Cloud computing for complex geothermal simulations using the parallel software Waiwera

Ken Dekkers[1], Adrian Croucher[1], Michael Gravatt[1], Theo Renaud[1], Ru Nicholson[1], Oliver Maclaren[1], John O'Sullivan[1], Michael O'Sullivan[1]

[1]Geothermal Institute, University of Auckland, 70 Symonds Street, Grafton, Auckland, 1010, New Zealand

ken.dekkers@auckland.ac.nz

Keywords: *Cloud Computing, High-Performance Computing, Waiwera, Uncertainty Quantification, Resource Assessment*

ABSTRACT

Numerical modelling of geothermal fields has been around for a few decades and has proven to be very insightful for sustainably managing geothermal projects. It is a great, cost-effective tool for evaluating various development plans using a realistic model of the field.

Simulating geothermal fields, however, is a time-consuming process, especially when the model become more complex and refined, and the simulation time increases. The geothermal simulator Waiwera, recently developed by the University of Auckland (UoA) in collaboration with GNS Science, has proven to be able to reduce simulation time drastically by using parallel computing on standard desktops, and even more so when using high-performance (cloud) computing. Simulation times reduce by a factor of 10-50 depending on the complexity and size of the model. Additionally, uncertainty quantification of a model, which requires running thousands of sample models, benefits hugely from high-performance cloud computing.

In this paper we explore three cloud computing platforms: NeSI (New Zealand eScience Infrastructure), AWS (Amazon Web Services) and Azure (Microsoft). Different options of user interaction with the cloud computing platforms will be discussed. Moreover, the simulation speed and costs will be compared for these three platforms.

1. INTRODUCTION

1.1 Background

In the early stages of numerical modelling of geothermal systems TOUGH (Pruess, 1987) was widely used, followed by TOUGH2 (Pruess, 1991), which is currently still one of the benchmark software for numerical geothermal flow simulations. These software packages capture the state of the whole geothermal reservoir by solving the energy and mass conservation equations and Darcy's law over time.

In these early stages of geothermal reservoir modelling the computation time using a single-core sequential software like TOUGH2 was acceptable. The geothermal models were relatively simple with a manageable number of blocks in the model grid (say, up to ~20,000). Simulation times of both the natural state (steady-state, pre-production) and the production history could range between a few minutes to a few hours, greatly dependant on the complexity and grid size of the model.

In the world of geothermal reservoir modelling, it is crucial to develop a model that represents the real geothermal reservoir accurately, which in turn gives confidence in the model predictions. Creating an accurate model is done by matching the model to the measured data, like temperature and pressure data. This calibration process can take tens to hundreds of iterative steps, changing one or more parameters at a time. A computation time of a few minutes is acceptable, but when each simulation takes a few hours, it may lead to an unacceptable timeline for a geothermal modelling project.

1.2 Waiwera

A breakthrough in geothermal modelling software was made with the development of Waiwera, a parallel open-source geothermal flow simulator, developed at the University of Auckland in collaboration with GNS Science and with support from Contact Energy Ltd (Croucher, 2020). This software solves the same mass and energy conservation equations as TOUGH2 but is now able to do this in parallel using multiple CPUs.

This opened the door for more complex geothermal models with more mesh refinements, which could represent the geothermal reservoir more accurately. Apart from computational speed up on a local PC, simulations could now also be run on HPC (high-performance computing) resources on the cloud. This option gives access to a virtually infinite number of computational processors. However, one should consider the diminishing returns of adding too many processors. This criterion varies with the reservoir model. It depending on the complexity and grid size of the model and communication speed between processors (and nodes).

The geothermal modelling team based within the University of Auckland has been able to successfully model geothermal fields with over 500,000 blocks in the model grid, which is possible thanks to Waiwera and cloud computing with access to hundreds of CPUs in parallel.

Another major advantage of using parallel geothermal simulation software like Waiwera in combination with cloud computing is the ability to apply uncertainty quantification to a geothermal project. The uncertainty quantification requires the running of thousands of models, varying parameters like the rock properties and deep hot upflows. The uncertainty quantification of geothermal model predictions is a crucial aspect in a geothermal project for the risk assessment of future development plans.

