Lower the Time-to-Results for Tightly Coupled HPC Applications on the AWS Cloud with the Elastic Fabric Adapter

INTRODUCTION

Every day, High Performance Computing (HPC) customers use AWS to run their most compute-intensive workloads in computer-aided engineering (CAE), computational chemistry, genomics, seismic imaging, and many more. You use AWS to power your infrastructure; to lower your cost; and to conduct your IT operations in a flexible, efficient, and secure way. Our integrated portfolio of services allows you to quickly, easily, and automatically build (and take down) and manage your HPC clusters on AWS and helps you to accelerate your time-to-results for your HPC workloads.

This White Paper examines AWS' next-generation, cloud-centric network technology, the Elastic Fabric Adapter (EFA). EFA was made available to (HPC) customers in the first half of 2019 and targets an important and challenging class of HPC applications: tightly coupled workloads. Tightly coupled applications are used in every industry vertical with examples in computational fluid dynamics (CFD), weather prediction, reservoir simulations, and even in new technologies such as in Machine Learning/Deep Learning (ML/DL).

Hyperion Research documents the Return-on-Investment (ROI) of HPC

Hyperion Research has found that businesses see $507 in incremental revenues and $47 in incremental profit for every dollar invested in HPC. At the same time, lost productivity is common among organizations that rely on on-premises HPC systems: [1]

72.8% of organizations can't keep up with user needs, resulting in delayed or canceled jobs

29% of organizations indicated that pent-up demand exceeded 50% of total yearly workload volume.

USE CASES

Computational Fluid Dynamics - With EFA, design engineers can easily scale out their simulations of complex flows leading to faster results.

Weather Modelling - EFA offers a fast interconnect that allows weather modelling applications to take advantage of the scaling capabilities of the cloud and get more accurate predictions in less time.

Machine Learning - Optimized for the NVIDIA Collective Communications Library (NCCL), EFA improves the throughput and scalability of Caffe, Caffe2, Chainer, MxNet, TensorFlow, and PyTorch training models on GPUs, which leads to faster results.

TIGHTLY COUPLED HPC WORKLOADS ARE HARD

Tightly coupled HPC applications solve the underlying mathematical model of a dynamic physical system - be it igniting a combustible mixture in an engine, extracting oil from a reservoir, or "the weather" — in a highly iterative and closely coupled fashion. These applications generally require locally computed data to be distributed globally (over the HPC cluster interconnect) for frequent recalculation until some convergence criterion is met. They are particularly challenging — for cloud providers and on-premises system OEMs alike — in that even small inefficiencies or disruptions in the network service can lead to disproportionately wasted compute resources, higher turnaround times, and increased cost or productivity loss.

An evaluation study of HPC infrastructure for running the Navy Global Atmospheric Model (NAVGEM) weather prediction code highlights the bottom-line impact of EFA on such tightly coupled applications. The US Naval Research Laboratory (NRL) concluded that for high resolution forecasts, using Amazon EC2 C5n Instances with the EFA network interface achieved results 74% faster with 27% lower costs than previous generation instances (C4) and network (TCP over ENA) on AWS.   On the other hand, note too, that at higher process count, NAVGEM on ethernet (without EFA), InfiniBand, and Omni-Path stops scaling and wall clock time (i.e., time-to-results) exhibits undesired and counterproductive scaling behavior.

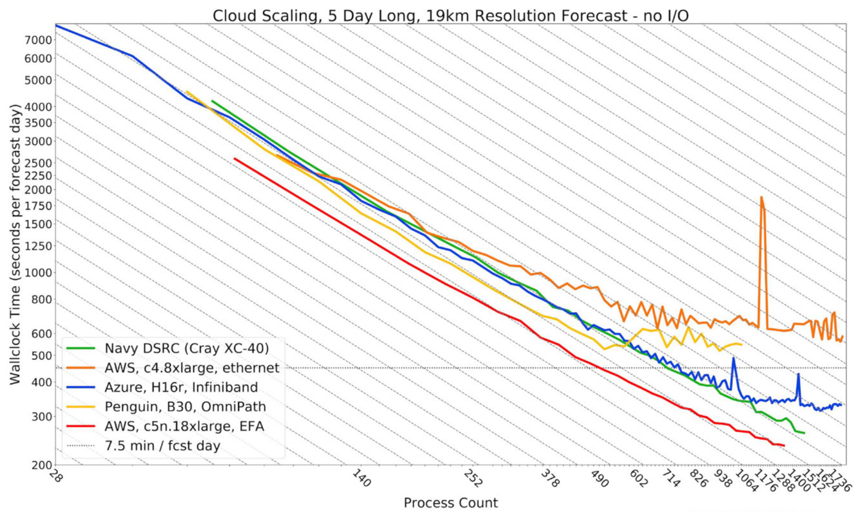

Figure 1: Evaluation of HPC infrastructure for running the Navy Global Atmospheric Model (NAVGEM).

