DELL EMC

Technical White Paper

Dell EMC Ready Solution for HPC Digital Manufacturing— Altair Performance

Abstract

This Dell EMC technical white paper discusses performance benchmarking results and analysis for Altair HyperWorks™ on the Dell EMC Ready Solution for HPC Digital Manufacturing.

April 2019

Revisions

Date	Description
April 2019	Initial release

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 various Altair HyperWorks™ products, including Altair OptiStruct™, Altair Radioss™, Altair AcuSolve™, and Altair Feko™, on the Dell EMC Ready Solution for HPC Digital Manufacturing, with benchmarking workload management by Altair PBS Professional™. This Dell EMC Ready Solution for HPC 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 Solution 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 Solution 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 Solution 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, and the benchmark test cases that were used to measure and analyze the performance of the Dell EMC Ready Solution for HPC Digital Manufacturing. Section 4 quantifies the capabilities of the system and presents benchmark performance for AcuSolve. Section 5 contains the performance data for Radioss. Section 6 contains the performance information for OptiStruct. Section 7 contains the performance data for Feko, and Section 8 contains performance information for the Basic Building Block.

2 System Building Blocks

The Dell EMC Ready Solution 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.

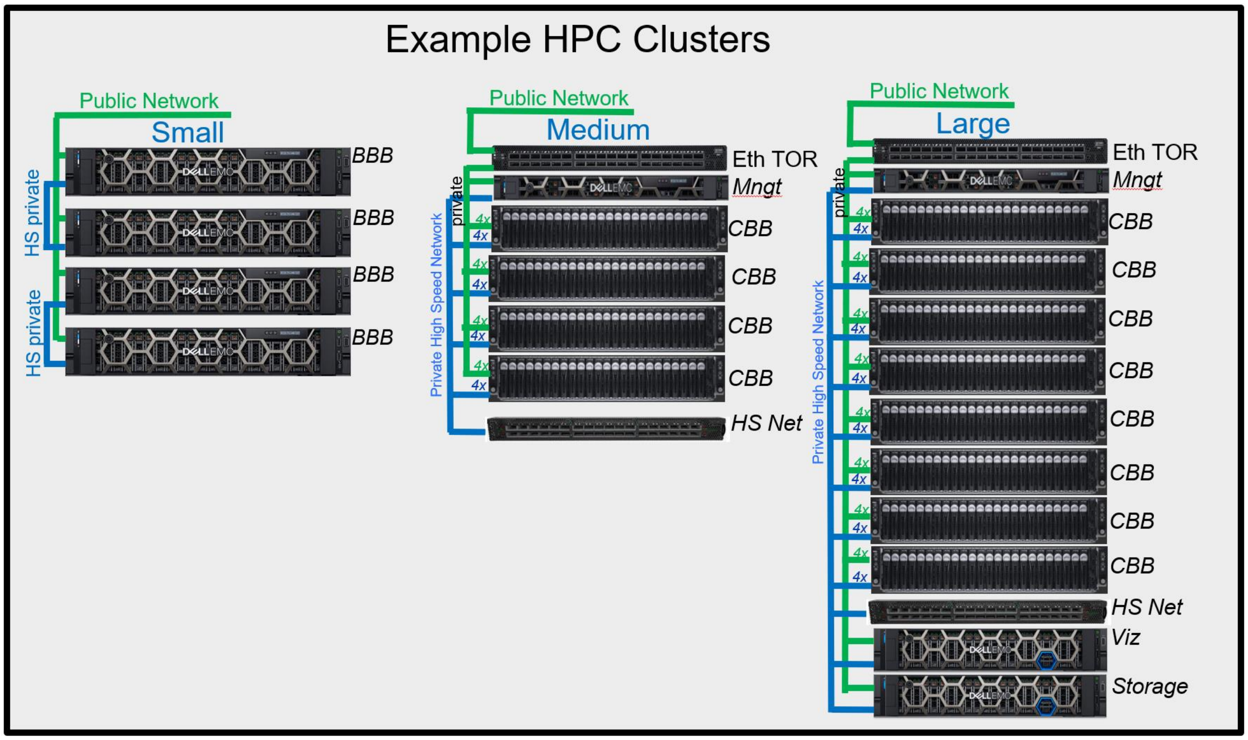


Figure 1 Example Ready Solutions for HPC Digital Manufacturing

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 R640 server
- Dual Intel® Xeon® Bronze 3106 processors
- 192 GB of RAM (12 x 16GB 2667 MTps DIMMs)
- PERC H330 RAID controller
- 2 x 480GB Mixed-Use SATA SSD RAID 1
- Dell EMC iDRAC9 Enterprise
- 2 x 750 W power supply units (PSUs)
- Mellanox EDR InfiniBand™ (optional)

The recommended base configuration for the infrastructure server is described as follows. The PowerEdge R640 server is suited for this role. Typical HPC clusters will only use a few infrastructure servers; therefore, density is not a priority, but manageability is important. The Intel Xeon Bronze 3106 processor, with 8 cores per socket, is a basic recommendation 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. 192 GB of memory provided by twelve 16 GB DIMMs provides sufficient memory capacity, with minimal cost per GB, while also providing good memory bandwidth. These servers are not expected to perform much I/O, so a single mixed-use SATA SSD should be sufficient for the operating system. For small systems (four nodes or less), an Ethernet network may provide sufficient application performance. For most other systems, EDR 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 a Dell EMC Ready Solution for HPC NFS Storage solution or a Dell EMC Ready Solution for HPC Lustre Storage solution.

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 Altair Hyperworks simulations. The best configuration for these servers depends on the specific mix of applications and types of simulations being performed by each customer. Since the best configuration may be different for each customer, a table of recommended options are provided that are appropriate for these servers. The specific configuration can then be selected based on the specific system and workload requirements of each customer. Relevant criteria to consider when making these selections 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 R640 Dell EMC PowerEdge C6420
Processors	Dual Intel Xeon Gold 6242 (16 cores per socket) Dual Intel Xeon Gold 6248 (20 cores per socket) Dual Intel Xeon Gold 6252 (24 cores per socket)
Memory Options	192 GB (12 x 16GB 2933 MTps DIMMs) 384 GB (12 x 32GB 2933 MTps DIMMs) 768 GB (24 x 32GB 2933 MTps DIMMs, R640 only)
Storage Options	PERC H330, H730P or H740P RAID controller 2 x 480GB Mixed-Use SATA SSD RAID 0 4 x 480GB Mixed-Use SATA SSD RAID 0
iDRAC	iDRAC9 Enterprise (R640) iDRAC9 Express (C6420)
Power Supplies	2 x 750W PSU (R640) 2 x 2000W PSU (C6400)
Networking	Mellanox® ConnectX®-5 EDR 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 4-socket Dell EMC PowerEdge R840 server.

The recommended configuration for BBB servers is:

- Dell EMC PowerEdge R840 server
- Quad Intel Xeon Gold 6242 processors
- 384 GB of RAM (24 x 16GB 2933 MTps DIMMS)
- PERC H740P 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 1600W power supply units (PSUs)
- Mellanox ConnectX-5 EDR InfiniBand (optional)
- Mellanox 25 GbE (optional)

The R840 platform is used to minimize server count and provide good compute power per server. Each server can contain up to four Intel Xeon processors, where each BBB is essentially two CBB's fused into a single server. The Intel Xeon Gold 6142 processor is a sixteen-core CPU with a base frequency of 2.6 GHz and a max all-core turbo frequency of 3.3 GHz. With four processors, a BBB contains 64 cores, a natural number for many CAE simulations. A memory configuration of 24 x 16GB DIMMs is used to provide balanced performance and capacity. While 384GB is typically sufficient for most CAE workloads, customers expecting to handle larger production jobs should consider increasing the memory capacity to 768GB. 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. The compute nodes do not normally require extensive OOB management capabilities; therefore, an iDRAC9 Express is recommended.