Recently, a new method has been developed to estimate the geothermal resource potential in a more realistic way including reservoir physics, wellbore physics and realistic production scenarios (Dekkers et al., 2022). This method presents an ensemble of possible resource estimates by running thousands of sample models and using an innovative production algorithm. This approach is also only feasible using Waiwera and cloud computing.

2. SETUP

2.1 Concept

The idea of cloud computing aimed at geothermal reservoir modelling is using HPC resources, that could be based

ISSN 2703-4275

anywhere in the world, for complex geothermal model simulations, while interacting with these resources through the internet. In this paper we will go into more details about the setup and use of Waiwera on NeSI, AWS and Azure (NeSI, 2023; AWS, 2023; Azure, 2023).

NeSI is the HPC platform based at the UoA and is free to use for researchers at the UoA. External users can use NeSI for a fee. AWS is the cloud computing platform, including a HPC resources, from Amazon. Azure is the cloud computing platform, including a HPC resources, from Microsoft.

2.2 Cloud Computing Options

There are different ways to interact with the HPC platforms, in this study we will discuss three options, ordered according to the required level of user interaction, that our geothermal reservoir modelling team, at the University of Auckland, has been using.

2.2.1 Virtual Machine

The lowest level of user interaction is the VM (virtual machine) linked with a local python script. An IT-architect can setup this local python script in a way that the user only has to specify their model and click the run button. Behind the scenes the local python script will start up the VM, gather the required model and geometry files, upload these files, run the simulation, post-process results (if required), download the result files and shut down the VM.

The IT-architect would have created this VM with the right configuration, installed the required software (simulator and processing-scripts) on the VM and setup online functions that link the local python script with the VM on the cloud.

Using a VM has a marginally small, fixed cost for the configuration of the on-demand VM and a variable on-demand hourly rate. The VM option is available on AWS and Azure. At present the fastest VM on AWS is the hpc6a.48xlarge (96 physical CPUs and 384GB RAM) with an hourly on-demand cost of 2.88 USD. On Azure the fastest VM is the HB120rs-v2 (60 physical CPUs and 456GB RAM) with an hourly on-demand cost of 3.60 USD.

2.2.2 Global Scheduler

The higher level of user interaction is the global scheduler, which requires more knowledge from the user compared to the use of a VM. Connecting to a HPC platform is done through the SSH protocol. Implementing this SSH protocol can be done in different ways. We use a combination of WinSCP and PuTTY, which are applications for Windows. This combination gives the user a file explorer and terminal.

In the case of a global scheduler the user needs to know more about Ubuntu (an open-source Linux operating system) and Slurm (an open-source workload manager and job scheduler). Linux commands are used for interacting with the models and data. Slurm is used for creating jobs that will send the model and simulator to the HPC resources. The job then goes into a queue based on priority, which can vary depending on the platform you are using and subscription plan. The wait can be from seconds up to more than 24 hours, which highly depends on how busy the platform is and the size of your job.

The advantage of this approach is that big jobs can be divided over multiple nodes, each containing up to 96 CPUs depending on the platform. Example are the simulation of a large complex model with over 500,000 blocks in the numerical grid or simulating thousands of sample models for uncertainty quantification (UQ). The disadvantage of the approach is that the resources are not on-demand and so the wait for results could be long, however, it will be faster than simulating big jobs locally. The global scheduler can be used on NeSI, AWS and Azure.

In this case the IT-architect would have installed the required software (simulator and processing-scripts) on the cloud.

2.2.3 Cluster

A cluster requires the same level of user interaction and knowledge as the global scheduler. The cloud interaction can be implemented as well through, for example, a WinSCP and PuTTY combination. However, the IT-architect requires more knowledge about how to configure the cluster.

The user also needs to know about Ubuntu commands and Slurm in this case. In contrast to the global scheduler the HPC resources are always ready on-demand and can be seen as a private scheduler. There is one head-node that manages all resources that are configured for the cluster. All software should be installed through the head-node. The jobs should be submitted in a similar way as for the global scheduler using Slurm.

