Technical White Paper

Dell EMC Ready Solutions for HPC Digital Manufacturing with AMD EPYC™ Processors—ANSYS® Performance

Abstract

This Dell EMC technical white paper discusses performance benchmarking results and analysis for ANSYS® CFX® and Fluent® on the Dell EMC Ready Solutions for HPC Digital Manufacturing with AMD EPYC™ processors.

November 2019

Revisions

Date	Description
November 2019	Initial release with AMD EPYC™ 7002 series processors

Acknowledgements

This paper was produced by the following:

Authors: Joshua Weage
Martin Feyereisen

Table of contents

1 Introduction

This technical white paper discusses the performance of ANSYS® CFX® and Fluent® on the Dell EMC Ready Solutions for HPC Digital Manufacturing with AMD EPYC™ processors. This Ready Solution was designed and configured specifically for Digital Manufacturing workloads, where computer aided engineering (CAE) applications are critical for virtual product development. The Dell EMC Ready Solutions for HPC Digital Manufacturing uses a flexible building block approach to HPC system design, where individual building blocks can be combined to build HPC systems which are optimized for customer specific workloads and use cases.

The Dell EMC Ready Solutions for HPC Digital Manufacturing is one of many solutions in the Dell EMC HPC solution portfolio. Please visit www.dellemc.com/hpc for a comprehensive overview of the available HPC solutions offered by Dell EMC.

The architecture of the Dell EMC Ready Solutions for HPC Digital Manufacturing and a description of the building blocks are presented in Section 2. Section 3 describes the system configuration, software and application versions that were used to measure and analyze the performance of the Dell EMC HPC Ready Solutions for HPC Digital Manufacturing with AMD EPYC processors. Section 4 presents benchmark performance for ANSYS CFX and Section 5 presents benchmark performance for ANSYS Fluent.

2 System Building Blocks

The Dell EMC Ready Solutions for HPC Digital Manufacturing is designed using preconfigured building blocks. The building block architecture allows an HPC system to be optimally designed for specific end-user requirements, while still making use of standardized, domain-specific system recommendations. The available building blocks are infrastructure servers, storage, networking, and compute building blocks. Configuration recommendations are provided for each of the building blocks which provide good performance for typical applications and workloads within the manufacturing domain. This section describes the available building blocks along with the recommended server configurations.

With this flexible building block approach, appropriately sized HPC clusters can be designed based on individual customer workloads and requirements. Figure 1 shows three example HPC clusters designed using the Dell EMC Ready Solutions for HPC Digital Manufacturing architecture.

