Benchmark as decision support for cloudification: Moving CAE and HPC to the cloud increases quality and efficiency of simulations

Christopher Woll, GNS Systems GmbH

Abstract

A decision in favor of the cloud and the associated organizational changes are of central strategic importance for automotive OEMs and their suppliers. In advance, it is important to clarify whether existing organizational structures and processes can actually be optimized with the help of the cloud. To this end, we regularly carry out benchmarks for well-known customers as a decision-making aid. In the presentation, we will present a benchmark with real productive jobs based on existing data in the customer's specific environment. Mapping the workflow with selected CAE applications in the cloud provided a realistic basis for evaluation of their benefits and costs. In addition to comparing runtimes and results of selected benchmark jobs, our experts tested the quality of remote work in a workstation scenario for selected finite element analyses (FEA), Computational Fluid Dynamics analyses (CFD) and iterative analyses.

The benchmark on the High Performance Computing (HPC) reference architecture was based entirely on a cloud environment from GNS Systems. Within the available subscription, separate virtual networks were created for the scalable HPC cluster, the license server and the rest of the infrastructure. The core was formed by an orchestration system and a master node. Necessary machines for the simulation jobs were started flexibly via predefined images and so-called "scale sets". Demand was determined by the number and size of jobs in the queue. Unneeded nodes were shut down to optimize costs.

The benchmark results demonstrate the high efficiency of simulations in the cloud. Engineers also achieve higher quality simulation results with better computing power. The simulation jobs benefit from the increased number of CPU cores per node and the demand-oriented scaling of computing performance. The implementation was functional and error-free with a selected toolset. The engineers experienced the interactive work sessions on remote desktops with Linux and Windows as smooth and trouble-free. The benchmark clearly proves that performance on modern CPUs in the cloud does not lag behind conventional on-premise solutions. The results for this specific business case provided a reliable strategic decision-making basis for the cloudification of HPC and CAE processes.

1 Introduction

The share of cloud computing in industry and manufacturing has increased rapidly in the past few years. According to the current Cloud Monitor 2021 by KPMG and Bitkom Research [1], four out of five of the companies surveyed already rely on the cloud. Cloud computing makes the greatest contribution to the digitalization of internal processes and the automation of workflows. This is also the case in the automotive industry: the cloud is an important enabler for its use cases and is driving innovation. It is not only changing the way in which IT is used, but also the organization of processes around the technologies used. Global corporations are currently establishing the cloud as the central instance of their processes. One example is Volkswagen, where the data of all machines, facilities and systems from all of the group's factories are being brought together in the cloud, thus fundamentally optimizing organizational structures [2]. In the long term, the concept foresees including the Group's global supply chains and is pointing the way forward for the entire automotive industry.

Companies in the automotive industry can currently access a wide range of computing technologies to carry out their simulation tasks. In addition to classic on-premise technologies, the latest high performance (HPC) technologies in the cloud have increasingly come into focus in recent years. Although modern cloud HPC systems offer the same performance and technologies as on-premise clusters, many companies are interested in minimizing job execution times while reducing simulation costs when using on-demand resources. On the one hand, cloud HPC solutions meet the demand for massive computing power quickly and reliably through their architecture. But at the same time, they offer positive effects such as pay-per-use, lower capital expenditures (CAPEX), greater business flexibility, and higher-quality results due to increased standardization. Whether these come into play for a specific business case, however, is crucial for the risk assessment of cloudification processes. Accordingly, cloud performance should always be examined and evaluated in detail for the use case in advance. We outline below how this can be sensibly done in practice.

2 Project description

The benchmark performed by GNS Systems was based on real productive jobs based on existing data in the specific client environment. The department involved on the customer side in the area of CAE infrastructure and High Performance Computing previously preferred to use on-premise clusters for analyses. Mapping the workflow with selected CAE applications in the cloud was to provide a realistic basis for evaluation of their benefits and costs.