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 (HPC 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. Performance testing for BBB's has been done using Windows Server 2016. In general, Linux provides better overall performance, and an easier path to combining BBB's to create larger, more capable HPC clusters. 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 for most CAE simulations offers the best overall price/performance and is considered best practice. 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, typically in a sequential manner, 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 ends 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 R740xd is often an appropriate choice. A suitable equipped storage server may be:

- Dell EMC PowerEdge R740xd server
- Dual Intel® Xeon® Bronze 4110 processors
- 96 GB of memory, 12 x 8GB 2667 MT/s DIMMS
- PERC H730P RAID controller
- 2 x 250GB Mixed-use SATA SSD in RAID-1 (For OS)
- 12 x 12TB 3.5: nlSAS HDDs in RAID-6 (for data)
- Dell EMC iDRAC9 Express
- 2 x 750 W power supply units (PSUs)
- Mellanox EDR InfiniBand Adapter
- Site specific 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 Solution for HPC NFS Storage (NSS), shown in Figure 2, with up 840 TB of raw storage of storage may be appropriate:

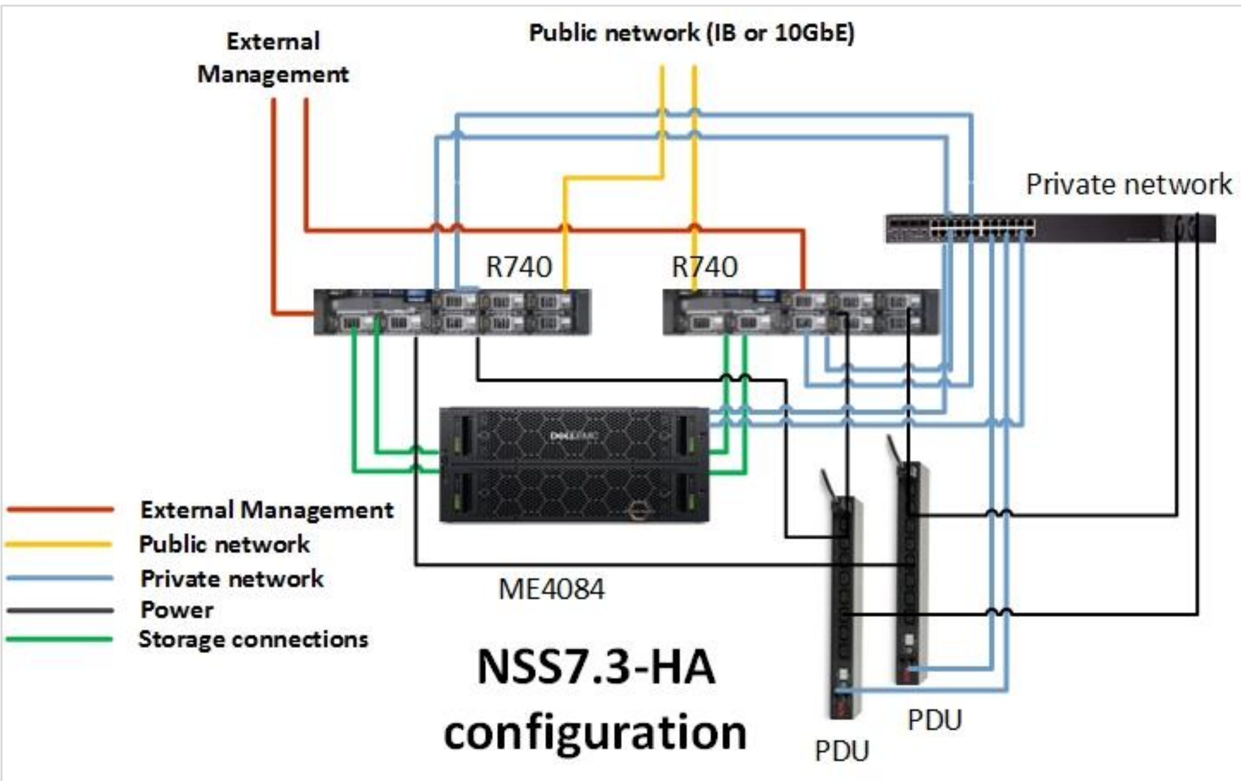


Figure 2 NSS7.3-HA Storage System Architecture

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

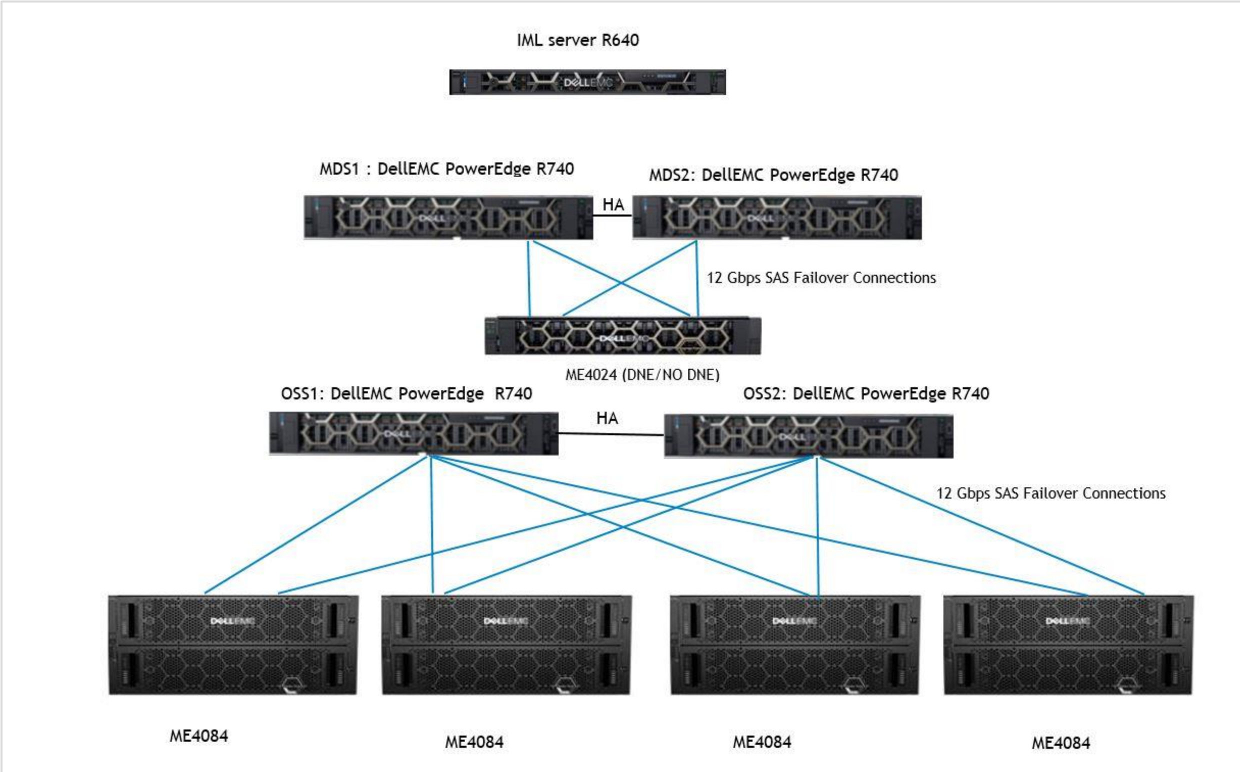


Figure 3 Dell EMC Ready Solution for Lustre Storage Reference Architecture

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 will also be used for IPMI hardware management. For infrastructure and storage servers, the iDRAC Enterprise ports may be connected to this network for OOB server management. The management network typically uses the Dell Networking S3048-ON Ethernet switch. If there is more than one switch in the system, multiple switches should 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 an EDR InfiniBand fabric. The fabric will typically be assembled using Mellanox SB7890 36-port EDR 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 Solution for HPC Digital Manufacturing is available with full hardware support and deployment services, including additional HPC system support options.

2.8 Workload Management

Workload management and job scheduling on the Dell EMC Ready Solution for HPC Digital Manufacturing can be handled efficiently with Altair PBS Professional, part of the Altair PBS Works™ suite. PBS Professional features include policy-based scheduling, OS provisioning, shrink-to-fit jobs, preemption, and failover. Its topology-aware scheduling optimizes task placement, improving application performance and reducing network contention. Many applications from Altair and other leading ISV's incorporate PBS Professional specific interfaces into their applications to improve the ease of use and performance for these applications with PBS Professional.

3 Reference System

The reference system was assembled in the Dell EMC HPC and AI Innovation Lab using the building blocks described in section 2. The building blocks used for the reference system are listed in Table 2.

Table 2 Reference System Configuration

Building Block	Quantity
Infrastructure Server	1
Computational Building Block (CBB) PowerEdge C6420 Dual Intel Xeon Gold 6242 192GB RAM 12x16GB 2933 MTps DIMMs Mellanox ConnectX-5 EDR adapter	2
Computational Building Block (CBB) PowerEdge C6420 Dual Intel Xeon Gold 6252 192 GB RAM 12x16GB 2933 MTps DIMMs Mellanox ConnectX-5 EDR adapter	8
Basic Building Block	2
Dell Networking S3048-ON Ethernet Switch	1
Mellanox SB7700 EDR InfiniBand Switch	1

The BIOS configuration options used for the reference system are listed in Table 3.