One of the advantages of a cluster is that you get the option to configure various nodes each for their own purpose. For example, you could configure 10 HPC compute nodes for simulations and 10 process nodes for running Python scripts. In this way your jobs will be done in the most efficient and cost-effective way. Another advantage of a cluster is the on-demand functionality. However, this also brings a disadvantage, as after a job has been submitted, the compute node requires 3 minutes of configuration time. This can be solved by keeping the node active, which then adds extra costs if the node is not used all the time.

The cluster option is available on both AWS and Azure, but not on NeSI.

2.3 Regions

The three platforms each have their own regions. While the resources of NeSI are only available in New Zealand, AWS and Azure have resources located all over the world. Being in another region means you can still connect to a resource in a specific region, however, there is a connection latency, which increases with longer distance.

The latency has a moderate effect on user interaction with the cloud resources. The upload and download speeds, however, are significantly affected. Furthermore, resources in specific regions vary in availability and price.

There are solutions to speed up your workflow, like pre- and post-processing on the cloud and transferring files on the back-end network of a platform. Nevertheless, all these things should be considered when choosing the right cloud computing solution.

3. COMPARISON

In this section we compare the three platforms that were discussed in the previous section based on purpose, options, speed, and pricing.

ISSN 2703-4275

3.1 Purpose and options

NeSI is a great cloud computing platform for researchers at the University of Auckland. The services and support are generally free, and connection and file transfer are fast when you are based in New Zealand. Our modelling team uses this cloud computing platform for research on geothermal reservoir modelling.

However, NeSI has limited scalability and is not on-demand which are requirements for modelling commercial geothermal projects on a tight schedule. The fee for using NeSI commercially, is also not competitive with the alternatives.

Therefore, our team has tested both AWS and Azure for geothermal modelling for commercial projects. We have tested both platforms simulating complex models as well as simulating thousands of sample models. Both platforms meet the requirements for our purpose. Fast, on-demand and highly scalable. Both platforms have various options for VM's (low-level user interaction) as well as cluster setups (high-level user interaction).

3.2 Speed and pricing

As a comparison between the three platforms and a desktop computer, we show the simulation time for some geothermal models in Table 1. As a reference on a single compute node NeSI has 40 CPUs, AWS has 96 CPUs and Azure has 60 CPUs. The price per hour for AWS and Azure is 2.88 USD and 3.60 USD, respectively. Each of the cloud computing platforms are a lot faster than a desktop computer. Computational speed-up can be up to 42 times.

The comparison from the table below shows that the order of computational speed from high to low is AWS, Azure and NeSI. In addition, the hourly compute node costs are lower for AWS than for Azure. This is amongst other reasons why AWS is the best option for our modelling team when working on commercial geothermal projects. However, as mentioned before, the user's requirements and purpose are important for choosing the right cloud computing solution for geothermal modelling.

Table 1: Speed comparison table.

Model (nr blocks)	PC 4 CPUs	NeSI 40 CPUs	AWS 96 CPUs	Azure 60 CPUs
Synthetic NS (11,764)	454 sec	78 sec	45 sec	53 sec
Synthetic PR (11,764)	36 sec	15 sec	12 sec	14 sec
Kotamobagu NS (39,994)	120 sec	36 sec	15 sec	25 sec
Kotamobagu PR (39,994)	100 sec	42 sec	30 sec	35 sec
Rotorua NS (94,701)	12 hours	50 min	17 min	30 min
Rotorua PR (94,701)	8 hours	25 min	7 min	10 min

4. APPLICATIONS

In the previous section we briefly mentioned the applications of cloud computing. In this section we will discuss examples of these applications and results of using cloud computing for geothermal modelling.

4.1 Model calibration

Geothermal reservoir models are typically used to support management to make sustainable develop plans. This is done by using the model as a replica of the real-world reservoir for testing potential development plans. It is crucial to develop accurate models in order to have confidence in the model predictions. Developing accurate models is done by a process called model calibration, which means fitting your model to measurements by adjusting model parameters like rock properties and deep hot upflows. An example of a well calibrated model can be seen in Figure 1. Cloud computing was an invaluable tool in speeding up the calibration process for this model.