With the aim of being independent of data centres and their limited resources in virtual development in the future, the benchmark examined the options of selected services in the cloud. In addition to comparing runtimes and results of specific benchmark jobs, GNS Systems tested the behaviour of CAE simulations for specific finite element analyses (FEA), computational fluid dynamics analyses (CFD) and iterative analyses in the cloud. In the best case scenario, cloud usage delivered better results than simulation analyses in a comparable on-premise environment. The practical test also confirmed that the cloud is capable of reliably and consistently mapping complete virtual product development workflows and not just processing HPC workloads. Accordingly, the cloud offers the potential to increase efficiency in prototyping, for example, and to drive innovation faster.

3 Structure and architecture of the benchmark

At its core, the benchmark aimed to provide comprehensive information on cloud performance for a defined business case. Accordingly, the aim was to compare the runtimes and results of selected benchmark jobs from finite element analysis (FEA), computational fluid dynamics analysis (CFD) and iterative analysis in the cloud. The basis was formed by real data from productive jobs in the customer-specific environment to be able to draw a realistic comparison of CAE simulations to the respective analyses in the cloud.
.

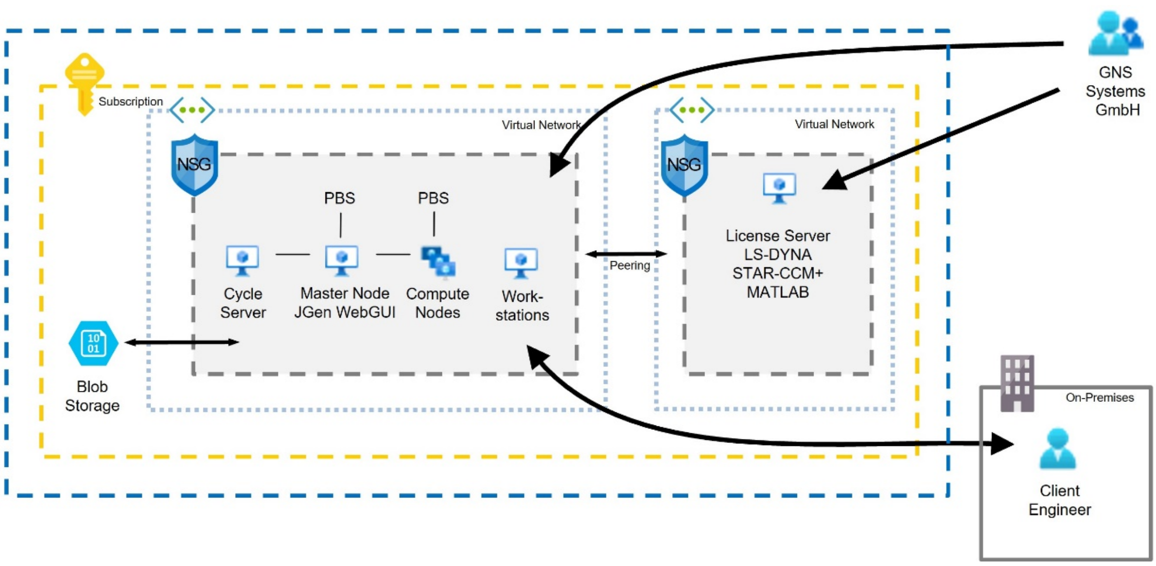

Fig.1: Overview of the Cloud Architecture

The benchmark on the High Performance Computing reference architecture was based entirely on a cloud environment in Microsoft Azure designed by GNS Systems. As part of the available subscription, separate virtual networks were created and deployed for the scalable HPC cluster, the license server and the rest of the infrastructure. The core was formed by an orchestration system and a master node. These are composed of a server running CycleCloud to orchestrate the nodes, a PBS master server to host job submission and a PBS compute account, and workstations for interactive access as well as job submission. The virtual machines required for the simulation jobs were started flexibly via predefined images and so-called "scale sets". The demand was determined by the number and size of jobs in the PBS queue. Unneeded nodes were shut down to optimize costs. In addition to Linux workstations with CentOS 7.8, the benchmark used a Windows workstation with Windows 10 Professional for the practical test in iterative analysis.