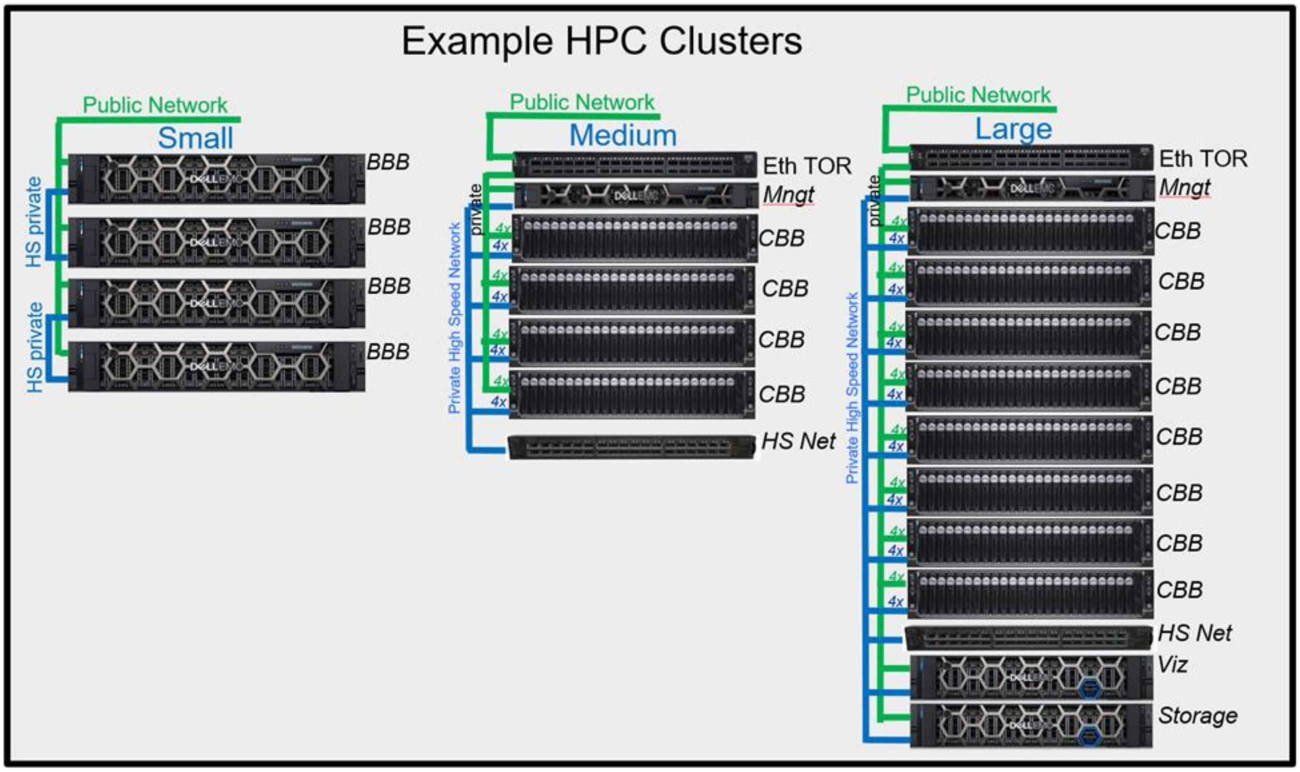

Figure 1 Example Ready Solutions for HPC Digital Manufacturing

2.1 Infrastructure Servers

Infrastructure servers are used to administer the system and provide user access. They are not typically involved in computation, but they provide services that are critical to the overall HPC system. These servers are used as the master nodes and the login nodes. For small sized clusters, a single physical server can provide the necessary system management functions. Infrastructure servers can also be used to provide storage services, by using NFS, in which case they must be configured with additional disk drives or an external storage array. One master node is mandatory for an HPC system to deploy and manage the system. If high-availability (HA) management functionality is required, two master nodes are necessary. Login nodes are optional and one login server per 30-100 users is recommended.

A recommended base configuration for infrastructure servers is:

- Dell EMC PowerEdge R6515 server
- AMD EPYC 7302P processor
- 128GB of RAM (8 x 16GB 3200 MTps DIMMs)
- PERC H345 RAID controller
- 2 x 480GB Mixed Use SATA SSD RAID 1
- Dell EMC iDRAC9 Enterprise
- 2 x 550W Power Supplies
- Mellanox ConnectX-6 InfiniBand™ HCA (optional)

The recommended base configuration for the infrastructure server is described as follows. The PowerEdge R6515 server is suited for this role. Typical HPC clusters will only require a few infrastructure servers; therefore, density is not a priority, but manageability is important. The AMD EPYC 7302P processor, with 16 cores per socket, is suitable for this role. If the infrastructure server will be used for CPU intensive tasks, such as compiling software or processing data, then a more capable processor may be appropriate. 128GB of RAM provided by eight 16GB DIMMs provides sufficient memory capacity, while also providing good memory bandwidth. These servers are not expected to perform much I/O, but reliability is important, so two mixed use SATA SSDs in RAID1 configuration are recommended for the operating system. For small systems (typically four nodes or less), an Ethernet network may provide sufficient application performance. For most other systems, InfiniBand is likely to be the data interconnect of choice, which provides a high-throughput, low-latency fabric for node-to-node communications or access to Dell EMC HPC Storage Ready Solutions.

2.2 Compute Building Blocks

Compute Building Blocks (CBB) provide the computational resources for most HPC systems for Digital Manufacturing. These servers are used to run the ANSYS CFX or Fluent simulations. The best configuration for these servers depends on the specific mix of applications and types of simulations being performed. Since the best configuration may be different for each customer, a table of recommended options are provided that are appropriate for these servers. The configuration can be selected based on the specific system and workload requirements of each customer. Relevant criteria to consider when selecting a system configuration are discussed in the application performance chapters of this white paper. The recommended configuration options for the Compute Building Block are provided in Table 1.

Table 1 Recommended Configurations for the Compute Building Block

Platforms	Dell EMC PowerEdge R6525 Dell EMC PowerEdge C6525
Processor Options	Dual AMD EPYC 7402 (24 cores per socket) Dual AMD EPYC 7452 (32 cores per socket) Dual AMD EPYC 7502 (32 cores per socket) Dual AMD EPYC 7552 (48 cores per socket) Dual AMD EPYC 7702 (64 cores per socket)
Memory Options	256 GB (16 x 16GB 3200 MTps DIMMs) 512 GB (16 x 32GB 3200 MTps DIMMs)
Local Storage	1 x 480GB Mixed Use SATA SSD
iDRAC	iDRAC9 Enterprise (R6525) iDRAC9 Express (C6525)
Power Supplies	2 x 800W PSU (R6525) 2 x 2400W PSU (C6525)
Networking	Mellanox® ConnectX®-6 InfiniBand™ adapter

2.3 Basic Building Blocks

Basic Building Block (BBB) servers are selected by customers to create simple but powerful HPC systems. These servers are appropriate for smaller HPC systems where reducing the management complexity of the HPC system is important. The BBB is based on the dual socket Dell EMC PowerEdge R6525 server.

The recommended configuration for BBB servers is:

- Dell EMC PowerEdge R6525 server
- Dual AMD EPYC 7502 32-Core processors
- 256 GB of RAM (16 x 16GB 3200 MTps DIMMS)
- PERC H745 RAID controller
- 2 x 240GB Read-Intensive SATA SSD RAID 1 (OS)
- 4 x 480GB Mixed Use SATA SSD RAID 0 (scratch)
- Dell EMC iDRAC9 Enterprise
- 2 x 800W Power Supplies
- Mellanox ConnectX-6 InfiniBand Adapter (optional)
- Mellanox ConnectX-5 25 GbE Adapter (optional)

The R6525 platform is used to provide configuration flexibility and good compute power per server. Each server can contain up to two AMD EPYC processors. The AMD EPYC 7502 processor is a 32-core CPU with a base frequency of 2.5 GHz and a max boost frequency of 3.35 GHz. As configured, a BBB contains 64 cores, a natural number for many CAE simulations. A memory configuration of 16 x 16GB DIMMs is used to provide balanced performance and capacity. While 256GB is typically sufficient for most CAE workloads, customers expecting to handle larger production jobs should consider increasing the memory capacity to 512GB. Various CAE applications, such as implicit FEA, often have large file system I/O requirements and four mixed use SATA SSD’s in RAID 0 are used to provide fast local I/O.