Table 3 BIOS Configuration

BIOS Option	Setting
Logical Processor	Disabled
Virtualization Technology	Disabled
System Profile	Performance Profile
Sub NUMA Cluster	Enabled (CBB) Disabled (BBB)

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

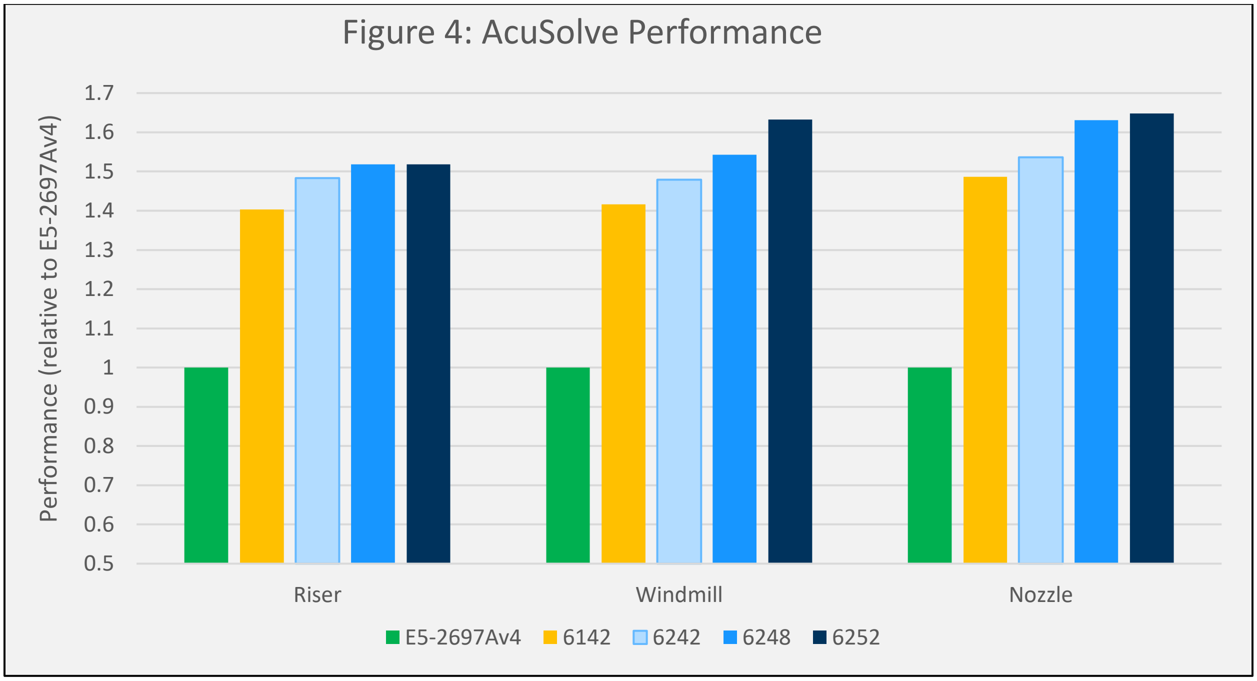
Table 4 Software Versions

Component	Version
Operating System	RHEL 7.6 Windows Server 2016 (BBB)
Kernel	3.10.0-957.el7.x86_64
OFED	Mellanox 4.5-1.0.1.0
Bright Cluster Manager	8.2
Altair PBS Professional	18.1.2
Altair OptiStruct	2017.2.3
Altair Radioss	2017.2.3
Altair AcuSolve	2017.2
Altair Feko	2017

4 Altair AcuSolve Performance

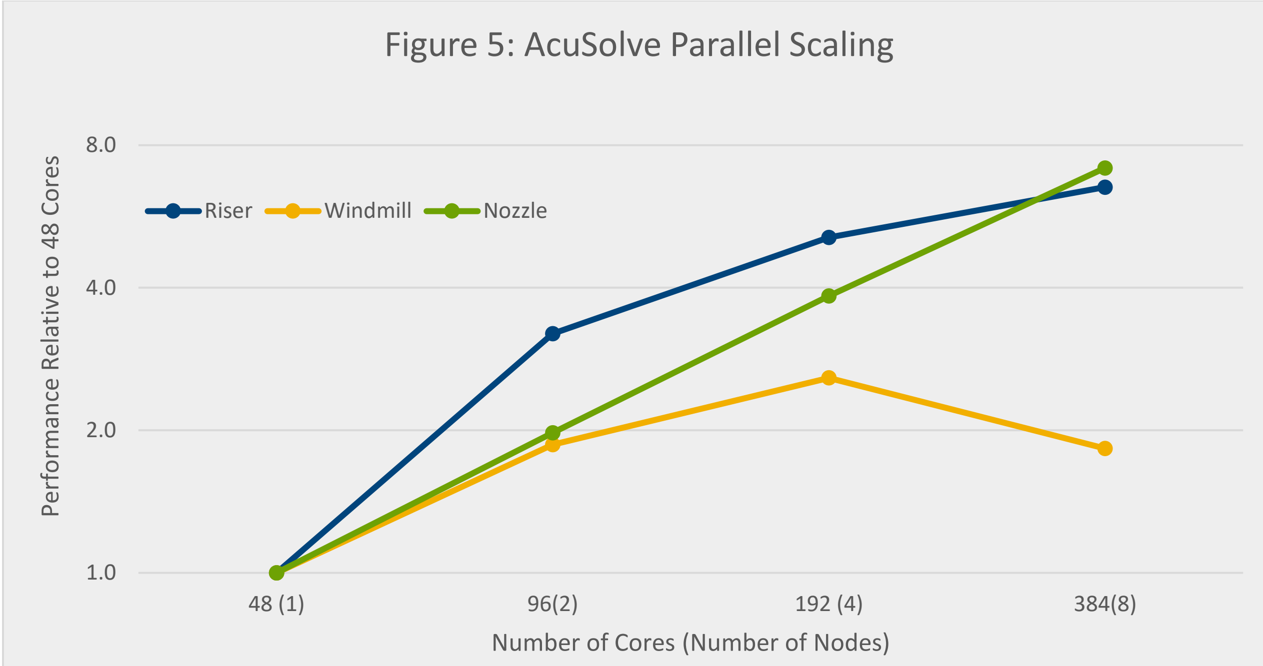
Altair AcuSolve is a Computational Fluid Dynamics (CFD) tool commonly used across a very wide range of CFD and multi-physics applications. AcuSolve is a robust solver with proprietary numerical methods that yield stable simulations and accurate results regardless of the quality and topology of mesh elements. It uses a single solver for all flow regimes, with very few parameters to adjust, and it produces rapid results by solving the fully-coupled pressure/velocity equation system using scientifically proven numerical techniques. 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, such as large transient analysis, may have greater I/O demands.

Figure 4 shows the relative single server performance of the CBB server types for three AcuSolve benchmarks.