In this paper, we first "zoom in" on the technical details of EFA and its three pillars and corresponding benefits. EFA leverages AWS' long-term investments in and build-up of its network infrastructure and we'll briefly review those for proper perspective. Then, we follow with measurements of EFA's impact on an industry-standard but synthetic benchmark. However, we also demonstrate real, application-level impact on a weather code and a CFD code. Equally important, EFA does not compromise on, and is fully integrated with, what you already like about the AWS HPC services: the elasticity to rapidly, easily, and automatically create or take down your HPC cluster of any size and the flexibility in choice of services and how to deploy them. Finally, we "zoom out" to reiterate the importance of the time-to-results metric for your HPC workloads and how the virtually unlimited capacity of the AWS cloud together with EFA brings a new perspective to this metric. After the conclusions, we show you how straightforward it is to get started with EFA.

THE CHALLENGE: EFFICIENTLY AND RELIABLY SCALING TIGHTLY COUPLED WORKLOADS

Up until recently, the main technology available at AWS to enable tightly coupled HPC workloads was a custom network interface (and accompanying drivers) called the Elastic Network Adapter (ENA). Even today, ENA reliably delivers high-throughput, reduced latency, and reduced variability network performance between EC2 instances. Indeed, today ENA is the network interface for virtually all Amazon EC2 instance types — from A1 to Z1d — and ENA provides up to 100 Gbps networking, depending upon the instance configuration. However, specifically for tightly coupled workloads, HPC customers observed and told us that TCP over ENA is good but not great, especially not at larger cluster sizes.

In response, in the first half of 2019, AWS introduced the Elastic Fabric Adapter (EFA). We describe EFA and its three pillars and corresponding benefits: implementing Operating System (OS) bypass to reduce overhead; continuing ENA improvements for enhanced bandwidth; and innovating in network transport -- the cloud-centric Scalable Reliable Datagram (SRD) -- for reliability.

AWS's commitment to network performance

AWS has always pushed network boundaries in the cloud: AWS provided 10 Gbps instance types in 2011 and a step function in network performance with Enhanced Networking in 2013. AWS was an early adopter of 25 Gbps networking in 2016 and released EC2 C5n — with 100 Gbps networking both between instances and to other AWS infrastructure such as Amazon Simple Storage Service (S3) - in 2018. For cloud providers, network interface cards (NICs) are virtualized and, at AWS and as far back as 2016, the AWS Nitro System is the preferred IO/network accelerator plus security chip plus hypervisor all-in-one "card." Since the 2016 debut of the Elastic Network Adapter (ENA) as part of our Nitro System, customers have been able to easily and transparently select instance types to choose a bandwidth and price point that works best for their application with minimal hypervisor overhead. EFA builds upon ENA to bring that same flexibility and upgradability to HPC and ML workloads.

CHARTING A MORE EFFICIENT PATH FORWARD: INTRODUCTION TO EFA

The keys to EFA's performance improvements are primarily related to three critical technologies. The first is the ability for the application to bypass the kernel and talk directly to hardware (OS bypass functionality), which also lowers variability in application performance. The second is tied to AWS' continued development and adaptation of ENA and device driver software for new high-bandwidth instance types and almost goes without saying. The third is a new, cloud-centric reliability layer known as the Scalable Reliable Datagram (SRD).

OS bypass

OS bypass is a well-developed way to bring small and efficient but specialized network software stacks (SW) to HPC applications. These SW stacks live in user space middleware libraries, such as the Message Passing Interface (MPI), but can communicate directly with network interface cards (NICs) and device drivers. Without OS bypass, MPI generally uses the heavy-weight operating system's TCP/IP stack and the NIC device driver to enable network communication between instances. With EFA's OS bypass design, MPI talks instead with the vendor neutral Libfabric API which bypasses the OS kernel and communicates directly with the NIC to put packets on the network.

