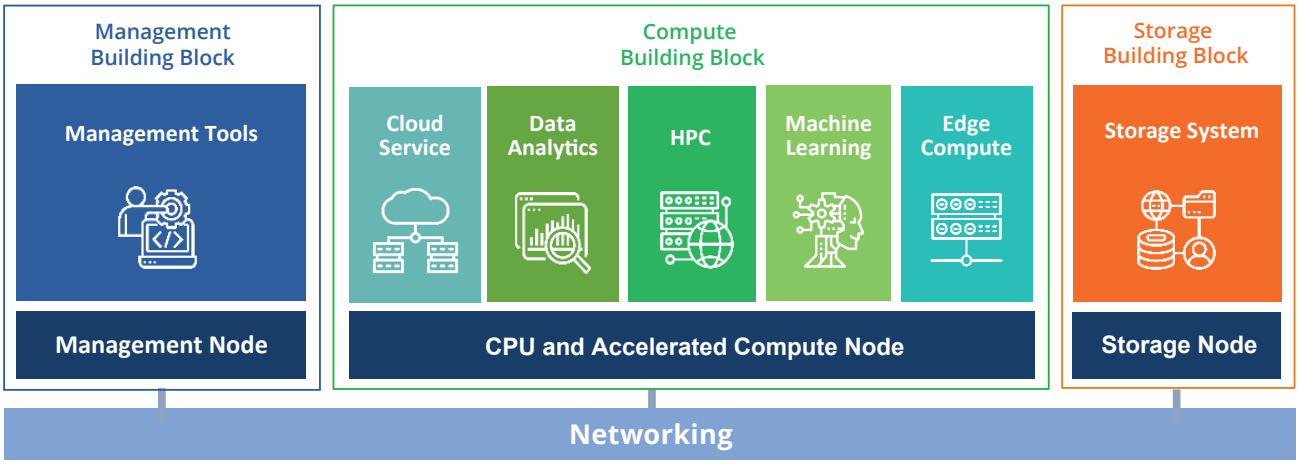
Adaptive HPC & AI Converged System

QCT® Platform on Demand (QCT® POD)

Flexible Workload-driven Design to Power Your Business

High-performance computing (HPC) is essential for scientific research modeling and artificial intelligence (AI) supports data analysis for informed decision-making. As data volumes surge in HPC and AI workloads, enterprises must adapt their infrastructure to meet processing demands. The need for efficiency and specialization in handling diverse tasks prompts the adoption of heterogeneous environments, combining various hardware architectures. However, building an integrated system to accommodate this heterogeneity is time-intensive and resource-demanding.

As a premier data center solutions provider, QCT® utilities our expertise to deliver a modern and adaptive infrastructure that meets the evolving needs of enterprises in HPC and AI. QCT® Platform on Demand (QCT® POD) is an integrated solution supporting both bare-metal and cloud-native environments for HPC and AI workloads. Featuring pre-validated and pre-configured system designs to meet diverse workload requirements, QCT® POD incorporates comprehensive tools and mechanisms to streamline cluster deployment and management processes.



QCT® POD Meets Your Demands with Workload-Driven HPC & AI Converged Solution

• **HPC & AI Convergence**	Supports both HPC and AI environments and offers toolkits to simplify the system management and software development
• **Workload-Driven**	Adaptive infrastructure tailored to meet specific workload requirements ensure cost optimization and enhanced resource utilization
• **Infrastructure Capabilities**	Comprehensive hardware portfolio with seamlessly integrated high-quality software and hardware systems

Meeting Industrial Workload Demands

The QCT® POD is a flexible, scalable solution tailored to diverse workload demands. Its adaptability allows adaptive building block designs based on specific industry requirements, aiding customers in overcoming challenges associated with limited IT resources.

Developers & End users

Pre-built workload packages and a variety of pre-installed development tools make it easier to streamline AI workflows and improve HPC workload performance.

- **Workload Package**

  QCT® offers diverse fine-tuned HPC workloads to accelerate model simulation.

