Available online at www.sciencedirect.com

ScienceDirect

Energy Procedia 87 (2016) 157 – 164

Energy Procedia

5th International Workshop on Hydro Scheduling in Competitive Electricity Markets

Use of parallel processing in applications for hydro power scheduling – current status and future challenges

Geir Warland, Arild Lote Henden, Birger Mo *

SINTEF Energy Research, Energy systems, Sem Sælands vei 11, Trondheim 7034, Norway

Abstract

This paper gives an overview of models for hydro power scheduling and market simulation developed at SINTEF Energy Research. All of the models are in operative use by power producers, transmission system operators, consultants and regulators operating in the Nordic power market. Several of the models have been adapted to use of parallel processing to decrease computation time. The paper gives an overview of existing models where parallel processing has been applied.

Finally future challenges and the need for higher level of parallelization are discussed and exemplified by two new models for power market simulation that use two level parallelization. Both of these models have the potential for efficient utilization of hundreds or even thousands of processors (or cores) and can be run on large compute clusters.

Peer-review under responsibility of SINTEF Energi AS

Keywords: Hydro power applications; hydro-thermal scheduling; power market simulation; parallel processing

1. Introduction

Operation of hydro systems involves a range of different tasks. Power producers have to find the optimal use of hydro resources and manage the economic risk of operation. Power producers and transmission owners perform system analyses and plan system expansion, refurbishment and other investments and plan for maintenance. All these tasks require decision tools for scheduling of hydro resources.

Hydro-thermal scheduling is a very complex task where the purpose is to find the optimal use of the available resources over a given study period. In purely thermal systems the scheduling problem is primarily decoupled in

---

* Geir Warland. Tel.: +47 -73-59-72-00 fax: +47-73-59-72-50.
*E-mail address:*geir.warland@sintef.no

Peer-review under responsibility of SINTEF Energi AS
doi:10.1016/j.egypro.2015.12.346

time and each generating unit have a direct operating cost and this cost does not depend on the costs of other plants. In hydro-thermal systems, particularly when the reservoirs have significant storage capacity, the problem becomes coupled in time. The hydro plants can use the "free" energy stored in the reservoirs to meet demand thereby avoiding the cost of using thermal units. The hydro resource is however limited by the storage capacity of the reservoirs and inflow, and this introduces a time dependency in operating costs. It can either be used today to reduce current operating costs by reducing use of thermal units today, or stored for use at later stages to reduce future operating costs. Thus the water resource, although having no direct costs, has an indirect opportunity cost, also called water value (WV). The operating decision for hydro power is to produce today or store for future use.

There are several future uncertainties that will have influence on the operation of the hydro system. Especially the future inflow is important for the operation of the hydro system, but also uncertain parameters like temperature, fuel prices and supply of other energy sources may have a large impact.

In addition to the dynamic and stochastic nature of the hydro-thermal system the problem may be even more complex due to factors such as losses, transmission limitations, start-up costs of thermal as well as hydro plants, ramping rates of plants and transmission, topology and time delays. All this make the decision problem a very complex and time consuming task to solve.

This paper gives an overview of computer applications for hydro-thermal scheduling and power market simulation developed at SINTEF Energy Research. All applications are in use by commercial market actors in the Nordic power market. The paper will also present some details on how parallelization has been introduced and further present some experience and thoughts on future development using multilevel parallelization.

2. Parallel processing

Traditionally, software has been written for serial computations. The problem is broken into discrete series of instructions which are executed sequentially one instruction at a time on a single processor as illustrated in Fig. 1. a)

Almost all new computers and all servers are parallel from a hardware perspective. Currently most CPU's have multiple execution units or cores and servers have multiple CPU's installed. Also other hardware technologies, e.g. internal memory and intercommunication allow parallel access. Servers can also be set up to communicate on a local network in parallel compute clusters.

If a given problem can be broken into separate parts which can be solved independently, this problem can be solved by using parallel computing to speed up calculations. Each part is broken into series of instructions, and the instructions for a part are executed on a given processor/core as is illustrated in Fig. 1 b).

There are several models for parallel programming in common use, for example threads using shared memory and distributed memory using message passing. Distributed memory and message passing is typically a manually developed parallel code, while threads and shared memory can be either automatically, imbedded in the compiler, or manual using e.g. OpenMP. This is an industry standard available on many platforms and programming languages.