Additionally, two BBB's can be directly coupled together via a high-speed network cable, such as InfiniBand or Ethernet, without need of an additional high-speed switch if additional compute capability is required for each simulation run (BBB Couplet). BBB's provide a simple framework for customers to incrementally grow the size and power of the HPC cluster by purchasing individual BBBs, BBB Couplets, or combining the individual and/or Couplets with a high-speed switch into a single monolithic system.

We did not carry out any explicit performance testing on BBB configurations for this paper. For Linux based systems, the single node and two-node couplet BBB clusters with InfiniBand would be comparable to the results reported for the two-node 7452 based CBB benchmarks below. For Windows based clusters, the use of InfiniBand for node-to-node connectivity in a two-node couplet requires complex setup and administration, likely beyond the intended scope for most customers seeking a Windows based solution. It is recommended that Windows based two-node couplets be networked with high speed Ethernet, such as 25GbE. Our experience with Windows for HPC workloads indicates the performance differential between Windows and Linux can be highly variable and problem dependent, making the use of standard benchmarks as an indication of projected performance of limited value. Customers wishing for the highest level of performance and potential cluster expansion would be advised to use Linux as an operating system.

2.4 Storage

Dell EMC offers a wide range of HPC storage solutions. For a general overview of the entire HPC solution portfolio please visit www.dellemc.com/hpc. There are typically three tiers of storage for HPC: scratch storage, operational storage, and archival storage, which differ in terms of size, performance, and persistence.

Scratch storage tends to persist for the duration of a single simulation. It may be used to hold temporary data which is unable to reside in the compute system's main memory due to insufficient physical memory capacity. HPC applications may be considered "I/O bound" if access to storage impedes the progress of the simulation. For these HPC workloads, typically the most cost-effective solution is to provide sufficient direct-attached local storage on the compute nodes. For situations where the application may require a shared file system across the compute cluster, a high-performance shared file system may be better suited than relying on local direct-attached storage. Typically, using direct-attached local storage offers the best overall price/performance and is considered best practice for most CAE simulations. For this reason, local storage is included in the recommended configurations with appropriate performance and capacity for a wide range of production workloads. If anticipated workload requirements exceed the performance and capacity provided by the recommended local storage configurations, care should be taken to size scratch storage appropriately based on the workload.

Operational storage is typically defined as storage used to maintain results over the duration of a project and other data, such as home directories, such that the data may be accessed daily for an extended period of time. Typically, this data consists of simulation input and results files, which may be transferred from the scratch storage or from users analyzing the data, often remotely. Since this data may persist for an extended period, some or all of it may be backed up at a regular interval, where the interval chosen is based on the balance of the cost to either archive the data or regenerate it if need be. Archival data is assumed to be persistent for a very long term, and data integrity is considered critical. For many modest HPC systems, use of the existing enterprise archival data storage may make the most sense, as the performance aspect of archival data tends to not impede HPC activities. Our experience in working with customers indicates that there is no 'one size fits all' operational and archival storage solution. Many customers rely on their corporate enterprise storage for archival purposes and instantiate a high-performance operational storage system dedicated for their HPC environment.

Operational storage is typically sized based on the number of expected users. For fewer than 30 users, a single storage server, such as the Dell PowerEdge R7515 is often an appropriate choice. A suitably equipped storage server may be:

- Dell EMC PowerEdge R7515 server
- AMD EPYC 7302P processor
- 128GB of RAM (8 x 16GB 3200 MTps DIMMs)
- PERC H745 RAID controller
- 2 x 240GB Mixed Use SATA SSD in RAID-1 (For OS)
- 12 x 12TB 3.5” NLSAS HDDs in RAID-6 (for data)
- Dell EMC iDRAC9 Enterprise
- 2 x 750W Power Supplies
- Mellanox ConnectX-6 InfiniBand Adapter (optional)
- High-speed Ethernet Adapter (optional)

This server configuration would provide 144TB of raw storage. For customers expecting between 25-100 users, an operational storage solution, such as the Dell EMC Ready Solutions for HPC NFS Storage (NSS), shown in Figure 2, with up 1008TB of raw storage capacity may be appropriate:

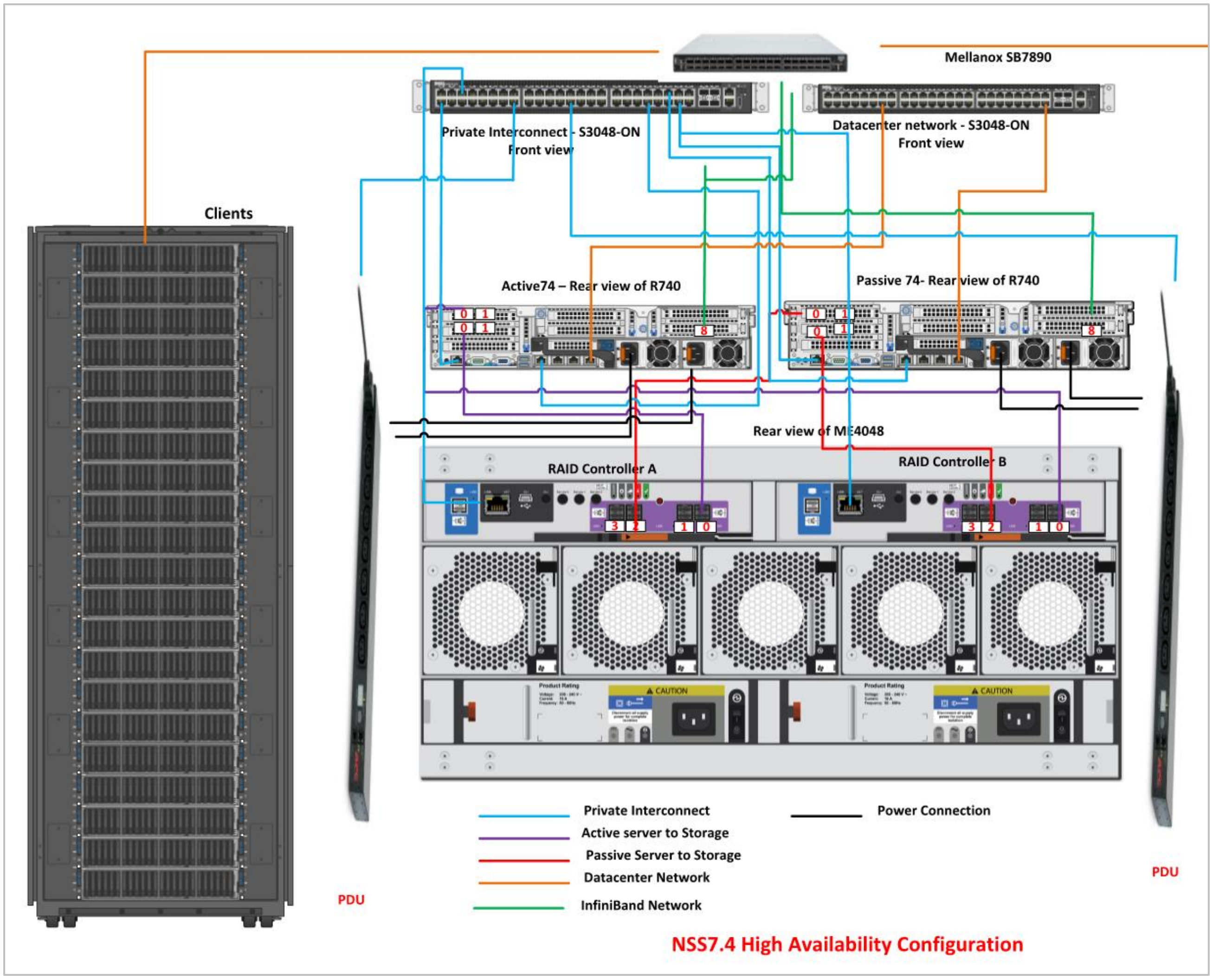

Figure 2 NSS7.4-HA Storage System Architecture

For customers desiring a shared high-performance parallel filesystem, the Dell MC Ready Solutions for HPC Lustre Storage solution shown in Figure 3 is appropriate. This solution can scale up to multiple petabytes of storage.