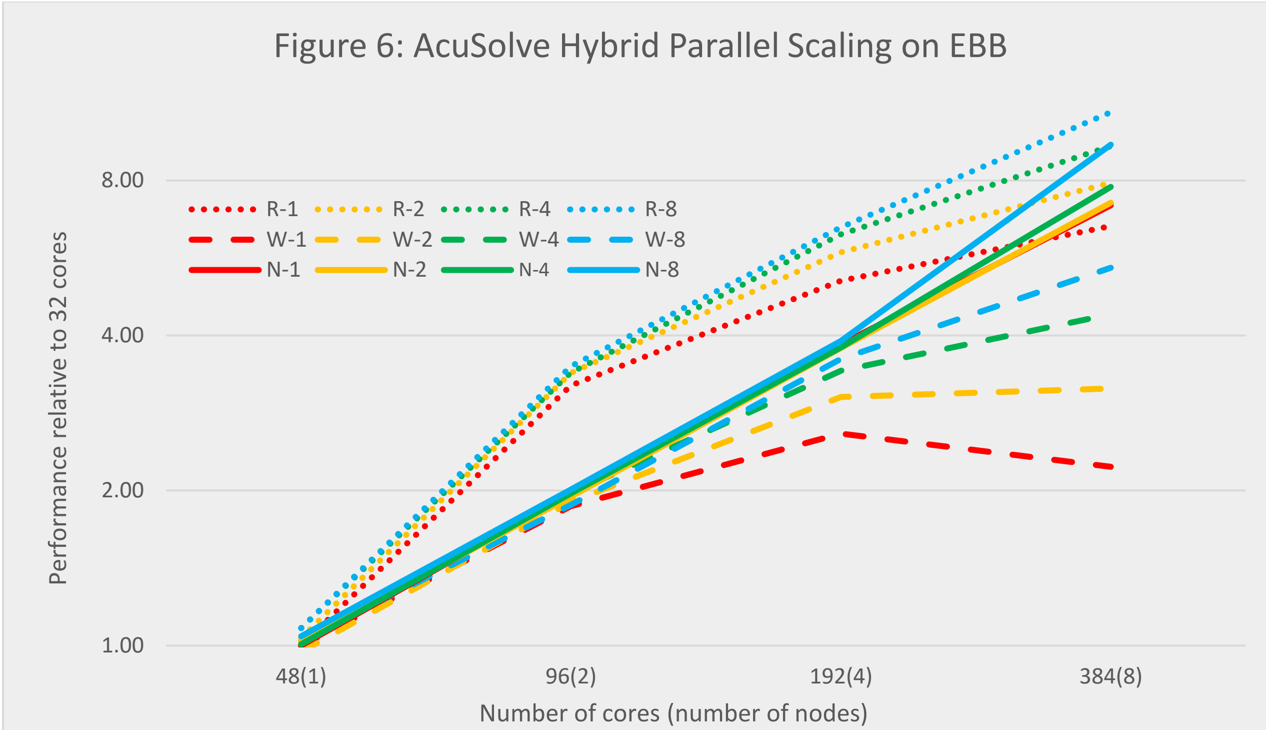
For comparison, the figure also includes performance data for prior generations of the Ready Solution for HPC Digital Manufacturing the 13th generation system using 16-core Intel E5-2697Av4 processors, and the 14th generation system using Intel Xeon Gold 16-core 6142 processors. Overall, the Intel 61xx (Skylake) and 62xx (Cascade Lake) 14G based servers demonstrate significantly better performance than the E5-2697Av4 based 13G system. For all benchmarks, the 16-core 6242 displayed noticeably better performance than its 6142 counterpart and the 20-core 6248 and 24-core 6252 demonstrated even better performance.

Figure 5 shows the parallel performance improvement for AcuSolve when running in parallel across multiple servers.



These benchmarks were carried out on a cluster of eight servers, each with 6252 processors. The results are presented in relative performance compared with the single node results. On the surface these results appear surprising for the “Riser” model where the performance increases by more than a factor of 2X going from one to two nodes. However, this behavior can be explained by “cache effects”, where when the data set is distributed among a greater number of nodes, there can be a point where the entire problem can fit into cache, and the speed of the solver can increase dramatically. Such cache effects are highly problem specific. In general, there is a tradeoff in distributed memory parallelism where the cache performance typically improves as the problem is distributed to more nodes, but the communication overhead also increases, counteracting the increased performance from the caching benefit. Overall the datasets show excellent parallel speedup up to 4 nodes. The largest model “Nozzle” displays nearly linear parallel scaling up to 8 nodes.

AcuSolve is a hybrid parallel application, where it is possible to use both shared memory parallelism within a node and distributed memory parallelism both within a node an across nodes. Finding the proper balance between shared memory and distributed memory parallelism within a node can be daunting. Figure 6 shows the parallel performance for these models where the number of shared memory parallel threads is adjusted from 1 to 8 threads per domain, where the number of domains per server was the divisor of the total number of cores per server with the number of threads per domain.

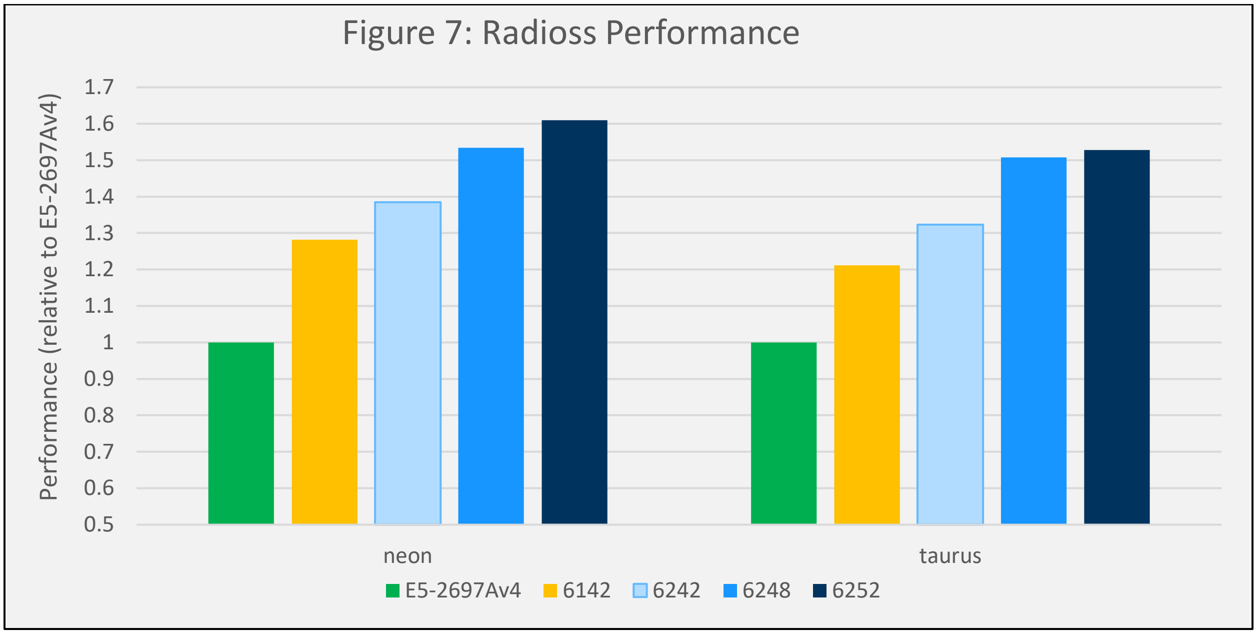


Again, the Riser(R) model shows better overall parallel scaling than the larger Windmill(W) and Nozzle(N) models, primarily from cache effects. All models display similar behavior when the number of shared memory threads is varied. There is little benefit in using multiple threads until four nodes are used. At eight nodes, the benefits of multiple shared memory threads are noticeable, where typically the more threads the better, up to a certain point. It would appear that a good rule of thumb for using thread parallelism would be to use one thread for the number of nodes used in the run (i.e. 4 threads for 4-node runs). This may not be optimal for every situation but should give reasonable performance.

5 Altair Radioss Performance

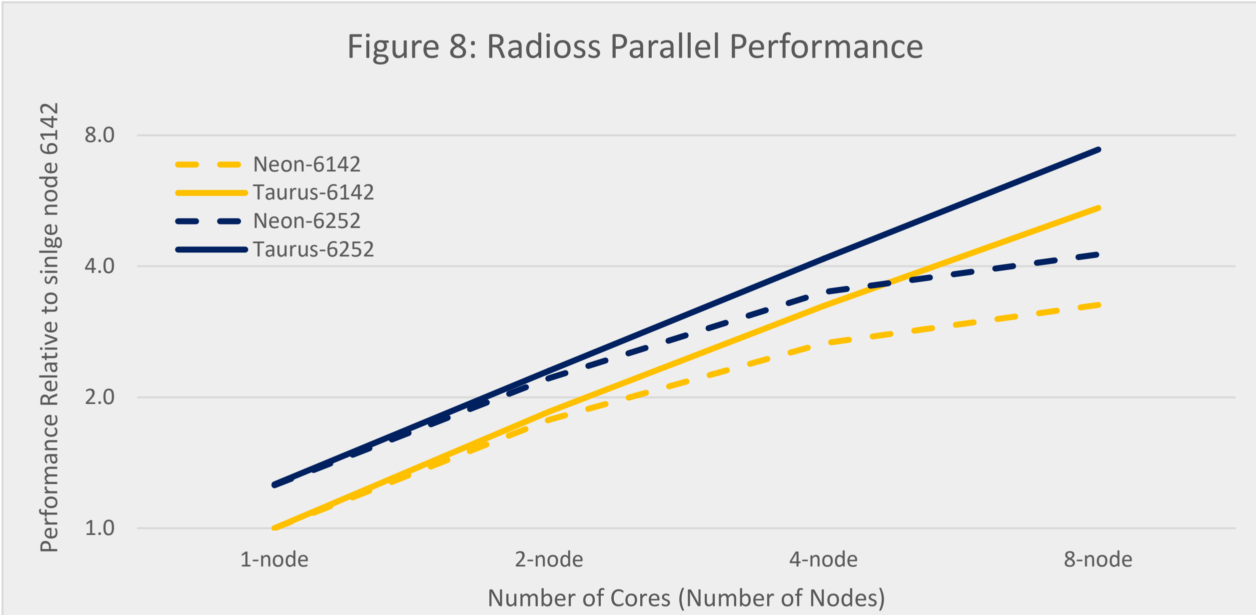
Altair Radioss is a leading structural analysis solver for highly non-linear problems under dynamic loadings. It is used across all industries worldwide to improve the crashworthiness, safety, and manufacturability of structural designs. Radioss is similar to AcuSolve in that it typically scales well across multiple processor cores and servers, has modest memory capacity requirements, and performs minimal disk I/O while in the solver section. It is tightly integrated with OptiStruct and it comes with a comprehensive material and rupture library.