The Scalable Reliable Datagram (SRD) protocol

SRD is a reliable, high-performance, lower-latency network transport designed specifically for AWS. It's a significant investment in how data moves across our data center network fabrics and is implemented as part of our 3rd generation Nitro chip.

SRD is inspired by the InfiniBand Reliable Datagram but with changes and improvements needed to operate workloads at cloud scale. SRD takes advantage of cloud resources and characteristics, e.g., AWS's complex multi-path backbone, to support new transport strategies. Key features that contribute to its value for tightly coupled workloads include:

- ***Out-of-order delivery*** - Relaxing the need for in-order message delivery eliminates head-of-line blocking. The messages become independent and in most cases the application/middleware can restore ordering if/when it's needed. To support applications that require such strong ordering guarantees, AWS implemented a packet reordering engine in the EFA user space software stack.
- ***Equal-cost multi-path routing (ECMP)*** - There are hundreds of possible paths between two EFA instances. We use the consistent flow hashing properties of our large multi-path network and SRD's ability to rapidly react to network conditions to find the most effective paths for a message. Packet spraying prevents hot spots and allows for fast and transparent recovery from network failures.
- ***Fast packet drop response*** - SRD responds much more quickly to packet drops than any high-level protocol. Occasional packet loss, especially for your long-running HPC application, is part and parcel of normal network operations, not an aberrant condition.
- ***Scalable transport offload*** - With SRD, and unlike other reliable protocols like InfiniBand Reliably-Connected (IBRC), applications do not maintain a connection with dedicated resources for each pair of communicating processes in the application. A single process can create and use a single Queue Pair (QP) to communicate with any number of peers.

InfiniBand	SRD
Messages	Messages
In-order	Out-of-order
Single (ish) path	ECMP spraying with load balancing
Static user-configured timeout (log scale)	Dynamically estimated timeout (usec resolution)
Semi-static rate limiting (limited set of supported rates)	Dynamic rate limiting
Transport offload with scaling limitations	Scalable transport offload (same number of Queue Pairs (QPs) regardless of cluster size)

Table 1: How SRD compares to InfiniBand, at a glance

MEASURING THE IMPACT OF EFA ON TIGHTLY COUPLED WORKLOADS

From benchmarks to customer experiences, EFA significantly increases network performance on EC2 instances with direct benefits in decreased time-to-solution or lower costs.

As a warm-up, Figure 2 compares multi-stream bandwidth of MPI across multiple generations of EC2 instances using the OSU Benchmark Suite.