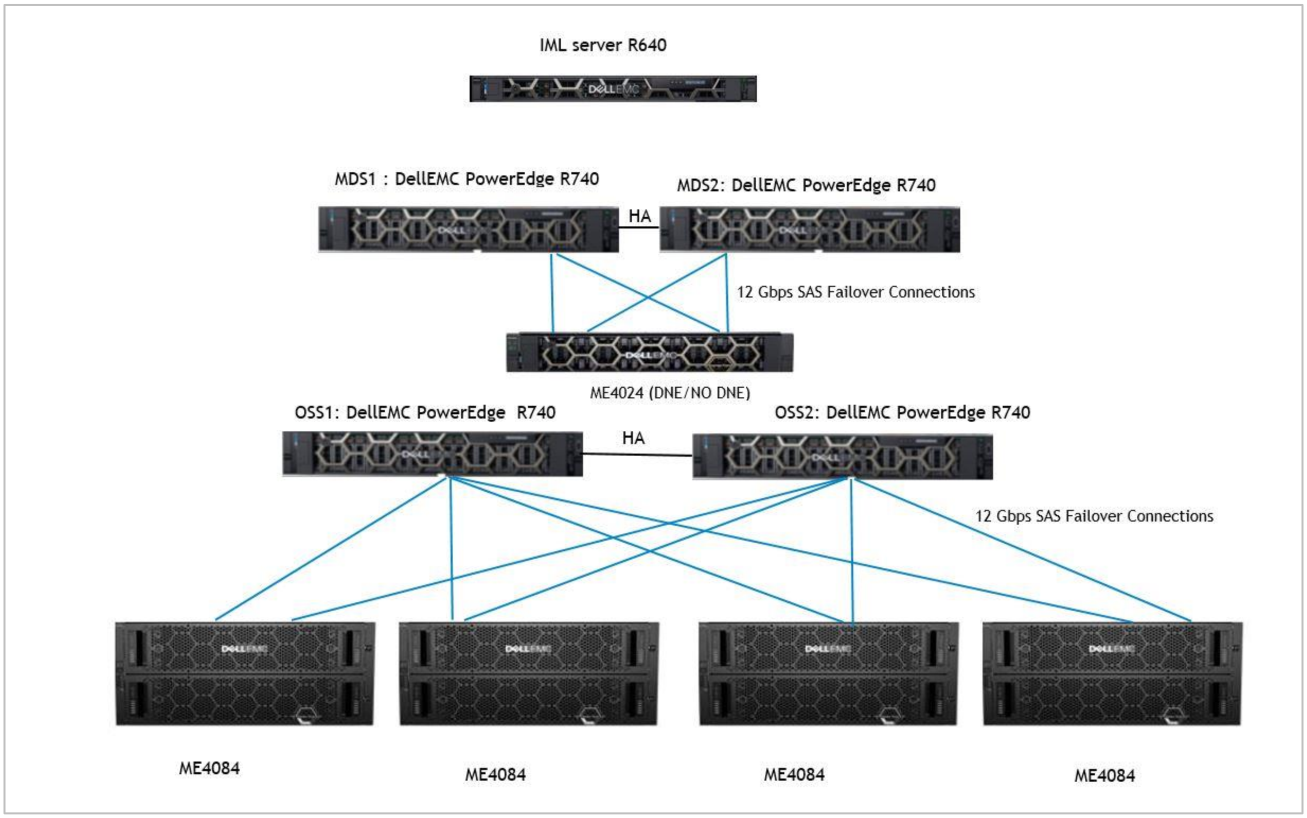

Figure 3 Dell EMC Ready Solutions for Lustre Storage Reference Architecture

2.5 System Networks

Most HPC systems are configured with two networks—an administration network and a high-speed/low-latency switched fabric. The administration network is typically Gigabit Ethernet that connects to the onboard LOM/NDC of every server in the cluster. This network is used for provisioning, management and administration. On the CBB servers, this network is also normally used for IPMI hardware management. For infrastructure and storage servers, the iDRAC Ethernet interfaces may be connected to this network for OOB server management. The management network is typically deployed using Dell Networking S3048-ON Ethernet switches. If there is more than one switch in the system, multiple switches can be stacked with 10 Gigabit Ethernet cables.

A high-speed/low-latency fabric is recommended for clusters with more than four servers. The current recommendation is a Mellanox HDR or EDR InfiniBand fabric. The fabric will typically be assembled using Mellanox QB8700 series or SB7800 series InfiniBand switches. The number of switches required depends on the size of the cluster and the blocking ratio of the fabric.

2.6 Cluster Management Software

The cluster management software is used to install and monitor the HPC system. Bright Cluster Manager (BCM) is the recommended cluster management software.

2.7 Services and Support

The Dell EMC Ready Solutions for HPC Digital Manufacturing is available with full hardware support and deployment services, including additional HPC system support options.

3 Reference System

Performance benchmarking was performed in the Dell EMC HPC and AI Innovation Lab using system configurations as listed in Table 2.

Table 2 Benchmark System Configurations

Building Block	Quantity
Compute Building Block (CBB) PowerEdge C6525 Dual AMD EPYC 7452 32-Core Processors 256GB RAM 16x16GB 3200 MTps DIMMs Mellanox ConnectX-6 HDR100 InfiniBand	8
Compute Building Block (CBB) PowerEdge C6525 Dual AMD EPYC 7402 24-Core Processors 256GB RAM 16x16GB 3200 MTps DIMMs	1
Compute Building Block (CBB) PowerEdge C6525 Dual AMD EPYC 7502 32-Core Processors 256GB RAM 16x16GB 3200 MTps DIMMs	1
Compute Building Block (CBB) PowerEdge C6525 Dual AMD EPYC 7702 64-Core Processors 256GB RAM 16x16GB 3200 MTps DIMMs	1
Mellanox QM8700 InfiniBand Switch	1

The BIOS configuration options used for the benchmark systems are listed in Table 3.

Table 3 BIOS Configuration

BIOS Option	Setting
Logical Processor	Disabled
Virtualization Technology	Disabled
System Profile	Performance Profile
CCX as NUMA Domain	Disabled
NUMA Nodes Per Socket	4

The software versions used for the benchmarks are listed in Table 4.

Table 4 Software Versions

Component	Version
Operating System	RedHat Enterprise Linux 7.6
Kernel	3.10.0-957.27.2.el7.x86_64
OFED	Mellanox 4.6-1.0.1.1
Bright Cluster Manager	8.2
ANSYS CFX	2019R3
ANSYS Fluent	2019R3

4 ANSYS CFX Performance

ANSYS CFX software is a computational fluid dynamics (CFD) application recognized for its accuracy, robustness and speed with rotating machinery applications. CFD applications typically scale well across multiple processor cores and servers, have modest memory capacity requirements, and typically perform minimal disk I/O while in the solver section. However, some simulations may have greater I/O demands, such as transient analysis. Figure 4 shows the measured performance of the standard CFX benchmarks using CFX 2019R3 on a single server. The performance for each benchmark is measured using the solver elapsed time.