Figure 7 shows the relative performance for two Radioss benchmarks on single servers. For this comparison, all processor cores were utilized while running Radioss.



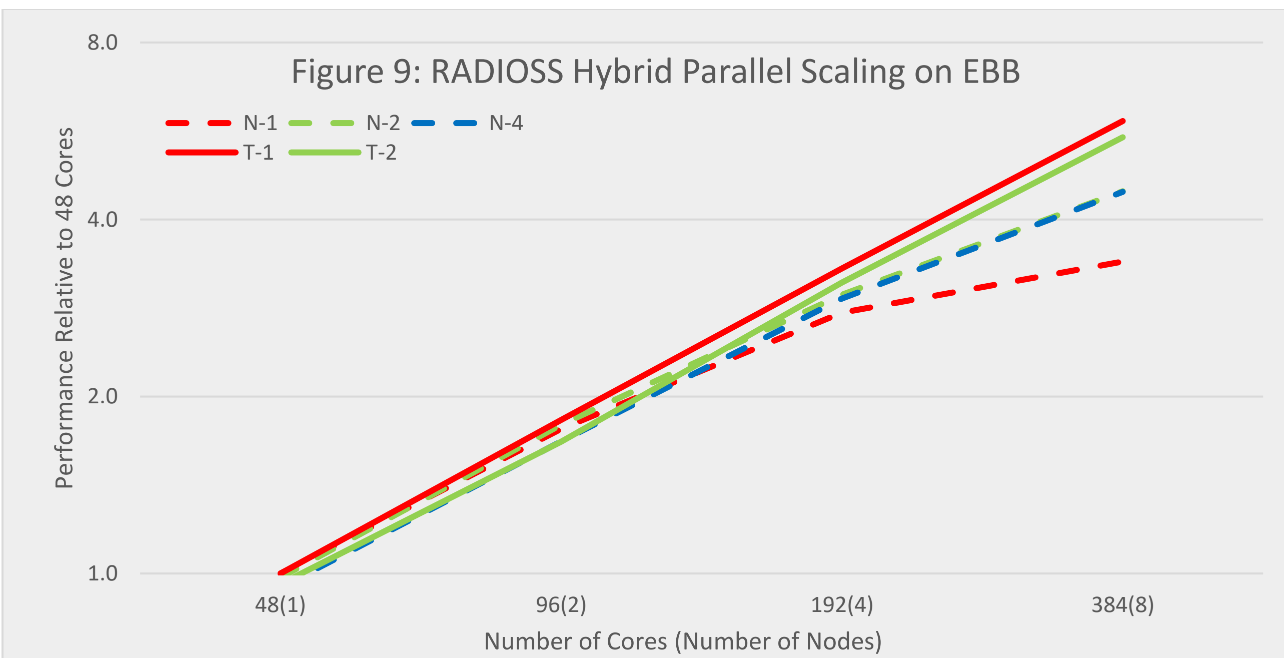
Here the performance is measured based on the elapsed solver time with the reference of 1.0 for the Dell EMC 13G Dell EMC R630 with dual Intel Xeon E5-2697Av4 16-core processors. The results are similar to the AcuSolve results, where the increase in performance from the 13G to the 14G servers is significant, particularly when 62xx series processors with a lot of cores are used.

Figure 8 presents the Radioss parallel performance with parallel benchmarks run across up to 8 servers.



The figure uses the performance reference of 1.0 for a single node equipped with the Intel 6142. The overall parallel scalability for these models is good, behaving as expected, where the larger model shows better overall scalability. The single node performance improvement of the servers equipped with 6252 processors over the servers equipped with the 6142 processors is maintained for these benchmark runs up through eight nodes.

Like AcuSolve, it is also possible to use Radioss in hybrid parallel mode. Figure 9 shows the parallel performance for both Radioss benchmarks models using (1,2,4) shared memory threads on up to 8 compute nodes.



Here, the results are significantly different than the results obtained with the hybrid parallel version of AcuSolve. For the larger Taurus (T) model, the best performance was always obtained using a single shared memory thread (the same as non-hybrid distributed memory MPI). For the smaller Neon model, there was a small benefit using more than one thread at four or more nodes. These results indicate that users should carefully test their models for the potential benefit of using multiple threads when running on larger number of nodes, particularly with smaller models.

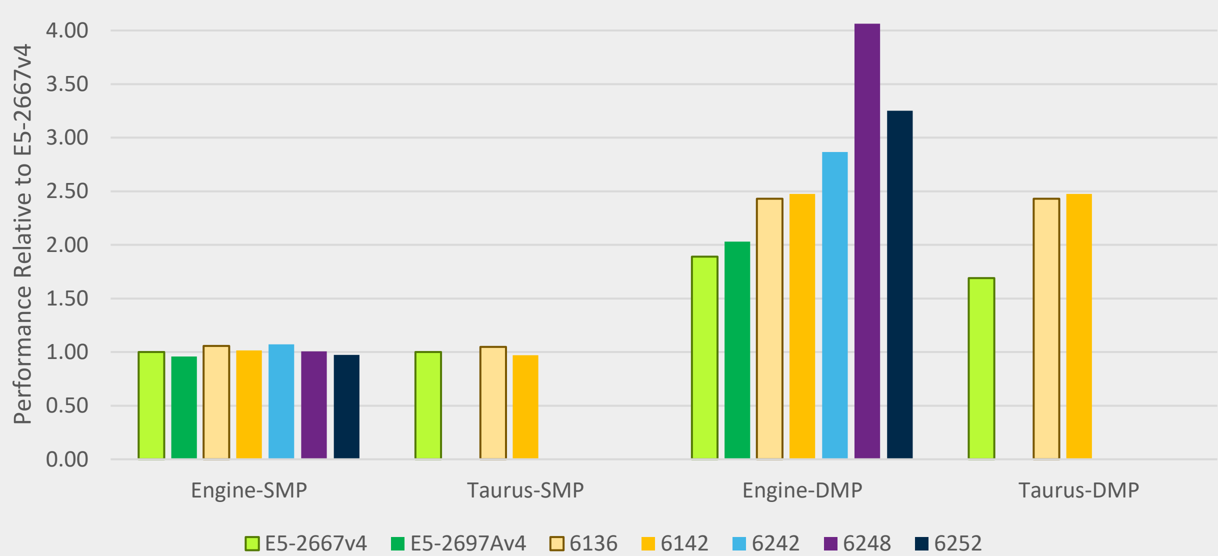
6 Altair OptiStruct Performance

Altair OptiStruct is a multi-physics Finite Element Analysis (FEA) solver commonly used in multiple engineering disciplines. Based on finite-element and multi-body dynamics technology, and through advanced analysis and optimization algorithms, OptiStruct helps designers and engineers rapidly develop innovative, lightweight and structurally efficient designs. It accurately handles nonlinearity of materials, geometries, and contact for applications including gasket analysis, bolt pre-tensioning, rotor dynamics, and thermo-structural analysis.

Two types of solvers are available with OptiStruct: Distributed Memory Parallel (DMP) and Shared Memory Parallel (SMP). In general, the DMP solver offers equivalent or better performance than the SMP solver particularly when all of the cores on a processor are used. Depending on the specific problem types, FEA simulations may or may not scale well across multiple processor cores and servers. Implicit FEA problems often place large demands on the memory and disk I/O sub-systems. Given the varying system requirements for different types of FEA problems, benchmarking for OptiStruct was performed using a variety of server configurations.

For system testing, two benchmark problems provided by Altair (Engine=5MDOF, Taurus=25MDOF_FRF). provided by Altair, are shown. Benchmark results from various systems using both the SMP solver and the DMP solver where possible are show below in Figure 10 below.