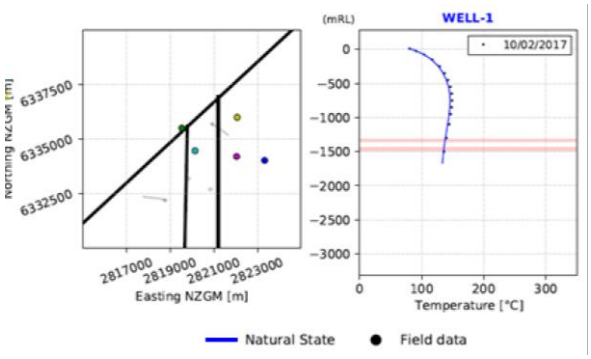


Figure 1: On the left a map of various well locations. On the right the downhole temperature of Well-1.

4.2 Uncertainty Quantification

A well calibrated model can give accurate predictions for the behaviour of a geothermal reservoir. However, a model should always be seen as a useful tool and not as the truth. In addition, there are often few field measurements for model calibration, and these measurements can also be inaccurate.

These approximations and inaccuracies are uncertainties, which we can include in our model, and therefore, model predictions. In contrast, to a single calibrated model with a single model outcome, carrying out an uncertainty quantification gives a range of possible model outcomes. These results are invaluable for managing the reservoir based on worst- and best-case scenarios. Typically, the uncertainty of the results is dependent on the available data for the reservoir.

The uncertainties are translated to statistical distributions for model parameters like rock properties and deep hot upflows, such as those shown in Figure 2 and Figure 3, respectively. The parameters are often centred around the calibrated values and then sampled based on a gaussian distribution.

ISSN 2703-4275

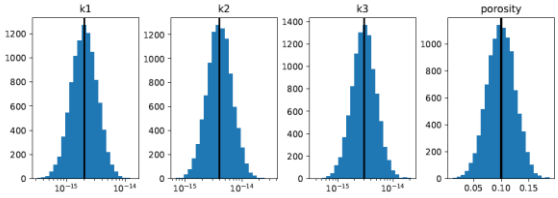


Figure 2: Example of rock permeability and porosity uncertainty.

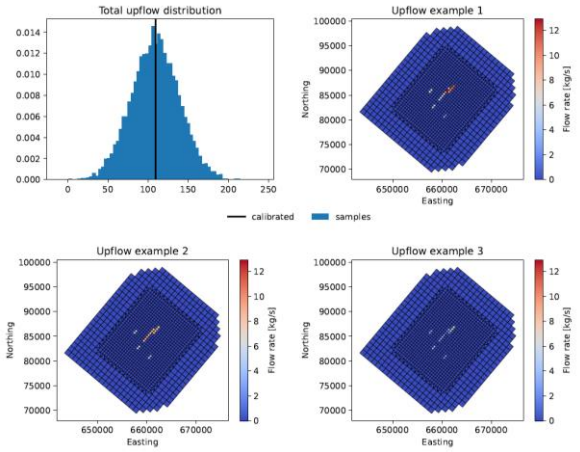


Figure 3: Example of uncertainty in the magnitude and location of the deep hot upflow.

Thousands of sample models are then generated using these sampled parameters. All these models are then simulated on the cloud computer, which can be done in just several hours. The results from these sample models give a range of possible temperatures, pressures, and other model outputs. Figure 4 shows an example of downhole temperature UQ results. The grey lines represent all the sample model temperatures. However, since there is data available for this model a filter can be applied to get a more accurate uncertainty quantification. The black crosses are downhole temperature measurements. We now condition the initial sample models on these measurements by selecting the sample models that fit the data the best, which is the filtering process. The red lines are the leftover filtered sample models, which provides a more accurate range of possible models. All types of field data can be included in the filtering process, for example, production enthalpy or surface outflow data. The method used for this filtering process is the Approximate Bayesian Computation (ABC). The ABC is implemented, amongst other uncertainty quantification tools, in the Python package *ccandu* (conditional composition and uncertainty quantification), which is developed in-house at the University of Auckland.