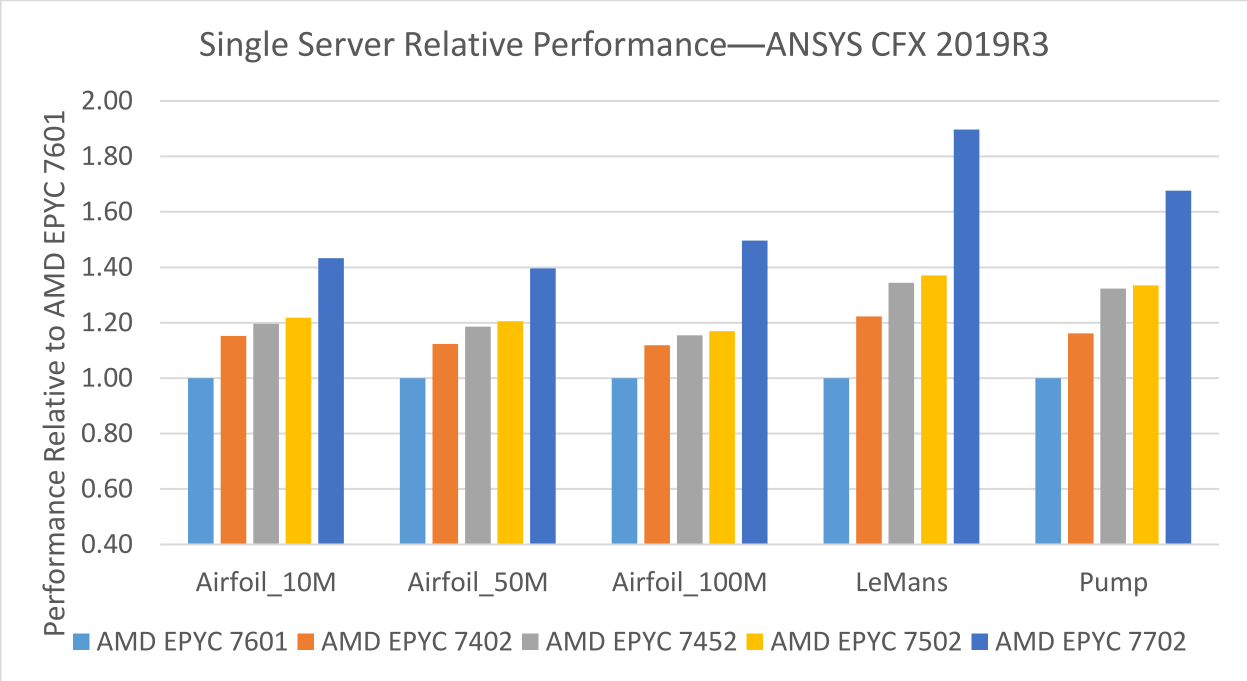

Figure 4 Single Server Relative Performance—ANSYS CFX

The results presented in Figure 4 are single-server performance results, with the benchmark run using all processor cores available in the server. The results are plotted relative to the performance of a single server configured with AMD EPYC 7601 processors. A larger value indicates better performance. These results show the performance advantage available with AMD EPYC 7002 series processors. The 32-core AMD EPYC 7452 and 7502 processors provide very good performance for these benchmarks. The 64-core AMD EPYC 7702 provides a significant performance advantage over the 32-core processors, particularly for the LeMans and Pump benchmark cases. This may be due to the larger L3 cache provided by the 7702. The effect of cache can be significant, and the relative performance may change as more servers are used and more of the data is contained in cache. Because of this behavior, the relative performance of the various processor models depends on the specific benchmark case and the number of servers used for the simulation.

