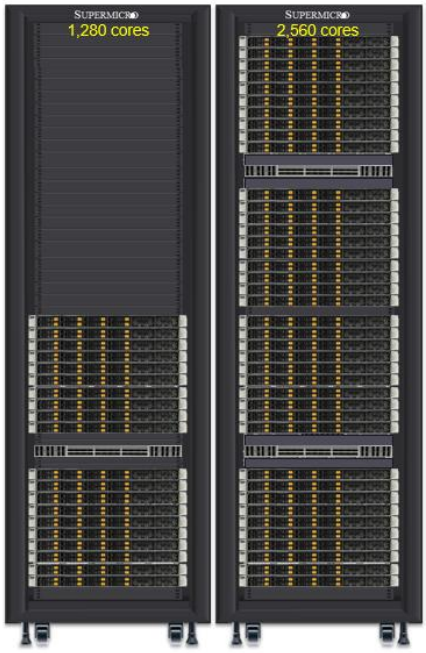
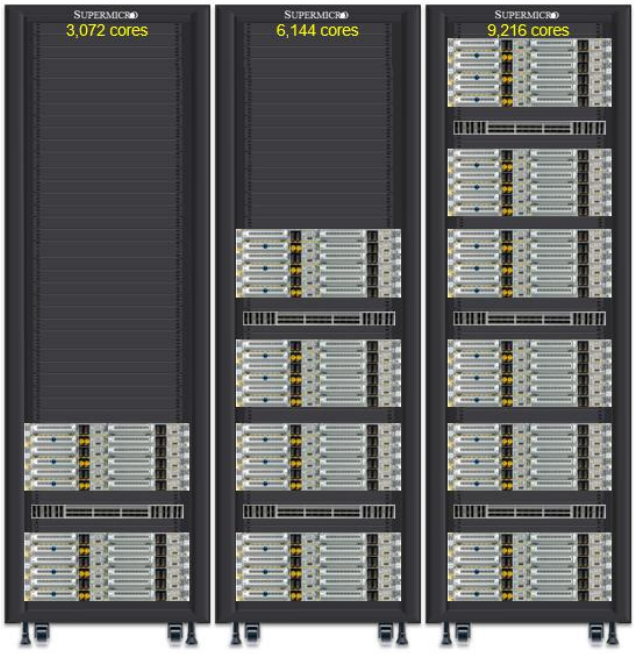
SOLUTION BRIEF

SUPERMICRO SCALABLE HIGH-PERFORMANCE COMPUTING CLUSTERS POWERED BY CORNELIS

Supermicro Hyper and FlexTwin Systems with Cornelis CN5000 Omni-Path



Air-cooled



Liquid-cooled

TABLE OF CONTENTS

Executive Summary

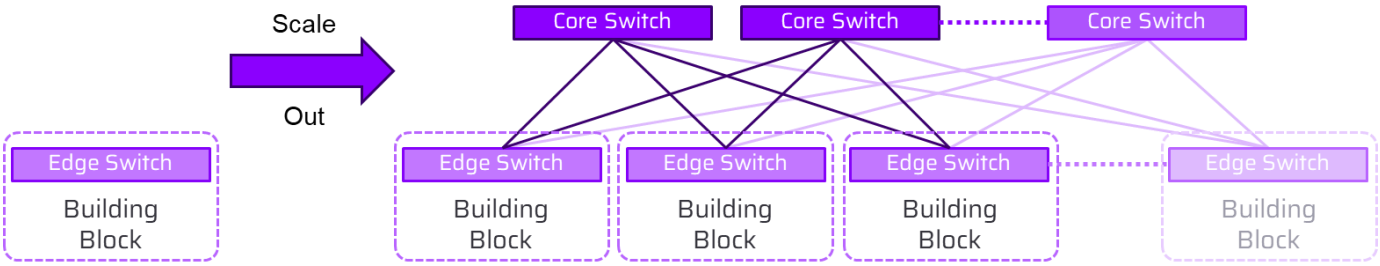
A densely populated rack of advanced server CPUs is a cornerstone of modern high-performance computing (HPC) infrastructure. With high-clock frequency cores, high-bandwidth memory, and support for advanced instruction sets such as AVX-512, the cluster delivers fast, complex simulations, genomic sequencing, and fluid dynamics modeling. Adding liquid cooling supports sustained peak performance, up to 4X compute density, while maximizing throughput per watt. Liquid cooling can remove up to 95%[1] of the heat generated by systems and up to 100%[1] of the heat generated by switches. This architecture enables organizations to tackle grand-challenge

problems faster and at lower operational costs, providing superior price-performance and more compute for a given data center power budget. The networking fabric is critical to the cluster's scale and efficiency.

Cornelis CN5000 Omni-Path unlocks the full potential of the rack by eliminating network bottlenecks. CN5000 delivers up to 400 Gbps of bandwidth per port with ultra-low latency (<1 microsecond for MPI operations), the industry's highest message rate (800 million packets per second), lossless packet delivery, and advanced congestion-avoidance architecture. The result is dramatically improved cluster efficiency, reduced time-to-insight for research breakthroughs, and a compelling return on investment through higher overall system productivity.