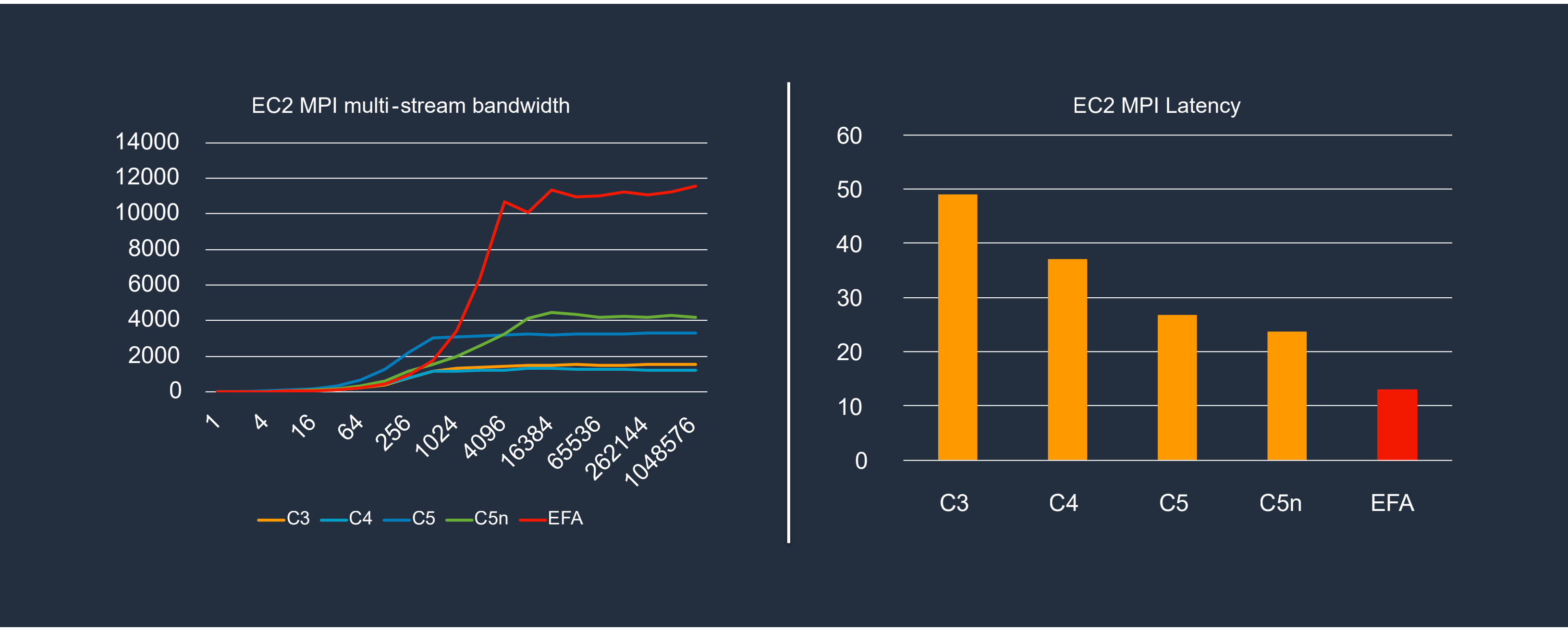

Figure 2: MPI benchmarks. C5n results in orange in the right pane are without EFA.

As seen in Figure 2, EFA significantly enhances bandwidth utilization (close to the line rate, 100 Gbps, in fact) while also meaningfully reducing single-packet latency.

More importantly, when it comes to applications scaling, EFA delivers as well; as already pointed out for the NAVGEM weather code (Figure 1). In addition, we also show a CFD performance measurement exercise — the ANSYS® Fluent® benchmark that models the "External Flow Over a Formula-1 Race Car" . In Figure 3 we show the official performance "rating" of a Cray XC50 versus AWS HPC cluster (EC2 C5n.18xlarge instances with EFA). ANSYS defines this rating as "the number of benchmarks that can be run on a given machine (in sequence) in a 24-hour period." Higher is better.

