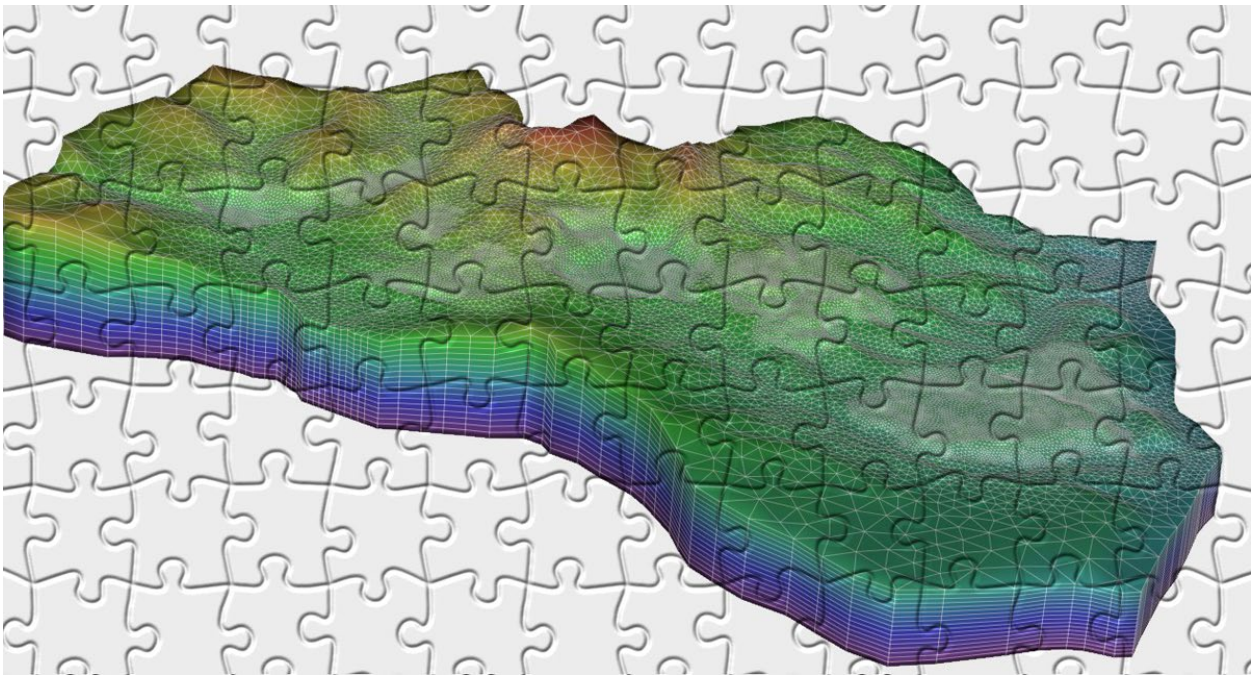


Too Much of a Good Thing?

A test case of FEFLOW scalability in an Azure environment

The Introduction

Parallel computing of groundwater models, i.e., the simultaneous utilization of processors when solving numerical problems, can reduce run times and subsequently reduce cost. Yet not all groundwater modeling simulators that can accommodate parallel computing are designed to do so efficiently. Only few simulators were designed with high-performance computing in mind, where hundreds or thousands of processors can be used effectively at one time (so-called “embarrassingly” or “massively” parallel simulators).

At some point, scalability – i.e., reduction of run time as a function of adding more processors – becomes less effective due to slow-down caused by passing of distributed information from the various processors and “stitching” all this information back together. The ability of a numerical simulator to scale up affects the choice of hardware or cloud computing configuration.

The Test

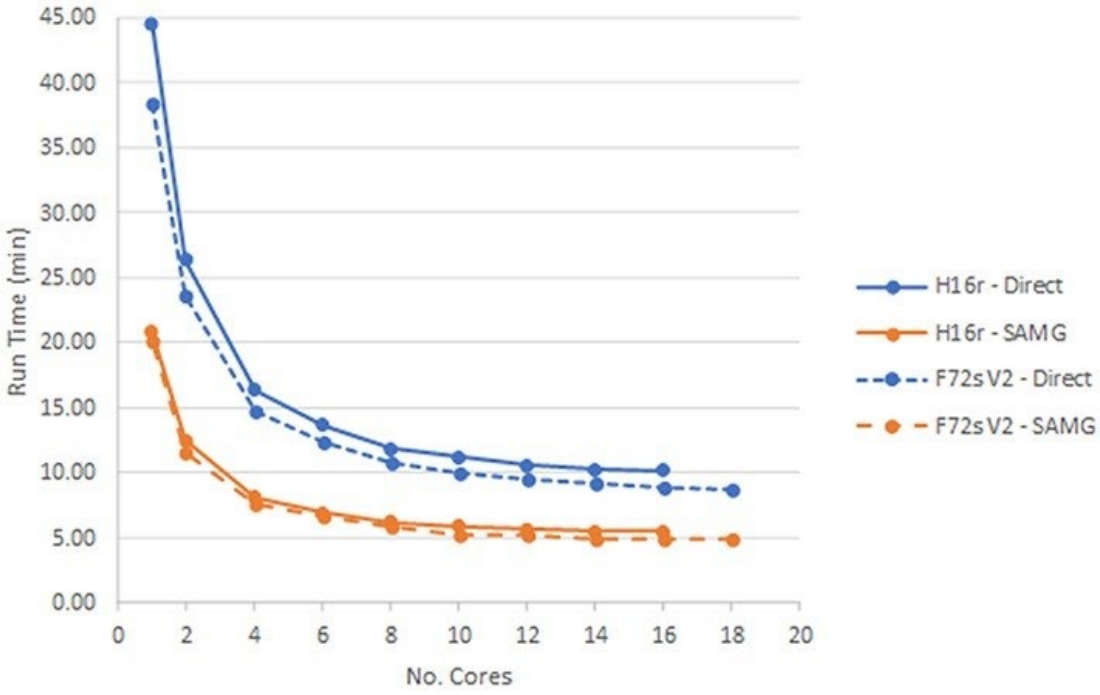
For recent groundwater modeling work, SRK modelers tested the scalability of a transient FEFLOW model on several Microsoft Azure configurations, utilizing different FEFLOW solvers. The test was automated by running FEFLOW from a batch script, which also recorded the run times. The solvers tested were the Direct and the algebraic multigrid methods for systems (SAMG). The Azure architectures tested were of “Hr” series (hard-disc drives, Intel Xeon E5-2667 V3 processors) and the “Fs V2” series

(solid-state drives, Intel Xeon 8168 processors). Note that Microsoft updates and retires architectures periodically, so these specific configurations may not be available anymore.

The Results

The figure below plots run times as a function of number of cores used in the simulation. Several observations stand out from this test:

- There was little difference in performance – approximately 10% - between the Fs V2 and Hr series. The Fs V2 series performed slightly better than the Hr series, probably owing to newer processors and use of SSD instead of HDD.
- Overall, run times did not improve much when 10 or more cores were used. At 10 cores, the model took a quarter of the time to run, compared to when only a single core was utilized.
- The SAMG solver performed better than the Direct solver for this specific model.



The Conclusion

For this particular FEFLOW model, in these particular tested environments, the use of more than 10 cores would be simply inefficient. Other tests performed by SRK have shown that the model actually ran slower with 20 cores than with 10 cores, and even slower with 40.

Can there be too much of a good thing? In parallel computing of numerical groundwater models, apparently yes.

Assaf Wunsch | Consultant / Hydrogeology | SRK Consulting (U.S.), Inc.
awunsch@srk.com