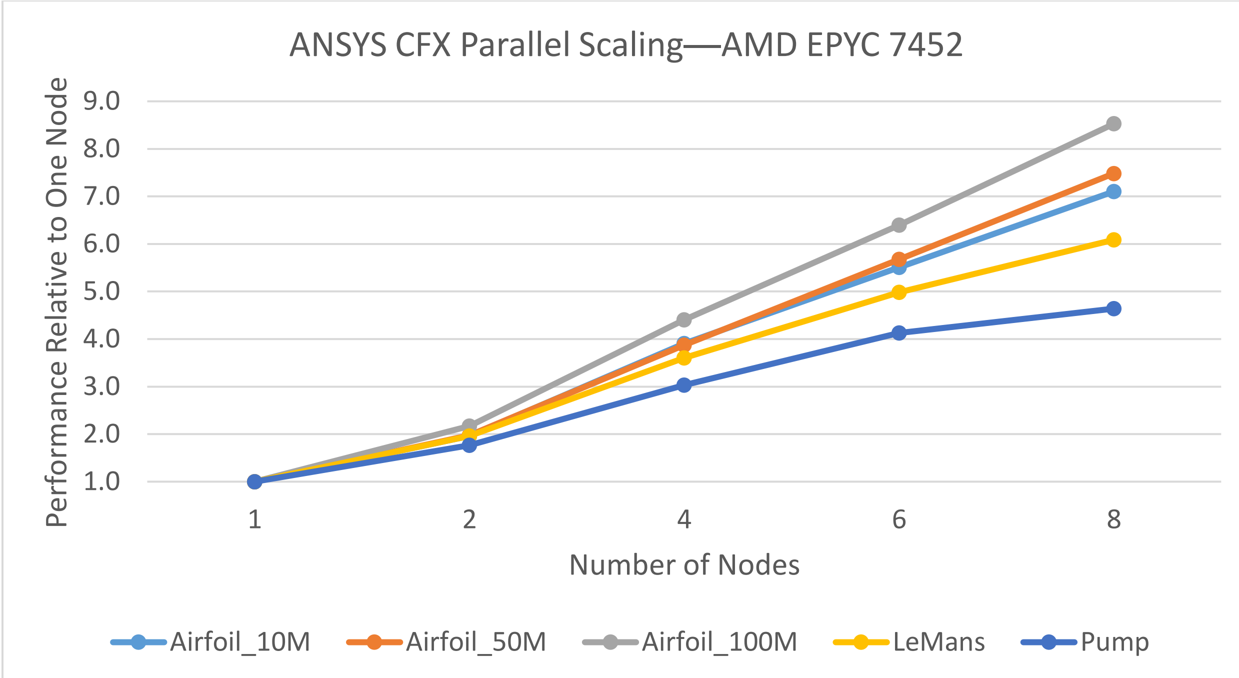

Figure 5 ANSYS CFX Parallel Scaling—AMD EPYC 7452

Figure 5 presents the parallel scalability when running CFX with up to eight CBB nodes configured with AMD EPYC 7452 processors and ConnectX-6 HDR100 InfiniBand. All processor cores in each server were used when running these benchmarks. The performance at each node count is presented relative to the performance of a single node.

The parallel scalability for these benchmarks is as expected, with the largest model demonstrating super-linear scaling. Super-linear scaling is typically observed when more of the active data fits into the processor cache at larger node counts. For the Airfoil_100M case, super linear scaling is observed starting at four nodes.

5 ANSYS Fluent Performance

ANSYS Fluent is a computational fluid dynamics (CFD) application commonly used across a very wide range of CFD and multi-physics applications. CFD applications typically scale well across multiple processor cores and servers, have modest memory capacity requirements and typically perform minimal disk I/O while in the solver section. However, some simulations may have greater I/O demands, such as transient analysis.

Fluent benchmark performance is measured using the Solver Rating metric which is the number of 25 iteration solves that can be completed in a day. That is, (total seconds in a day)/(25 iteration solve time in seconds). A larger value represents better performance. Figure 6 shows the measured performance for six of the ANSYS Fluent benchmarks on a single server.

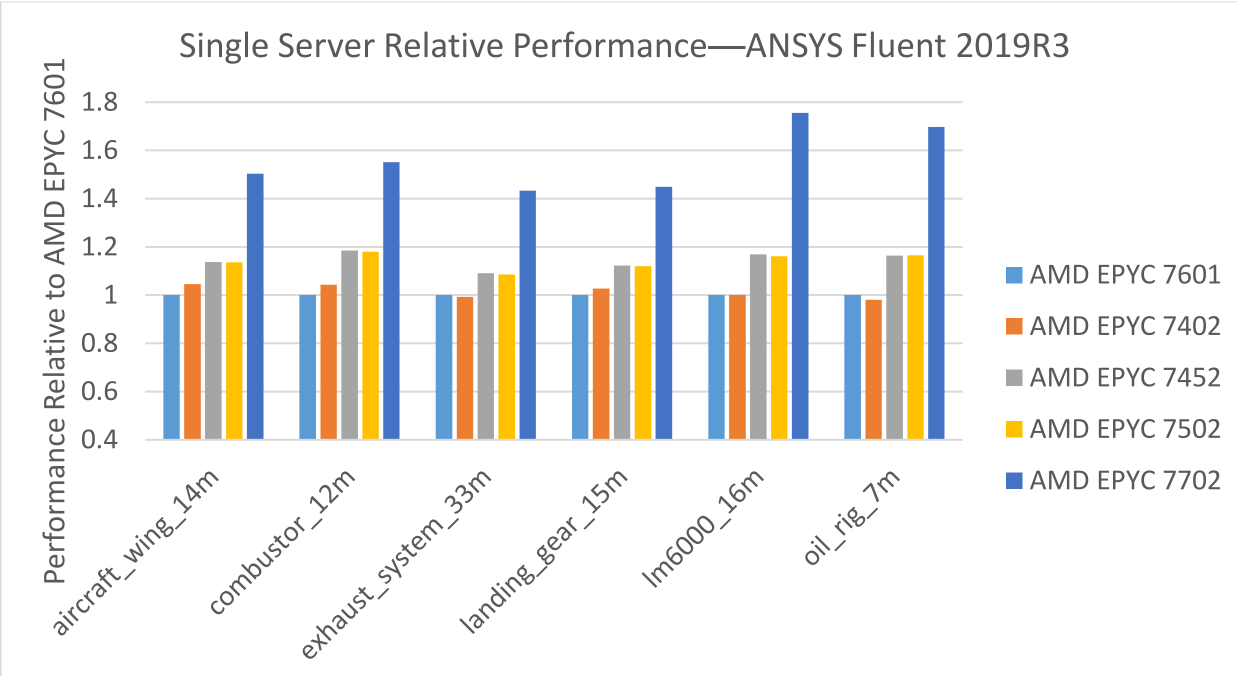

Figure 6 Single Server Relative Performance—ANSYS Fluent

The results presented in Figure 6 are single-server performance results, with the benchmark run using all processor cores available in the server. The results are plotted relative to the performance of a single server configured with AMD EPYC 7601 processors. A larger value indicates better performance. These results show the performance advantage available with AMD EPYC 7002 series processors. The 32-core AMD EPYC 7452 and 7502 processors provide good performance for these benchmarks. Per server, the 64-core AMD EPYC 7702 provides a significant performance advantage over the 32-core processors for these benchmark cases. Note that the relative performance of the AMD EPYC 7702 depends on the specific benchmark case. This may be due to the larger L3 cache provided by the 7702. The effect of cache can be significant, and the relative performance may change as more servers are used and more of the data is contained in cache. Because of this behavior, the relative performance of the various processor models depends on the specific benchmark case and the number of servers used for the simulation.