Figure 10: Single System Performance—OptiStruct

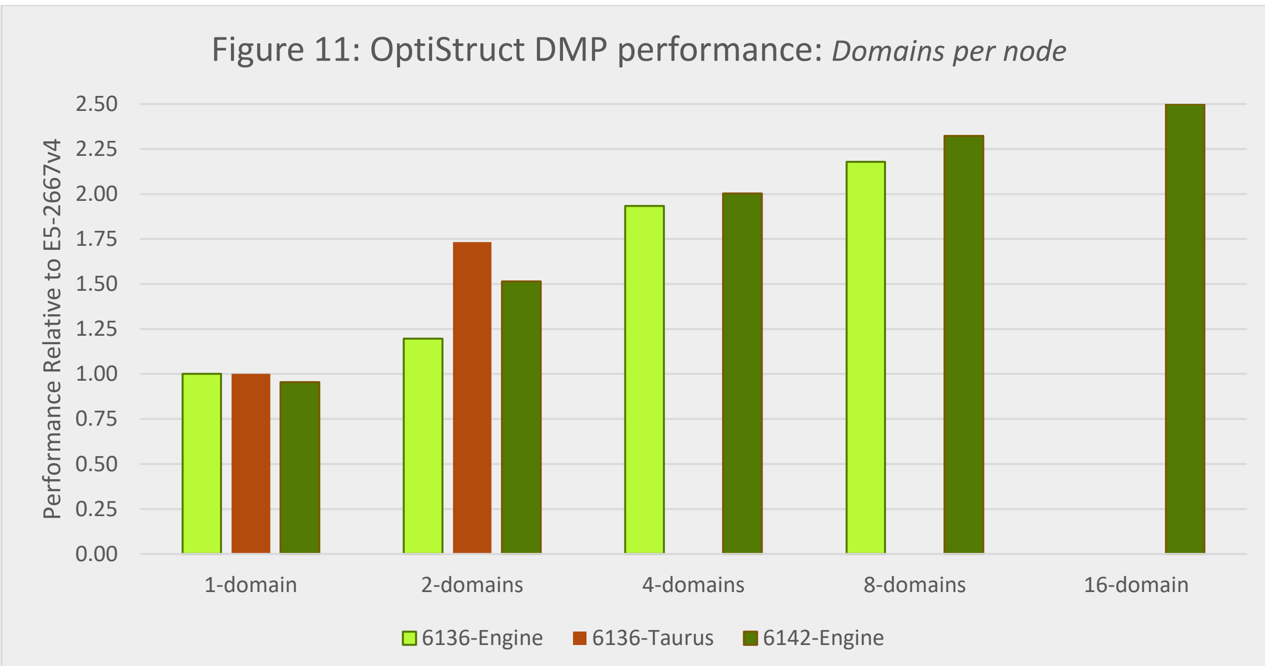


These results are shown relative to the performance of the 13G 8-core Intel Xeon E5-2667v4 based server, where higher indicates better overall performance. The E5-2667v4 based server had 512GB of memory with 2TB of local storage, the E5-2697Av4 server had 128GB of memory and 1TB of local storage, the 6136-based server had 384GB of memory and 2TB of disk space, and the remaining servers had 192GB of memory with 1TB of local storage.

The Engine model was small, making it was possible to carry out benchmarks on all test systems. The Taurus model was much larger and difficult to run on systems with modest memory and storage amounts, resulting in several missing table entries. The results shown for the DMP solver were those obtained using the optimal number of DMP partitions per node. The optimal number of partitions can vary based on the benchmark model and server. Typically, OptiStruct scales better using more DMP partitions per node with fewer SMP threads each as compared with using fewer DMP partitions and more SMP threads per partition. However, typically memory and I/O requirements increase with more DMP partitions, such that it is possible to run out of memory or create I/O bottlenecks with too many DMP partitions per node.

The overall observations show that for the small Engine model, the performance correlated well with the total number of cores on the system, particularly for the DMP solver, which showed significantly better overall performance than the SMP solver. The results from the Taurus model displayed fairly uniform behavior for the SMP results across the various servers, similar to what was observed with the engine model. For system with sufficient memory, the DMP version displayed similar increases in performance as a function of increasing number of cores per server.

The performance of the DMP solver is based on how the simulation is laid out on the server in terms of the number of DMP partitions and the number of SMP threads per partition. For practical considerations it is assumed that there would be no more than one active thread per physical core so the product of the number of DMP partitions on the node with the number of threads per partition was never greater than the total number of physical cores. Figure 11 shows the relative performance of the various domain layouts for servers with both benchmark models.

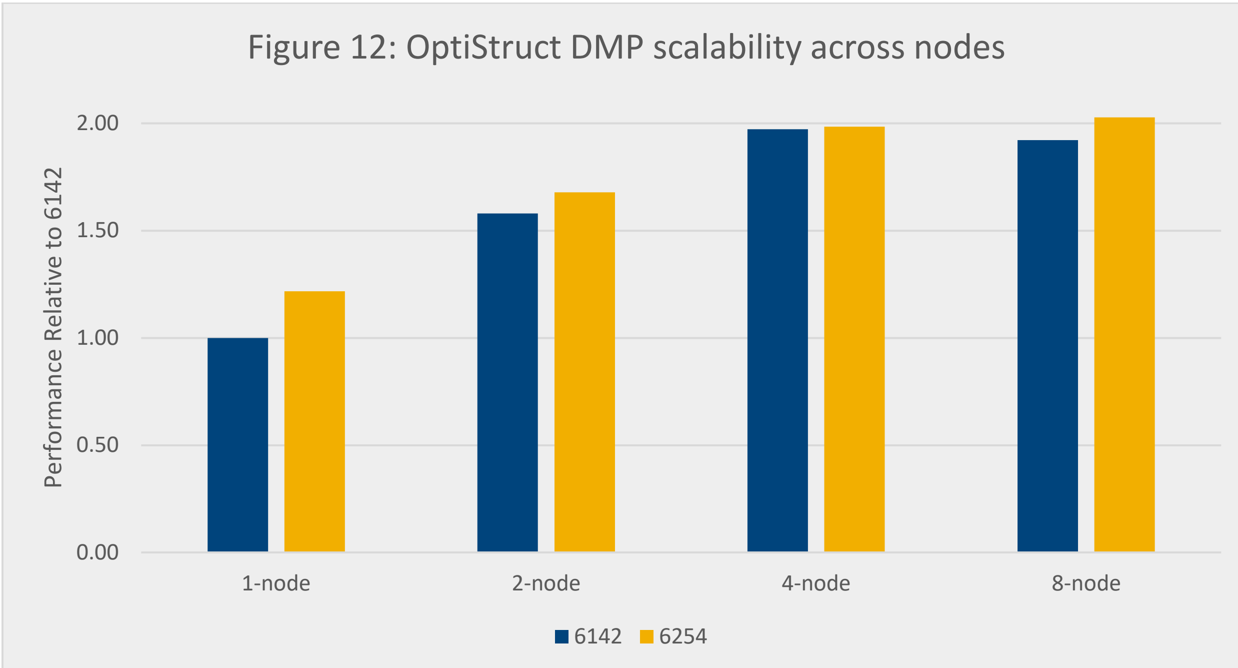


Here, the reference value of 1.0 was chosen to be the single domain for the 6136-based server. For the cases tested, eight domains appeared to be a good value, offering significantly better performance than two or four domains. However, the I/O storage requirements increase with the number of MPI domains such that there can be a practical limit of the number of domains possible for larger models. As an example, there was insufficient Disk space available on the 6142 based server with 1TB of scratch space to use four MPI domains with the Taurus benchmark.

Apart from providing better single-server performance from the SMP solver, the DMP solver allows simulations to be run using multiple servers. Our experience with MPI based CAE packages is that a high-bandwidth, low-latency network is typically required to carry out MPI based simulations across multiple nodes.

There are benchmarks which may perform well without such a high-performance network, but such a solution would not be considered robust for a variety of caseloads.

Figure 12 presents the relative performance for OptiStruct with the Engine model using the DMP solver across multiple servers.