In computer code presented in this paper, message passing is primarily applied using a standardized and portable message passing system called Message Passing Interface (MPI) [1]. This standard supports both point-to-point and collective communication between concurrent processes and is one of the most used models for high-performance computing today. Point-to-point communication is used to send messages between individual processes. Typically in applications described in this paper this is applied to communicate between a master process and a slave process. The master process is a process dedicated to administrate the parallel computations while slave processes are dedicated to solve different parts of the calculations. Collective communication is used to send data from the master process to all the slave processes, sending results from all slave processes to the master process or sending data from all processes to all other processes.

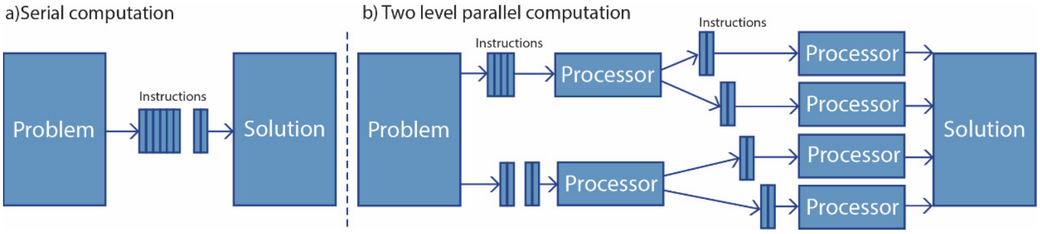


Fig. 1. Illustration of serial and parallel computations

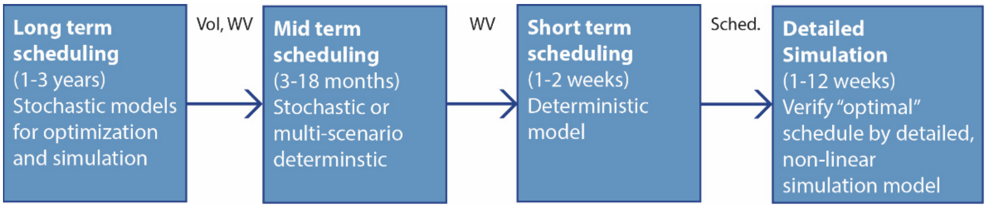
3. Hydro power scheduling hierarchy

In principle the short and long-term use of water in a hydro-thermal system could be solved as large stochastic optimization problem. However, this is in practice not feasible. The hydro system may comprise of tens, hundreds or even thousands of hydro reservoirs and power plants. There are several stochastic optimization techniques that can be applied to solve the problem, but all have a practical upper limit of how many states (or reservoirs) they can handle with a reasonable computation time. Thus the overall decision problem is decomposed in different stages. Usually the problem is decomposed into long-term, medium-term and short term scheduling, each being solved by different models and solution techniques [2]. This is shown in Fig. 2. Typically the computation time for the different stages differ, and reasonable computation times are in seconds, minutes or tens of minutes depending on which stage. Computation times must be faster the closer one comes to the operations.

3.1. Long term scheduling

The long term hydro scheduling finds the optimal strategy for the hydro system operation, and can be used directly for scheduling of hydro resources in a market environment. The long term scheduling normally has a planning horizon of more than one year. The planning horizon needed depends on hydro system characteristics.

Models for long term scheduling are typically stochastic models and often involve simplifications and approximations in order to make computation times acceptable, e.g. aggregation of hydro representation and few time increments. The stochastic parts in long term optimization are often based of stochastic dynamic programing (SDP) and stochastic dual dynamic programing (SDDP) [3]. The result of these simplifications is that the long term scheduling will not directly provide the boundary conditions for the short term models.



3.2. Medium term scheduling

The medium term scheduling is often seen as a link between the long term and short term scheduling, and provides boundary conditions for the short term scheduling. The planning horizon is shorter than for the long term scheduling, typically from a few months to one year. In the medium term scheduling the hydro description is approximately similar to that of the short term scheduling while the time increments is often equal to that of the long term. The medium term use multi-deterministic optimization or stochastic models (SDDP) [4].

Fig. 2. Hydro scheduling stages – decomposition principle.