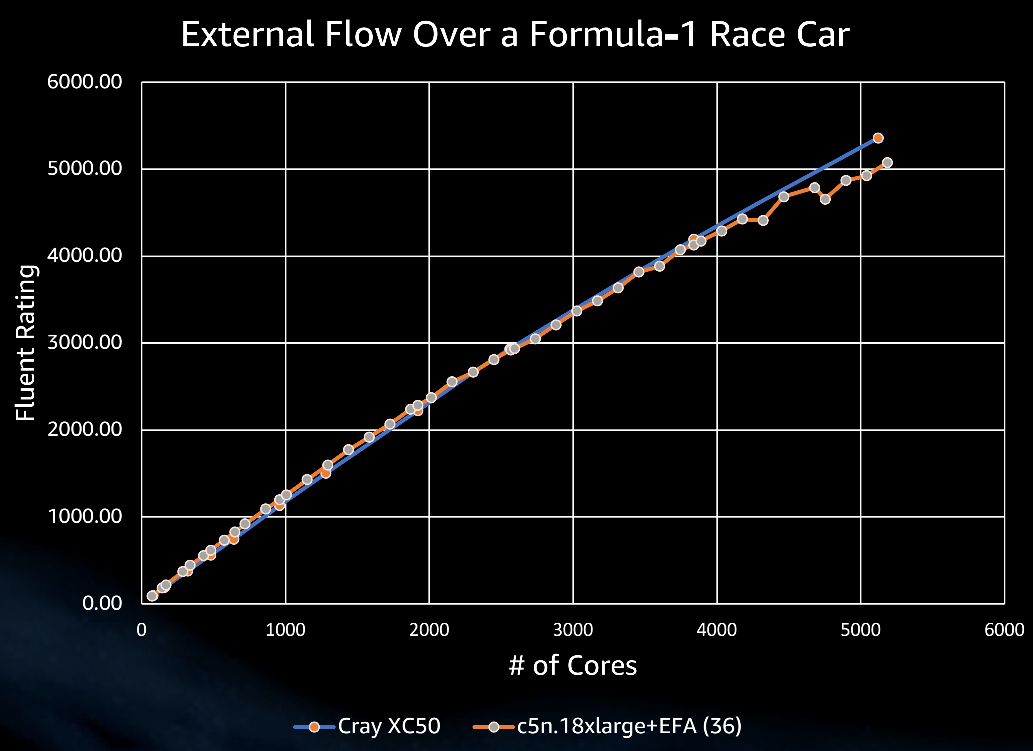

Figure 3: This chart shows C5n.18xlarge with EFA got a higher rating up to 2,400 cores and is essentially on par up to about 3,800 cores.

The combination of open source software (OpenFOAM), C5n instances, and EFA provides impressive HPC, at a price accessible to all.

CFD Direct, an organization that architects OpenFOAM, the well-known open source software for CFD, similarly performed benchmark simulations using EFA on various-sized clusters of C5n instances, and drew the following additional conclusions:

- C5n instances, with standard networking (ENA), deliver 70%-90% scaling at 504 cores.
- The addition of EFA delivered ~70% scaling at 1008 cores for meshes of 100 k cells per core.
- EFA delivered linear scaling at 1008 cores for a fixed mesh size of 97 m cells.
- With linear scaling, EFA enables faster solutions for the same cost simply by running on more cores.
- With open source software and C5n instances, EFA offers improved HPC at an affordable price point.

OpenFOAM v7.0 supports EFA for HPC workloads with Open MPI v3.1.4.

RELYING ON A BETTER METRIC FOR SUCCESS - TIME-TO-RESULTS

For some time, many HPC practitioners considered driving down latency as the best way to get data moving faster, similar to building a car that can get to the 100mph mark ever faster. AWS came at the problem from a different angle and focused on finding a way to move hundreds of millions of packets that real codes depend on — in other words, a better traffic infrastructure for all those drivers trying to get to work during rush hour — but not neglecting individual packet performance.

With EFA, you have the opportunity to revisit your metric for success on tightly coupled workloads and to re-examine previous conclusions about whether or not to -- or how to best -- move them to the cloud (Figure 4). On-premises solutions have finite capacity that usually result in long queues for jobs. With AWS, organizations can tap into massive capacity when needed to speed up time to results. EFA shortens the time-to-results of even your tightest coupled, still-on-premises HPC application making the decision to move to the cloud straightforward. And as how to quickly move those applications to the cloud, with EFA the answer is increasingly "as is," i.e., with no special effort. As always, you pay only for what you use.

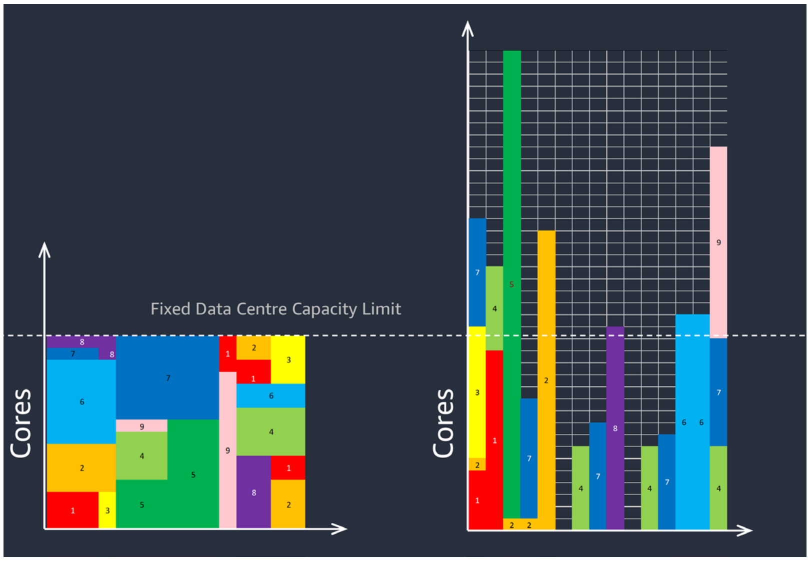

Figure 4: Time-to-results as the metric for success.