During the preparation phase, a license server in the cloud with fixed host identification parameters ensured the continuous availability of the license keys for the simulations. For this purpose, GNS Systems connected the initially manually set up license via a network to the automated HPC environment in the cloud. The transfer of customer benchmark test data and software packages to GNS Systems was ensured by a web portal in the cloud with direct access to a read-only, non-searchable blob memory.

Job submission was also done via GNS Systems' own job orchestration tool "JGen" on a test basis. In classic on-premise HPC, a common metric is "jobs per day." This hard limit becomes obsolete with the cloud: there is no fixed amount of computing resources, so theoretically an unlimited number of jobs per day can be run in the cloud. However, to stay within a budget, it is important to orchestrate the simulation jobs according to the available resources. This is where the "JGen" tool, which is in production at several customer sites, comes in: It supports engineers from simple job submission and job control to the creation of user-defined job workflows. A Linux application core with solver-specific modules can orchestrate the PBS jobs used in this benchmark. This is done in a completely user-friendly way via an integrated WebGUI, which can be accessed at any time from the workstation in the cloud. JGen thus forms a reliable alternative to the direct submission of benchmark jobs via the command line of the PBS master server.

In the following, we present in more detail the individual benchmarks of the selected finite element analysis (FEA), computational fluid dynamics analysis (CFD), and iterative analysis in the cloud and their results.

4 Details of the Benchmarks

In the area of finite element analysis (FEA), GNS Systems looked at the simulation program LS-DYNA. The software enables the investigation of highly nonlinear physical processes on the computer. In the automotive industry, LS-DYNA is used in crash test simulation and in specific problem solving with dummies (test dummies) and components such as airbags and seat belts. This benchmark used the LS-DYNA 9.3 (double precision) application together with Platform MPI 9.1.4.3, running on the Linux CentOS 7.8 operating system.

The compute nodes in the cloud were from the Azure HC44rs series with Intel processor and HB120rs_v2 with AMD processor. Both target HPC-like workloads and thus form a good equivalent for the on-premise reference architecture to be compared. For the use of the simulation analyses, 44 cores were available in the virtual machine (VM) for the HC44rs with 2x24 Intel cores and 120 cores per node for the HB120rs_v2 with 2x64 AMD cores (Remark: Some cores are reserved for the handling of virtualization overhead in Azure). Since the LS-DYNA test license is limited to 250 cores each, the maximum number of nodes that could be used was 2 nodes with AMD CPU and 5 nodes with Intel CPU. For the comparison of the different benchmark runs, GNS Systems only used the runtime reported by LS-DYNA and not the CPU time. Furthermore, the time for system provisioning and loading the input data were not considered, since the reference did not use these either.