3.3. Short term scheduling

Short term scheduling solves a deterministic optimization problem by using linear programming (LP), successive linear programming (SLP) or mixed integer programming (MIP). In a newer prototype version, stochastic optimization is also included. In this model the time increment is usually more detailed, typically with hourly time steps or less, and all physical parts which influence power production in a water course are described in detail. The coupling to the medium term model is based on a water value description for all reservoirs depending on price, reservoir level and week. The model is typically used to make bidding curves to market before clearing and to optimize the production contract after clearing.

3.4. Detailed simulation

Detailed simulation using non-linear physical description of the power system is used to verify the solution from short term optimization and to show how the system responds to unforeseen events.

4. Current computer software using parallel processing

4.1. SINTEF Energy Research's applications

SINTEF has developed several computer models for power market simulation and hydro power scheduling in hydro-thermal systems. The objective function for the models is minimization of operating costs which is the same as maximization of socio economic surplus. The hydro modelling varies from aggregate to fully detailed representation. The computer models cover all the aspects presented in chapter 3. The most important are presented in Table 1.

Table 1. Overview of operational models made and developed of SINTEF Energy Research

Application	Term	Description	Problem	Method	Parallel
EOPS	Long and medium	Single area hydro-thermal scheduling. Scheduling, use of reservoirs and expansion planning.	Stochastic	Optimization (SDP) and heuristic	No
EMPS	Long and medium	Multi area hydro thermal market model. Price forecasting, planning, expansion and power system analyses.	Stochastic	Optimization (SDP) and heuristic	Yes
Samlast & Samnett	Long	EMPS with physical power flow constraints.	Stochastic	Optimization (SDP) and heuristic	No
Seasonal model	Medium	Calculate individual water values, operational decisions, or input to short term model (SHOP).	Multi-deterministic	Optimization (LP)	Yes
ProdRisk	Long and medium	Single area hydro-thermal scheduling. Scheduling, use of reservoir, expansion planning and water values for short term model (SHOP).	Stochastic	Optimization (SDDP)	Yes
SHOP	Short and detailed simulation	Single water course. Scheduling, power market trade. Also include simulator for validation of the optimization.	Deter-ministic	Optimization (SLP, MIP) or simulation	No (Yes in MIP)

These models are used by the Norwegian regulator (NVE), most of the power producers, transmission system operators and several consultants in the Nordic area. SHOP and EMPS have also been sold to power producers outside the Nordic power market. Parallel processing using MPI has been applied to several of these applications.

These are discussed in the following subsections. Some LP/MIP-solvers also support parallel processing, but this is not discussed in the paper.

4.2. EMPS

4.2.1. Model concept

The model solves long term hydro-thermal optimization and simulation in two steps; strategy evaluation and simulation of operational detailed hydro decisions [5]. In the strategy evaluation, incremental water values (marginal costs for hydropower) are computed for each market area using a combination of SDP and a heuristic approach for treating the interconnection between areas. The strategy evaluation part also includes simulation of system operation and a convergence check.

In the simulation part hydro and thermal generation decisions are determined in a market clearing process based on the incremental water values found in the strategy part. Hydro production for individual plants is found using a rule-based reservoir drawdown heuristic for a detailed hydro system representation.

4.2.2. Parallelization of the strategy evaluation

The strategy evaluation contains three separate operations that can be parallelized:

- Water value calculation which is performed separately for each area using SDP. Typically the number of areas (N_{area}) is in the range of a few tens of areas. Parallelization is here applied per area. These calculations can be speeded up even further by applying parallel processing in the SDP-method but this has thus far not been implemented.
- A simulation of system operation using N_{year} historical inflow scenarios, where each scenario is independent of the others. Parallelization is applied per historical inflow scenario.
- The convergence test involves solving N_{task} independent subtasks where N_{task} is significantly larger than N_{area} and N_{year}. Parallelization is performed per subtask.

These three operations are executed in sequence. The overall efficiency and the optimal number of processors that can be applied for parallel processing will depend on the computational burden and timing of each operation. In the first operation only N_{area} processors will be fully utilized (no master process), in the second part and third part a maximum of N_{area} +1 and N_{task} +1 respectively can be fully utilized ($N_{area} < N_{year} << N_{task}$). A master – slave concept is applied for both of these operations.