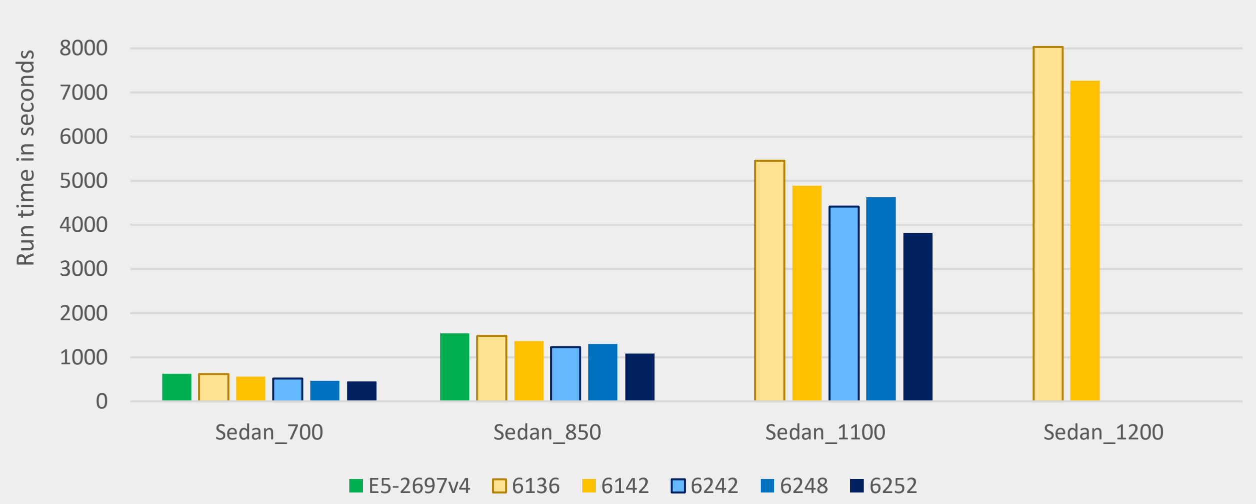
The reference for the plots was the single node 6142 server, where higher is better. One principle advantage of running OptiStruct jobs across nodes is that more memory and disk space is available when using multiple nodes, allowing larger models to be solved than could be solved using a single node. The results indicate good performance scalability when using multiple nodes.

7 Altair Feko Performance

Altair Feko is a comprehensive computational electromagnetics (CEM) solution used widely in the telecommunications, automobile, aerospace and defense industries. Feko offers several frequency and time domain EM solvers under a single license. Hybridization of these methods enables the efficient analysis of a broad spectrum of EM problems, including the analysis of antennas, microstrip circuits, RF components and biomedical systems. Feko also enables computation of antenna placement on electrically large structures, the scattering from large structures as well as electromagnetic compatibility (EMC) investigations.

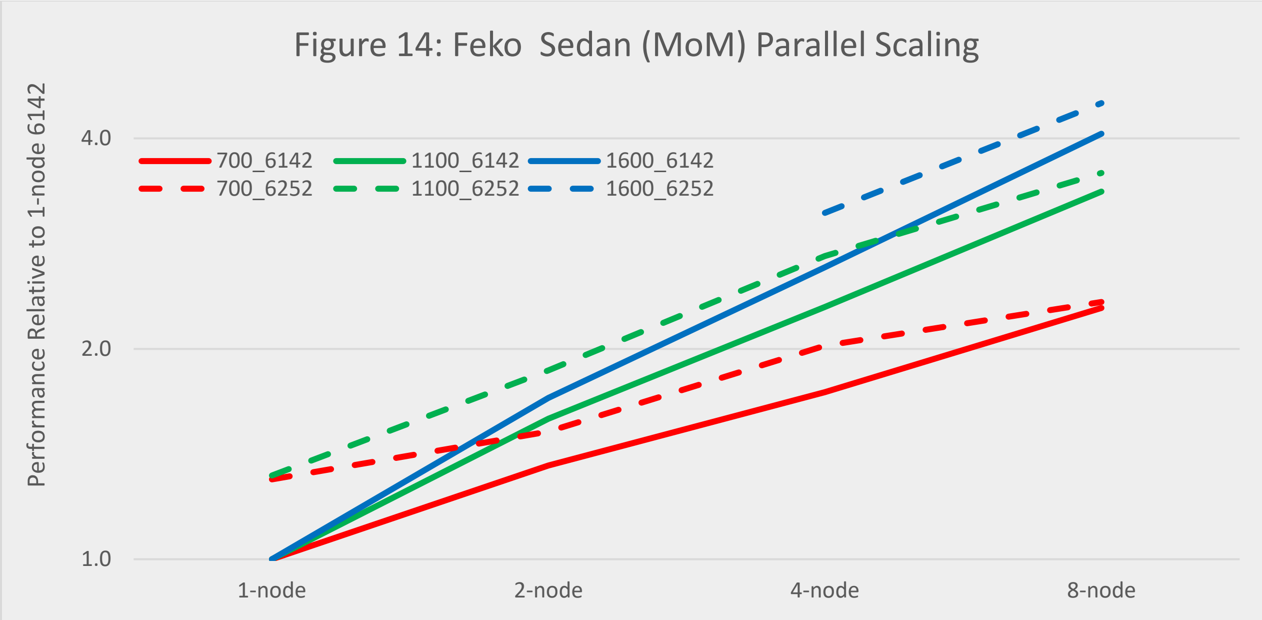
Figure 13 show the performance of a generic sedan model at various frequencies, which use the full wave method of moments (MoM), when run on various solver computer servers.

Figure 13: Single System Performance—Feko Sedan (MoM)



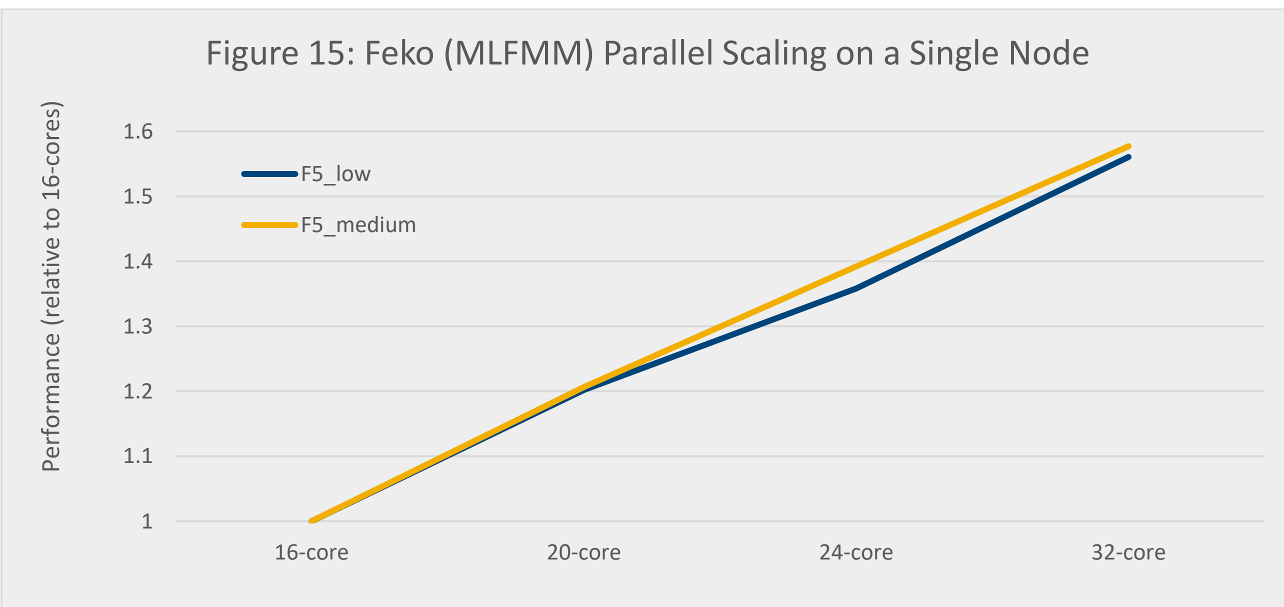
Full wave electromagnetic models require substantial memory to run when using this direct full wave MoM solver. Feko also has very memory efficient methods like UTD or PO or RL-GO (computations on the fly). Only the MoM solver used here requires that much memory. For these benchmarks, the E5-2697Av4 processor-based server had 128GB of memory, the 6136 based server had 384GB of memory, the 6142 server had 768GB of memory, and the 62xxx servers all had 192GB of memory. As expected, the simulation run times increased as a function of input mesh size. There was insufficient memory available on some of the servers to run all of the benchmarks models. The benchmarks were also carried out on a 6142 based server equipped with 192GB of memory and the performance difference between the 129GB and 1768GB servers was similar, indicating the amount of system memory does not have a measurable effect on performance.