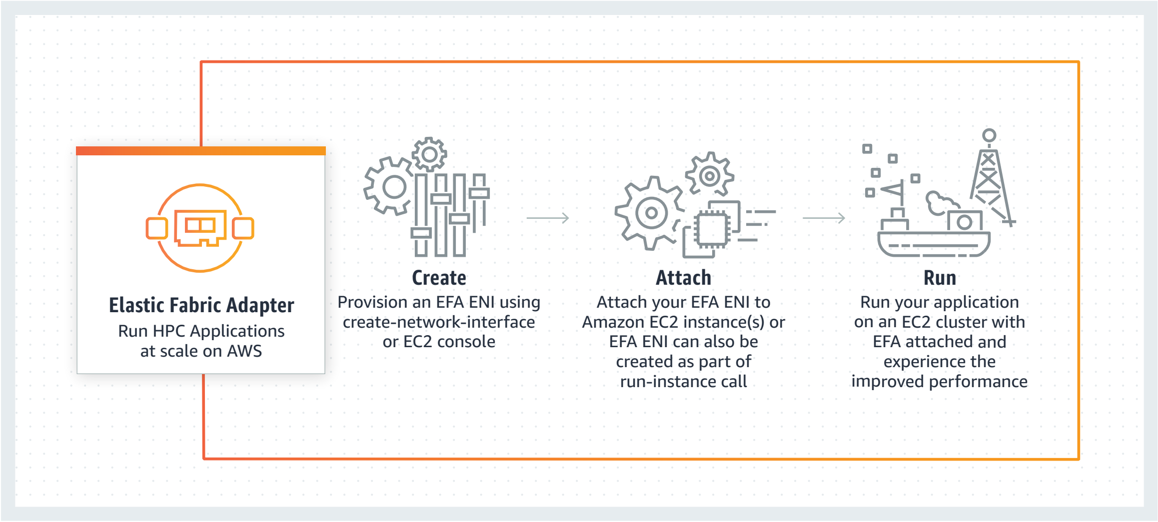

Figure 5: How to run your HPC application with EFA attached.

CONCLUSIONS

Using EFA to support tightly coupled workloads on AWS delivers new benefits while also maintaining the previous benefits of running HPC workloads on the AWS cloud.

- **Faster time-to-results -** Your tightly coupled HPC or distributed ML codes can now scale to thousands of cores thereby providing you with results faster.
- **Elasticity and flexibility -** Simply change your HPC cluster configurations as your needs change without prior reservations or much upfront planning and continue to use AWS services such as AWS ParallelCluster on AWS Batch for provisioning and scheduling.
- **Seamless migration -** Migrate your existing HPC applications to the AWS cloud with little to no modifications.

START AND LEARN MORE

HPC applications can benefit from EFA in a few easy steps and you need not be aware of implementation details. You can start with EFA very similarly to how you did with ENA. You can do so in all regions where EFA is generally available, an ever-growing list; for many select instance types, at no additional cost to the instance price; and for many supported Amazon Machine Images (AMIs), including Amazon Linux 2.

Note that if your HPC application supports the Message Passing Interface (MPI), you can use EFA with no application changes and have the choice between the open-source implementation of MPI, Open MPI, or the widely used Intel MPI implementation.

For applications not using MPI, EFA supports the vendor neutral Libfabric interface as a networking programming interface; many common HPC and ML packages natively support Libfabric.

AWS also provides a plug-in to enable your GPU ML application that uses NVIDIA's Collective Communication Library (NCCL) to use Libfabric as a network provider.

MPI and NCCL are commonly used packages in science, engineering, and machine learning applications and, to a lesser extent, distributed databases.

To learn more about AWS EFA, visit: https://aws.amazon.com/hpc/efa/

[1] Hyperion Research ROI study as funded by the U.S. Department of Energy.
[2] NAVGEM on the Cloud: Computational Evaluation of Commercial Cloud HPC with a Global Atmospheric Model, U.S. Naval Research Laboratory, 2019.
[3] For the expert, this case has around 140-million Hex-core cells and uses the realizable k-epsilon turbulence model as well as the Pressure-based coupled solver and the Least Squares cell based, pseudo-transient solver.
[4] https://www.ansys.com/de-de/solutions/solutions-by-role/it-professionals/platform-support/benchmarks-overview/benchmarking-terminology
[5] "OpenFOAM HPC with AWS EFA," CFD Direct, 2019.