Until now the water value calculations and the simulation part has been the dominant part in terms of calculation time, thus the number of processors applied have been in the range between N_{area} and N_{year}. This has, however, changed as the modelling details have increased; new functionality, finer timer resolutions, etc. Thus the number of processors that can efficiently be utilized has increased slightly.

4.2.3. Parallelization of the simulation

The simulation of detailed hydro operation in EMPS is in terms of parallelization no different from the simulation using aggregate hydro representation in the strategy evaluation. The only difference is the use of the rule-based reservoir drawdown model. Each historical weather scenario is simulated in sequence, week by week, on a single process, but all (N_{year}) weather scenarios are independent. The simulation is parallelized using N_{year}+1 processes, one master process and N_{year} slave processes.

4.3. Parallelization of the multi-deterministic medium term model

The multi-deterministic model for seasonal planning in EMPS, uses LP to find individual water values for each reservoir in the system for a given initial reservoir. The water value for a given reservoir is given by the shadow price (from the LP-solution) for the actual reservoir constraint. The water value is found as the average for all N_{year} weather scenarios, by solving independent LP's for each watercourse and weather scenario. The total number of watercourses is N_{water}. The water values can be found for a large number (N_{init}) of initial reservoirs. The total number of independent LP problems is $N_{init}*N_{water}*N_{year}$. A master-slave scheme is used to solve these problems using parallel processing.

4.4. ProdRisk – SDDP based application for medium and long term scheduling

ProdRisk is a model for long and medium-term hydro scheduling for a general detailed hydro representation using formal optimization (SDDP) [4]. Electricity market prices can either be described endogenously, i.e part of the solution, or the prices can be given exogenously which is the case for normal use of the model. In both cases the model consists of two parts; a backward recursion and forward simulation. In the backward recursion, the weekly decision problem (formulated as an LP problem) is solved for each exogenous price nodes $N_{price,}$ for all state combinations N_{scen} simulated in the previous forward pass for all discrete inflow combinations N_{inflow}. In total this gives $N_{price} * N_{scen} * N_{inflow}$ LP problems for each week in the backward pass. In our implementation the N_{scen} loop is parallelized in the backward pass. Therefore, in theory N_{scen} parallel processors could be utilized. N_{scen} typically ranges from 50 to 240, depending on the length of the inflow records and availability of exogenous price scenarios. In the forward pass the N_{scen} scenarios are simulated with perfect decoupling and can therefore easily be parallelized.

4.5. User experience

The parallel processing versions of the computer software are used by several users operationally; a few users have approximately 10 years of experience. They have typically been using servers with up to 16 or 32 cores. Some customers have recently also acquired compute cluster and have lately been running the EMPS model and medium term multi-deterministic model with up to 100+ cores.

One way to evaluate the efficiency of parallel processing is to measure by a scaling factor $E = T_s / (N*T_p)$, where T_s is the serial computation time, T_p is parallel computation time and N is number of processes. Ideally this number should equal 1, but will often be lower due to aspects discussed in the next chapter. Adequate scaling is often assumed if $E = 1/\log(N)$. The EMPS simulation has given scaling factors in this range for number of processors up to around half the number of historical scenarios (N_{year}). The scaling factor will decrease by further increasing the number of processors. The scaling of the simulation has been considered acceptable by the users. Also the parallelization of the strategy evaluation has been considered acceptable by the users even though here scaling factors close to 1 is harder to achieve due to the difference in number of tasks to be solved in the different parts of the process. Use of parallelization has reduced calculation times in EMPS from hours to minutes.

The parallelization of the medium term multi-deterministic model has been considered adequate by customers. This model has the potential of utilizing a vast number of processors but has thus far only been tested up till around 80 processors with adequate results (with scaling efficiency around 0.5). Similarly as for the EMPS, calculation time has been reduced from several hours to minutes.

The ProdRisk model is the model using parallel processing with the highest number of users. This model does not scale well for more than 12 processors. To increase calculation speed and get acceptable scaling for a larger number of processors, a new parallel scheme will be implemented in the future, probably by introducing a master–slave concept.