Figure 14 shows the performance speedup obtained when carrying out these benchmarks over multiple 14G EBB servers.



All the models show good scaling when running a single job with up to 8 nodes. It should be noted that Feko can efficiently distribute the problem datasets across nodes for larger direct full wave problems. The larger sedan examples solved at 1.6 GHz which could not fit within a single 192GB node. At 1.6GHz, four nodes were required to carry out the job. The performance benefit of the larger cores processors, such as the 6254 diminishes as the problem is run across more nodes, but this effective relative speedup is better maintained for the higher frequency models.

Figure 15 shows the performance of the benchmark model, “F5” at both low and medium range frequencies. This model is much larger than the Sedan model described above, and benchmarks were carried out using the iterative Multi Level Fast Multipole Method (MLFMM). The MLFMM is very efficient at solving very large full wave problems and require a very fast interconnect. The problem size only warranted presenting results computed on a single node. These benchmarks were carried out on a 6142-based server equipped with 768GB of memory, as neither simulation would fit in 192GB of memory.

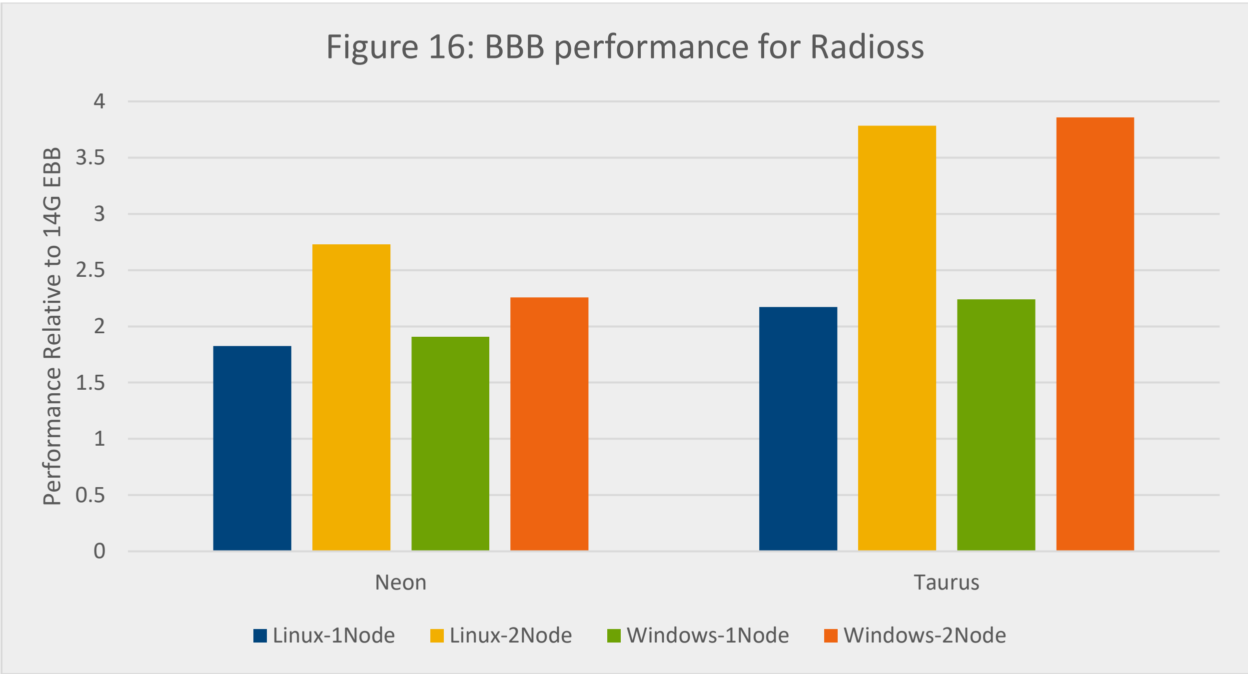


Both simulations showed a significant performance improvement up to all 32 cores on the system.

8 Basic Building Block Performance

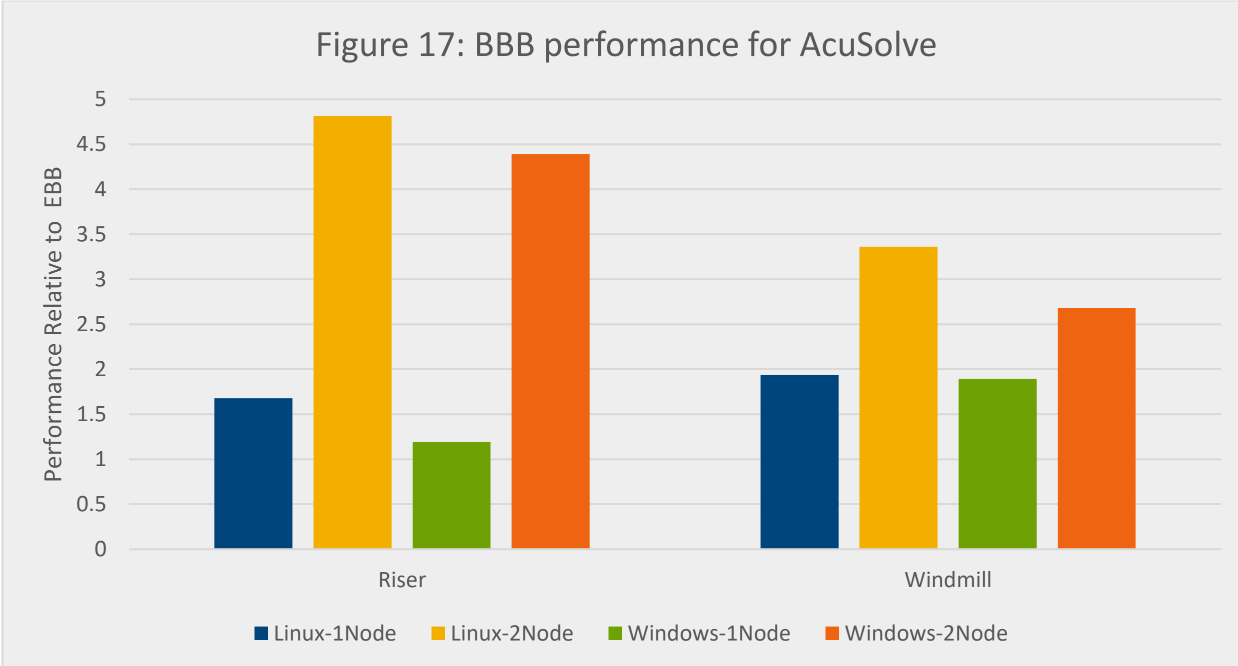
We tested the performance of systems created with Basic Building Blocks (BBB's) for Altair AcuSolve, Radioss, and OptiStruct. We tested performance on both a single BBB and a BBB couplet composed of two BBB's with a direct high-speed network connection. We tested with both RHEL Linux and Windows Server 2016 Enterprise Edition. For Linux, our results were obtained using an EDR network, for Windows, they were obtained using a 25 GbE network, since this is how we envision customers using these systems.

Figure 16 shows the results we obtained for the Radioss benchmarks with the system configurations mentioned above.



The results indicate that the Basic Building Blocks offer comparable performance to the EBB based building blocks within a conventional HPC cluster running Linux using an InfiniBand network, when compared on a per core basis. This appears to hold true for both the Linux and Windows based solutions. The smaller Neon model has some performance limitations when run across the couplet, particularly with Windows, but as noted above this model offers limited overall parallel scalability at higher core counts. A simple solution of BBBs could be created such that modest jobs could be run on a single server and larger jobs run across the couplet.

Figure 17 displays similar performance data for AcuSolve on the BBB.



Here, the two node results for the Riser benchmark are artificially high, due to the effect of being able to fit this model into cache on two BBB's, but not into the reference 14G EBB server. This effect can be seen as well for the parallel speedup of this model on a system of EBB in Figure 4. Like Radioss, the BBB shows excellent overall performance both with Windows and with Linux. Again, the Linux based couplet offers a noticeable performance benefit over its Windows counterpart.

At the writing of this paper, we have yet to complete OptiStruct benchmarks with the BBBs. Overall, we believe that OptiStruct workloads will be suitable for BBBs, but not as optimal as they would be as the IBBs, which have a better balance of CPU, memory, and I/O for a typical OptiStruct workload than the BBB. However, customers wishing to use the BBB for a diverse CAE workload could use BBBs for such a purpose.

9 Conclusion

This technical white paper presents the Dell EMC Ready Solution for HPC Digital Manufacturing. 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 system performance with Altair HyperWorks software.