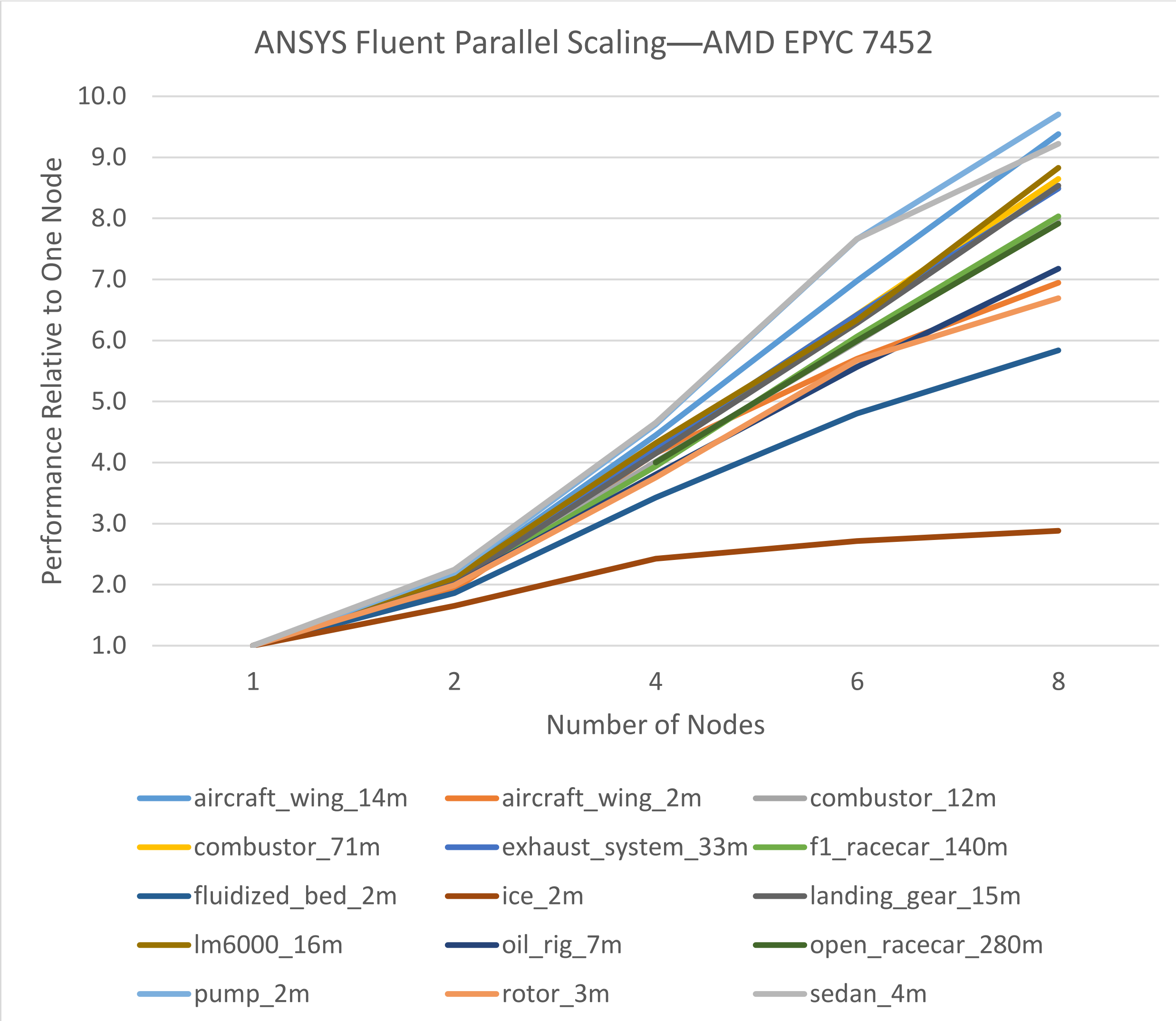

Figure 7 ANSYS Fluent Parallel Scaling—AMD EPYC 7452

Figure 7 presents the parallel scalability when running Fluent with up to eight CBB nodes configured with AMD EPYC 7452 processors and ConnectX-6 HDR100 InfiniBand. All processor cores in each server were used when running these benchmarks. The performance at each node count is presented relative to the performance of a single node.

The parallel scalability for these benchmarks is as expected, with some models demonstrating super-linear scaling. Super-linear scaling is typically observed when more of the active data fits into processor cache as the number of nodes used increases. The ice_2m benchmark does not scale as well as the other benchmarks, but this is expected as it is a small model which includes dynamic mesh and combustion simulation.

6 Conclusion

This technical white paper presents the Dell EMC Ready Solutions for HPC Digital Manufacturing with AMD EPYC 7002 Series processors. The detailed analysis of the building block configurations demonstrate that the system is architected for a specific purpose—to provide a comprehensive HPC solution for the manufacturing domain. Use of this building block approach allows customers to easily deploy an HPC system optimized for their specific workload requirements. The design addresses computation, storage, networking and software requirements and provides a solution that is easy to install, configure and manage, with installation services and support readily available. The performance benchmarking bears out the solution design, demonstrating the performance of the solution with ANSYS CFX and Fluent.