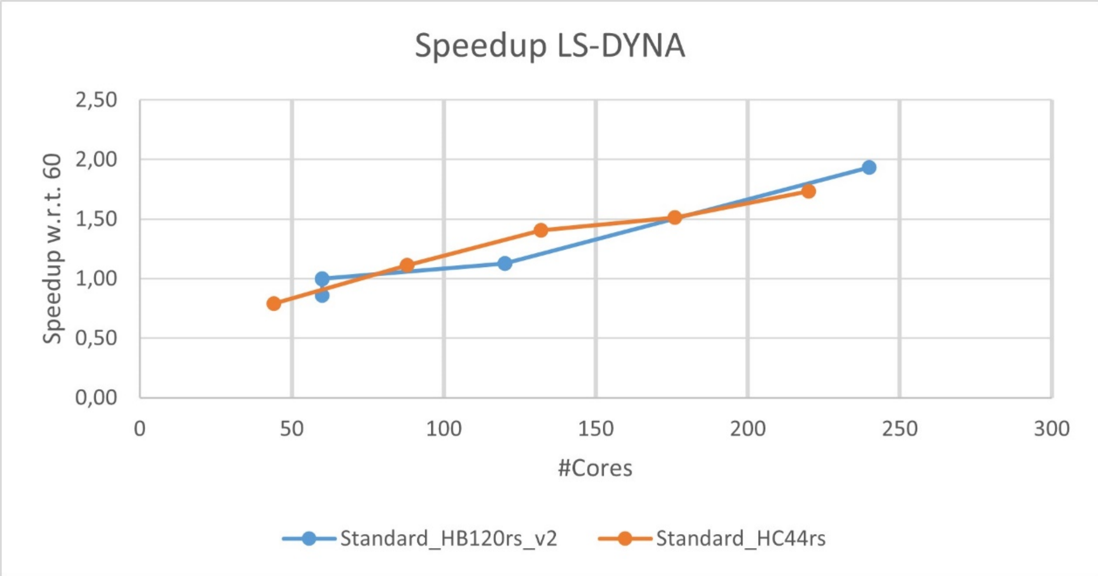

Fig.2: For the comparison of the different benchmark runs, the runtime reported by LS-DYNA without system provisioning and job data loading time was used. Per core, the Intel processor has a better performance than the AMD processor.

In a computational fluid dynamics (CFD) analysis, GNS Systems investigated the Star-CCM+ simulation software. The tool is used to model and analyse a range of engineering problems in flow, heat transfer, stress, particle flow and electromagnetism. It provides a complete multiphysics solution for simulating products and designs under real-world conditions. The benchmark was performed using Star-CCM+ version 15.04.008-R8 and Platform MPI under the Linux operating system CentOS 7.8. As in the LS-DYNA benchmark, the compute nodes in the cloud were from the Azure HC44rs series with Intel processor and the HB120rs_v2 series with AMD processor. However, in the initial simulation analyses, scaling up the AMD nodes was shown to be counterproductive compared to the Intel nodes. To nonetheless ensure comparability, the maximum number of nodes used was limited to three in the HB120rs_v2 series.

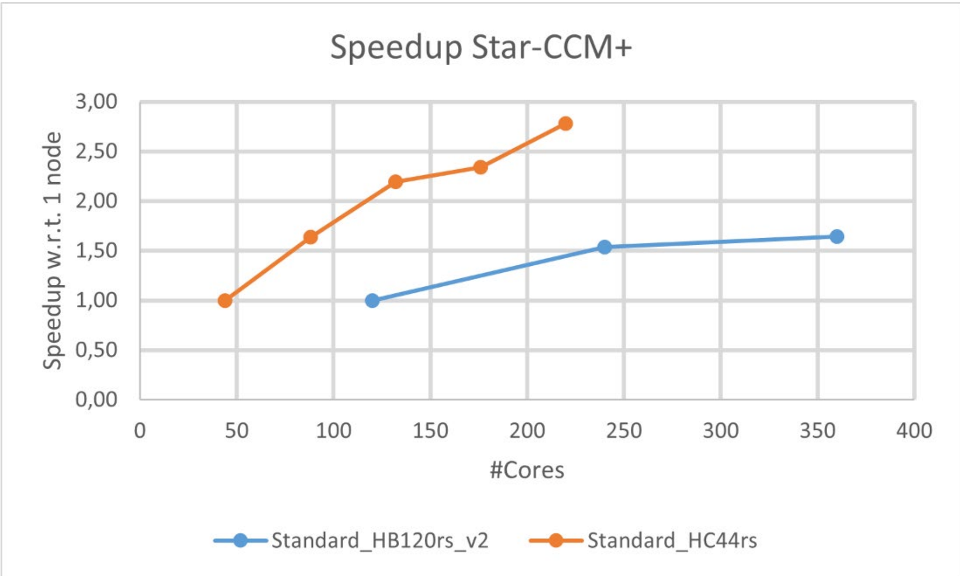

Fig.3: In terms of performance per core, the Intel CPUs are faster. Generally, the SpeedUp in the cloud doubles with the number of CPUs. The customer-side job model used here showed its limitations, in that it scaled negatively for this benchmark.