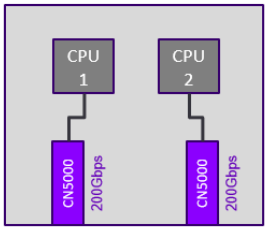
Supermicro and Cornelis have established rack-scale building blocks for building high-performance, low-latency CPU clusters, available in air-cooled or liquid-cooled configurations. Multiple units of these building blocks can be used to scale out compute capabilities, with low-latency interconnect across all CPUs, providing scalability in high-performance computing deployments.

Reference Architecture Scalability



The Reference Architecture allows multiple Building Blocks to be deployed to scale out the compute cluster's capacity and performance. Starting from a single building block with a single CN5000 48-port Edge switch, the fabric can be expanded by adding a Core tier of CN5000 48-port switches to create a fat-tree topology, enabling the seamless integration of additional Edge switches and, therefore, Building Blocks.

Air-Cooled Latency Optimized Compute Building Block AS -1126HS-TNR-01-G2



Balanced Connectivity
- HFI per CPU reduces average latency
- Double peak message injection rate

Consumes same number of switch ports as 1x 400gbps connection

This architecture provides dual 200Gbps CN5000 Omni-Path connections per node. It prioritizes fabric connectivity to reduce average latency per node and double the peak message rate, while still delivering 400 Gbps per node and using the same number of switch ports as a single 400 Gbps design.

- 20 nodes, 23U, 33 KW (estimated)
- 1,280 cores / 3.55GHz (5th Gen AMD EPYC™ 9355 CPUs)
- 30TB (6400MT/s) total memory
- 37.5TB Total Flash Storage
- Cornelis CN5000 Omni-Path 48-port edge switch
- 2x 200Gbps Dual rail, non-blocking, low-latency cluster network
- 25GbE connectivity to outside of cluster
-

Table 1 Scaling Air-Cooled Compute Building Blocks

Air cooled	1 building block (min)	2 building blocks	N building blocks	48 building blocks (max)
Compute	1,280 cores, 20 nodes	2,560 cores, 40 nodes	N*1280, N*20 nodes	61,440 cores, 960 nodes
Network bandwidth	2x 200Gbps	2x 200Gbps	2x 200Gbps	2x 200Gbps
U height	23U	46U	N*23U + 4N + N	~1130U
# Cornelis 48p Edge switches	1	2	N	48
# Cornelis 48p Core Switches	-	- (back-to-back)	N/2	24

Liquid Cooled High-Density Compute Building Block
AS -2126FT-HE-ALC

This architecture provides a single 400Gbps CN5000 Omni-Path connection per node and prioritizes compute density to deliver the highest FLOPs per rack

- 24 nodes, 15U, 20 KW (estimated)
- 3,072 cores / 3.3GHz (5th Gen AMD EPYC™ 9575F CPUs)
- 36.9TB (6400MT/s) total memory
- 23TB Total Flash Storage
- Cornelis CN5000 Omni-Path 48-port edge switch
- 400G single-rail, non-blocking, low-latency cluster network
- 25GbE connectivity to outside of cluster

Table 2 Scaling Liquid-Cooled Compute Building Blocks

Liquid cooled	1 building block (min)	2 building blocks	N building blocks	48 building blocks (max)
Compute	3,072 cores, 24 nodes	6,144 cores, 48 nodes	N*3,072, n*24 nodes	147,456 cores, 1152 nodes
Network bandwidth	400Gbps	400Gbps	400Gbps	400Gbps
U height	15U	30U	N*15U + N/2 + N	~750U
# Cornelis 48p Edge switches	1	2	2	48
# Cornelis 48p Core Switches	-	- (back-to-back)	N/2	24

Conclusion

Supermicro systems and Cornelis have set up rack-scalable HPC building blocks that include low-latency interconnect switches and infrastructure, along with high-performance CPU and memory, to help customers quickly deploy solutions for complex engineering, genomic, and scientific problems. By adding building blocks, more computing processors are added to the network, while retaining low-latency connectivity across all processors. Customers can scale the cluster to the footprint they need to meet their computing requirements, while maintaining a very low total cost of ownership.

For More Information

To get more information, visit:
https://www.supermicro.com
https://www.cornelis.com

Footnote:

1 – Based on Supermicro simulations and measurements

SUPERMICRO

As a global leader in high performance, high efficiency server technology and innovation, we develop and provide end-to-end green computing solutions to the data center, cloud computing, enterprise IT, big data, HPC, and embedded markets. Our Building Block Solutions® approach allows us to provide a broad range of SKUs, and enables us to build and deliver application-optimized solutions based upon your requirements. Visit www.supermicro.com

CORNELIS

Cornelis is a technology leader delivering purpose-built, high-performance fabrics accelerating high-performance computing, high-performance data analytics, and artificial intelligence workloads in the cloud and in the datacenter. The company's products enable scientific, academic, governmental, and commercial customers to solve some of the world's toughest challenges by efficiently focusing the computational power of many processing devices at scale on a single problem, simultaneously improving both result accuracy and time-to-solution for their most complex application workloads. Visit www.cornelis.com