fahad Zia, Elhoussin Elbouchikhi, and Mohamed Benouzid. 2018. Microgrids energy management systems: A critical review on methods, solutions, and prospects. Applied energy 222 (2018), 1033–1055.