For iterative analysis in the cloud, GNS Systems investigated the MATLAB software. MATLAB is proven in solving mathematical problems and integrates numerical simulation as well as data acquisition, analysis and evaluation. Unlike the previous benchmarks, GNS Systems conducted this interactive real-world test on a Windows 10 workstation running MATLAB version R2017b. Due to the special requirements, the benchmark used an Azure NV6 instance with an NVIDIA Tesla M60 GPU. Access to the Windows VM was provided by the RDP protocol.

5 Comparison of benchmark results

The three benchmark runs show the same or at least similar performance compared to on-premise use. In particular, the LS-DYNA benchmark showed a higher efficiency for simulation analyses in the cloud. Per core, the Intel processor delivers better performance than the AMD processor. The runtimes of the 120-core AMD job and the 2-node Intel job with only 88 cores were equally fast compared to the customer reference architecture. Especially in job submission, the "JGen" tool achieved more efficient utilization of available resources. The implementation was functional and error-free at all times with JGen. The advantage over local computing resources: jobs can be scaled better and executed faster in the cloud than with rigid on-premise systems. To achieve the same results on-premise would usually require oversized systems.

The benchmark with Star-CCM+ showed no significant differences between on-premise and the cloud. In principle, however, these simulation jobs also benefited from the increased number of CPU cores per node in the cloud and the scaling of computing power according to demand. The practical test with MATLAB confirms the assumption made earlier: It showed good performance and a recognizably better benchmark performance than the on-premise solution.

In principle, therefore, it can be said that engineers can achieve higher-quality simulation results with better computing performance in the cloud. The benchmarks clearly prove that the performance on modern CPUs in the cloud does not lag behind conventional on-premises solutions.

From a cost perspective, the results from the LS-DYNA and Star-CCM+ benchmarks present themselves as follows: The AMD-based compute nodes offer a better cost-benefit ratio compared to the Intel hardware. This can mainly be attributed to the higher number of CPU cores per node. However, for a complete cost analysis, more data - for example, that of the license costs - needs to be considered to create a valid basis for decision-making.

6 Conclusion

The results from the benchmarks show the various options for selected services from the cloud. It becomes clear that performance in the cloud is in no way inferior to established on-premise architectures. Against the background of establishing virtual product development independent of data centers and their limited resources in the future, these results are an important basis for decision-making in change processes. As before, many of the CAE simulations to be executed require powerful computing resources that are maximally scalable and can be flexibly accessed at any time. The increasing complexity of these calculations is likely to push this need even further in the coming years. For companies, the choice of the right HPC environment - on-premise or in the cloud - therefore depends in many cases on the underlying tasks and processes. The following applies: The more irregular the load distribution of the resources, the more worthwhile it is to move CAE simulations to the cloud.

For companies in the automotive industry, correspondingly hybrid models are suitable, in which existing on-premise HPC clusters are flexibly extended by cloud services as required. Particularly in the case of computationally intensive analysis techniques, the cloud then provides the HPC capacity for simulation analyses that the on-site resources cannot provide. Companies benefit during peak loads from this maximum scalability as well as the highest time and cost efficiency in on-demand HPC environments in the cloud - without the high ongoing costs for technology, operation and personnel. For example, efficiency can be significantly increased in the design of prototypes, and innovations in the automotive industry benefit from a faster time-to-market.

7 Literature

[1] KPMG: Cloud-Monitor 2021 – Die goldenen Zwanziger für die Cloud, 2021, Seite 5
[2] Volkswagen: Volkswagen bringt weitere Partnerunternehmen in die Industrial Cloud, Press release, 23.07.2020