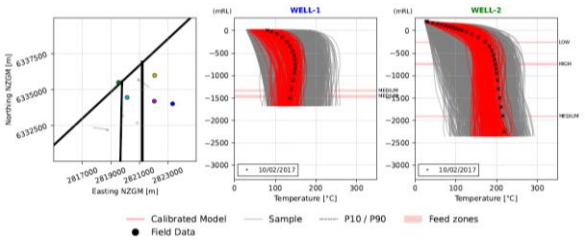


Figure 4: Uncertainty results for temperature profiles down the well. The grey lines indicate the initial sample models. The crosses are the field data that were used to condition the sample models. The red lines are the filtered sample models that are left after conditioning on the field data.

The ensemble of filtered models, conditioned on the available data can now be used to test future development plans of the reservoir. Figure 5 shows the UQ results of a future scenario. The figure shows the production mass flow and power output of the reservoir from 2017 to 2030. On the left the production mass flow is shown with all samples following the historic data up to 2020. From 2020-2030 the UQ shows a range of possible production mass flows and power outputs which is a more robust model prediction compared to a single prediction from a calibrated model. The whole UQ process including simulating thousands of models can now be done in the time frame of a single day with the use of cloud computing. We as a team have made it standard best practice to do an uncertainty quantification for each geothermal project. The UQ results gives the reservoir management worst- and best-case scenarios that help support decisions on their development plans.

We have already applied our UQ methodology to research projects as well as several commercial projects, like Lihir (Newcrest Mining Ltd in Papua New Guinea), Karaha, Bukit Daun and Sungai Penuh (PGE in Indonesia) and East Brawley (Phoenix Lithium in the USA).

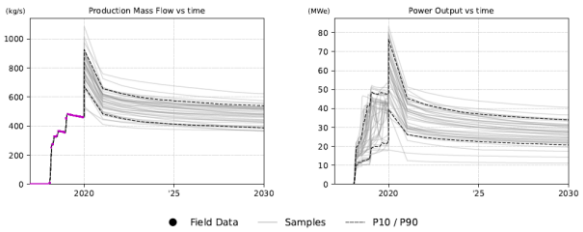


Figure 5: Model predictions with a range of possible models. The production mass flow is shown on the left and the power output is shown on the right.

4.3 Resource Potential Estimation

The gold standard to estimate the resource potential used to be a stored heat calculation. This method estimates the stored heat of the reservoir and applies a fairly arbitrary recovery factor to obtain the power potential of a green field geothermal project.

The availability of cloud computing resources has made it possible to estimate the resource potential of a green field project using full scale numerical modelling. Our modelling team has developed this new method which, in contrast to the stored heat calculation method, includes reservoir physics and wellbore modelling (Dekkers, K., et al. 2022).

For this method we also generate thousands of possible reservoir models based on statistical distributions of model parameter, as explained in the previous section. For green field projects there is typically not much data available. Often there is, however, magnetotelluric (MT) data that indicates the location of the clay cap. The clay cap is a zone of altered formations which happens at a temperature of 200 °C. The formation directly under the clay cap, which has not been altered yet, can therefore be estimated to be at around 190 °C. This information is used to condition the initial samples. Figure 6 shows the temperature of four blocks right under the clay cap with (in blue) the initial temperature distribution and (in red) the filtered temperature distribution, which is now

ISSN 2703-4275

centred around 190 °C. Other possible exploration data can include surface features and early stage well measurements. The more data that is available the more accurate is the resource potential estimation.

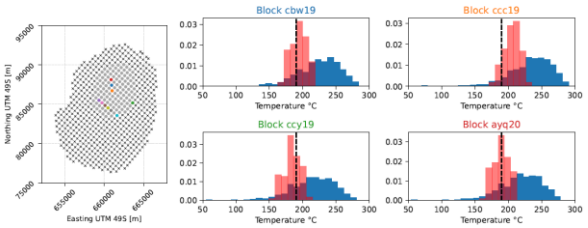


Figure 6: On the left the map with the location of the clay cap is shown. On the right the temperature distributions of four blocks under the clay cap are shown. In blue the initial distribution is shown and in red the filtered distribution.