5. Parallel development experience

More than ten years of implementation experience, using parallel processing, have given insight and experience, especially regarding parallelization of existing computer code. There are several possible pitfalls. The most obvious ones, yet still easily missed, are deadlocks and file access. Deadlocks will appear when several or all parallel processes have to reach a given synchronization point in the computer code but one or more fail to reach this point. File access is perhaps the most common cause of problems. This is linked to the problem of handling simultaneous access to file from multiple processes, and especially regarding writing to file. This can be handled by dedicating one process to writing data to files, but it does introduce the need to send large portions of data between processes. Communication and synchronization due to transfer of data between processors are some of the reasons why ideal scaling of computation is nearly impossible. It is vital to reduce the amount of synchronization in order to preserve good scalability. Collective communication may also in some cases be the cause of unnecessary synchronization as this communication will block further progress until all processes have reached this synchronization point.

All parallel code requires a minimum of communication between concurrent processes. The number of communications should be kept fairly low. This can be achieved by aggregating and sending larger portions of data. There may exist an optimal balance between package size and number of communications but this has not been investigated in further detail, but as a rule of thumb fewer and larger packages has been preferred. This will probably become more important when running compute cluster and the number of parallel processes increases.

One possible pitfall when introducing parallel processing in existing serial computer code is too high memory usage. This is specially the case for parallel code used on large compute cluster. Large effort to reduce memory use is vital. It is easier to achieve a goal of low memory usage per process when designing new computer software.

Scaling can also be impeded by unavoidable serial processing e.g. input and preprocessing of data needed by many of the processes. It is crucial that these tasks are made as efficient as possible. Another factor that limits good scaling is the fact that many models consist of separate sub-tasks that has to be executed in sequence. The degree of parallelization of each of these sub-tasks may vary considerably, as is the case for the EMPS–model.

In many of the models described above LP is a crucial part of the algorithm. In these cases the total calculation time may be highly affected by the solution time of very large LP's. Further reduction will in these cases often require improvements in the LP-solver, e.g. by using LP-solvers with parallel options when available (e.g. Barrier solver). For computer code solving sequences of LP problems, the distribution of sub-tasks between parallel processes is not random. The code should take into consideration the possibility of using warm-start in LP-solvers as this gives significantly shorter solution times.

In many cases parallel processing may benefit from parallelization in several levels, e.g. by running large sub-tasks in parallel where also each sub-task can be solved separately using parallel processing. This has not yet been applied in the models described above, but is exploited in new prototypes described in the next chapter.

6. Resent development and future challenges

The trend for hydro power scheduling, at least in the Nordic area, is to use finer time resolution in medium and long term hydro-thermal scheduling. With that more details may need to be included in the representation of the hydro and power market description. Examples of such details might be ramping rates, time delays in water courses, start- and stop constraints on hydro plants as well as thermal units. These modelling details and increased time resolution result in larger models and gives a large increase in calculation time. In many models there are unexploited possibilities of increasing the level of parallelization. By introducing parallelization in several levels of the calculations a larger number of processors may be utilized with potentially a significant increase in the size of problems that can be solved with reasonable time consumption. Two models for power market simulations currently under development as SINTEF Energy Research using a two level parallelization are briefly described below.

6.1. SOVN

The SOVN model simulates a sequence of problems referred to as scenario fan problem (SFP) [6]. The model solves the market clearing process for the whole power market with detailed representation of hydro. Each problem is, in principle, a large LP which is a deterministic equivalent representing the present and future states of the hydro-thermal power system. This problem includes coupling of hydro power production in time and space. It represents the future uncertainty using weather scenarios while the first week uses a given realization of the stochastic variables. The model simulates a large number of weeks for N_{year} independent weather scenarios, and the simulation of SFP is solved using parallel processing. At this level $N_{year}+1$ parallel processors can be utilized.

Each SFP is a two-stage stochastic LP problem. The first stage represents a given week of the simulation (weather scenario and week). The second stage is represented by N_{year} independent scenarios. The SFP is solved as a sequence of LP-problems, where the first stage is solved first providing the initial reservoir for the N_{year} independent future scenarios that represent the second stage. In the next iteration the first stage is re-solved with Benders cut representing the boundary conditions for the future. The solution of SFP is also solved using parallel processing, where one process solves the first stage and N_{year} processes solve the future scenarios.

The total number of parallel processors that can be utilized is $1 + N_{year} * (1 + N_{year})$. For typical cases N_{year} may in the range of 50–80. Thus the number of parallel processes can be very large. The model has been tested, with

N_{year}=8 weather scenarios giving a total of 73 parallel processes, on a compute cluster with 4 Compute nodes, each having two 10 core CPU's. Due to the sequential dependency in the solution of the SFP-problem, perfect scaling will not be obtained, but the parallelization scheme has given a huge reduction in calculation time.

6.2. ProdMarket

This model solves the hydro-thermal market problem, same as solved by EMPS and SOVN, using price decoupling and the SDDP methodology in and iteration process. Simulated market prices are used to decouple the overall problem into individual hydro scheduling problems that are solved for individual watercourses. This optimization is solved as an iteration process with a strategy phase and a simulation phase. In the strategy phase ProdRisk (see section 4.4) is used to optimize each water course for an exogenously given stochastic price. Results from the strategy phase are cuts for all N_{water} individual water courses. These cuts are used in the simulation part, where the whole power system is described. The potential for parallelization in the simulation part of ProdMarket is Nyears weather years. The simulation phase calculates new prices. The algorithm has converged when the difference between the new prices and the prices calculated in the previous iteration is within a given threshold.

ProdMarket is implemented using two-levels of parallelization. At the first level, all N_{water} water courses are solved in parallel. At the second level, each water course is solved using a parallel processing version of ProdRisk. ProdMarket is implemented using a combination of Fortran and Python, both programming languages having support for parallel processing. Parallel processing is applied in Python to solve the individual hydro scheduling problems. These are solved using parallel processing in ProdRisk. The potential for this model to utilized parallel processors is in principle almost unlimited. In practice however, with the current ProdRisk implementation, each water course can only efficiently utilize up to about twelve processors. A typical dataset of the Nordic hydro-thermal system may consist of more than 100 water courses. The solution time of the complete model is limited by the solution time of most complicated hydro system, and the number of available processors. In the market simulation phase where new prices are calculated based on the updated strategy the number of processes that can be utilized is limited by the number of historical weather years that are simulated.

Our experience is that the Python programming language makes it relatively easy to implement parallel processing based on existing program kernels that may themselves be parallelized. In our case, price decoupling is used to split a huge problem into smaller problems that are solved in parallel using existing models and administered using the Python language.

7. Summary

Parallel processing is the key to efficient handling of the complexity and computation times for hydro-thermal scheduling models. Our current commercial applications utilize only parallelization in one level, and have come close to their potential regarding efficiency. To further increase computation efficiency, and to fully utilize the present and future hardware having multiple compute-nodes in large compute cluster, and vast numbers of compute cores multilevel parallel processing needs to be applied.

References

[1] Gropp W., Lusk E., Thakur R., "Using MPI-2, Advanced Features of the Mesage-Passing Interface", Scientific and Engineering Computation Series, The MIT Press, Cambridge, Massachusetts, London, England, 1999.

[2] Fosso O. B., Gjelsvik A., Haugstad A., Mo B. and Wangensteen I.: “Generation Scheduling in a Deregulated System: The Norwegian Case”, IEEE Transactions on Power Systems 14, pp. 75-81, 1999.

[3] Pereira M.V.F., Pinto L.M.V.G., "Multi-stage stochastic optimization applied to energy planning", Math. Program. 52, pp.359-375, 1991

[4] Gjelsvik A., Mo B., Haugstad A., ”Long- and Medium-term Operations Planning and Stochastic Modelling in Hydro-dominated Power Systems Based on Stochastic Dual Dynamic Programming”, Handbook of Power Systems I, pp. 33-56, Springer Berlin Heidelberg, 2010.

[5] Wolfgang O., Haugstad A., Mo B., Gjelsvik A., Wangensteen I., Doorman G., ” Hydro reservoir handling in Norway before and after deregulation”, Energy 34, pp 1642–1651, 2009.

[6] A. Helseth, B. Mo and K. S. Gjerden. "On Designing a Stochastic Optimization Model for Detailed and Long-Term Hydro-Thermal Scheduling in the Nordic Power Market", XIII SYMPOSIUM OF SPECIALISTS IN ELECTRIC OPERATIONAL AND EXPANSION PLANNING, May – 18th to 21st, FOZ DO IGUAÇU (PR) – BRAZIL, 2014.