After conditioning on the available data, an innovative production extraction algorithm is applied to the filtered sample models. This algorithm selects target production and injection wells based on wellbore modelling, and the formations and structures of each sample model. Simulation is then used to estimate the full power potential of each sample model, which results in a range of potential resource estimations. Figure 7 shows an example of the estimation of the power output, production and injection mass flow, and production and injection enthalpy of the green field Kotamobagu system (Nagoro, B., 2023).

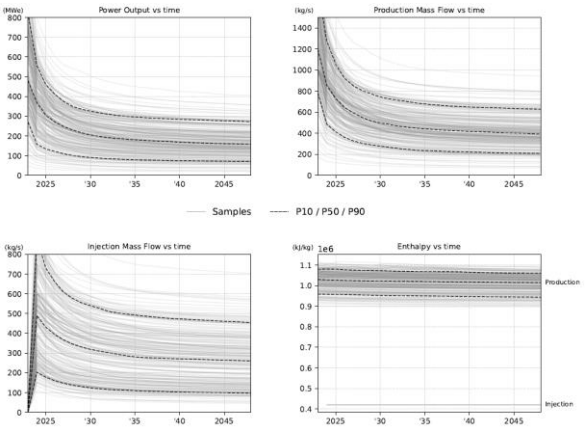


Figure 7: Resource potential estimation of hundreds of possible reservoir models. The grey lines are the sample models, and the dashed black lines are the P10 / P50 / P90 uncertainty percentiles.

5. CONCLUSIONS

Reservoir modellers now have different options for cloud computing each with their own advantages. Cloud computing in combination with Waiwera geothermal modelling has made it possible to innovate.

Geothermal models can be made more complex and refined, while simulation time has decreased. Complex models integrated with uncertainty quantification now give more accurate and robust model predictions.

Additionally, it is now possible to estimate the resource potential of green field geothermal projects using numerical modelling. The method includes reservoir physics and wellbore modelling and therefore should be preferred over a stored heat calculation.

Our team has applied uncertainty quantification to various projects and is now our standard best practice.

Future work will include researching different conceptual models for the UQ and the use of machine-learning for model calibration.

ACKNOWLEDGEMENTS

The author(s) wish to acknowledge the use of New Zealand eScience Infrastructure (NeSI) high performance computing facilities, consulting support and/or training services as part of this research. New Zealand's national facilities are provided by NeSI and funded jointly by NeSI's collaborator institutions and through the Ministry of Business, Innovation & Employment's Research Infrastructure programme. URL https://www.nesi.org.nz.

The author(s) would like to thank Amazon Web Services (AWS) for their ongoing support and collaboration.

REFERENCES

Amazon Web Services, Inc. (2023), https://aws.amazon.com/.

Croucher, A., O'Sullivan, M., O'Sullivan, J., Yeh, A., Burnell, J., & Kissling, W. (2020). Waiwera: A parallel open-source geothermal flow simulator. *Computers & Geosciences*, *141*, 104529.

Dekkers, K., Gravatt, M., Maclaren, O., Nicholson, R., Nugraha, R., O'Sullivan, M., Riffault, J., O'Sullivan, J. (2022). Resource Assessment: Estimating the Potential of a Geothermal Reservoir. *PROCEEDINGS, 47th Workshop on Geothermal Reservoir Engineering Stanford University, Stanford, California, February 7-9, 2022.*

Microsoft Azure (2023), https://azure.microsoft.com/en-us/.

Nagaro, B. (2023). Quantifying Geothermal Resource Potential and Uncertainty Analysis using a Natural State Model of Kotamobagu Geothermal Field in North Sulawesi, Indonesia. *Master Thesis.*

New Zealand eScience Infrastructure (2023). https://www.nesi.org.nz/.