  - Computation Fluid Dynamics
  - Molecular Dynamics
  - Numeric Weather Prediction
  - Quantum Chemistry

- **Development Environment**

  Customers can seamlessly access cloud-native and bare-metal development environments through various workspaces and utilize automated resource management tools.

- **Development Toolkit**

  Turnkey solution with a rapid deployment tool to accelerate your time-to-value Pre-configured and pre-validated HPC & AI runtime environment with relevant libraries, frameworks, and compiler tools.

User Workspaces

QCT® POD provides **multiple development environments** for users, offering comprehensive development tools accessible through **web-based and text-based interfaces.**

Resource Management

QCT® integrates an **advanced scheduler**, including a customized CLI tool (Qbatch), to enable users to **run multiple batched jobs in a cloud-native environment more easily and intuitively.** It also supports SLURM for **ensuring optimal resource utilization and offering flexible job scheduling capabilities** in a bare-metal environment.

Administrators

QCT® POD offers rapid deployment and powerful management tools to simplify administrative tasks for IT staff, accelerating time to value.

- **System Deployment**

  Empowers users to expedite the system deployment process, reducing setup time for a new cluster from days to hours through a text-based UI configurator.

- **System Monitoring**

  Delivers a web-based dashboard for real-time monitoring of system resource utilization and providing system status reports.

- **System Management**

  Encompasses user authentication, resource management, container orchestration tools, and data management, allowing administrators to effortlessly control user access and resource allocation.

Found at: www.QCT.io/wheretobuy

Powered by 5th/4th Gen Intel® Xeon® Scalable Processors.

Recommended Configuration for QCT® POD Solution

Building Blocks		AI				HPC				HPC & AI converged			
		Standard		Premium		Standard		Premium		Standard		Premium	
		Qty	Specification	Qty	Specification	Qty	Specification	Qty	Specification	Qty	Specification	Qty	Specification
Management nodes	Login node	1	D54Q-2U with accelerator	1	D54Q-2U with accelerator	1	D54Q-2U with accelerator	1	D54Q-2U with accelerator	1	D54Q-2U with accelerator	1	D54Q-2U with accelerator
	K8S master node			1	T43Z-2U					1	T43Z-2U	1	T43Z-2U
	Admin & deployment node			1	D54X-1U			1	D54X-1U			1	D54X-1U
Compute nodes	AI GPU node			1	D74H-7U								
	HPC & AI GPU node	1	D54U-3U							1	D54U-3U	2	D54U-3U
	HPC CPU node					4	D54X-1U	8	D54X-1U	4	D54X-1U	8	D54X-1U
Storage nodes	Capacity	1	S54S-1U	2	S24P-5U Dual nodes					3	S54S-1U		
	Performance			3	D54Q-2U All flashes			2	D54X-1U (MDS) D54Q-2U (OSS) JB4603-4U (OST)			2	D54X-1U (MDS) D54Q-2U (OSS) JB4603-4U (OST)
Network fabrics	Out-of-band management	1	T1048-LB9M	1	T1048-LB9M	1	T1048-LB9M	1	T1048-LB9M	1	T1048-LB9M	1	T1048-LB9M
	Service management	1	T4048-IX8D	1	T4048-IX8D	1	T4048-IX8D	1	T4048-IX8D	1	T4048-IX8D	1	T4048-IX8D
	Compute storage network			1	Mellanox NDR QM9700			1	Mellanox NDR QM9700			1	Mellanox NDR QM9700

Modern Hardware Platforms for QCT® POD

Compute · Storage · Management · Networking

QuantaGrid D54X-1U

- Powered by 5th/4th Gen Intel® Xeon® Scalable processors
- PCIe 5.0 & DDR5 platform ready
- Offers (16) E1.S NVMe flash drives or (12) 2.5" NVMe flash drives

QuantaGrid D54Q-2U

- Powered by 5th/4th Gen Intel® Xeon® Scalable processors
- PCIe 5.0 & DDR5 platform ready
- Support All (24) NVMe U.2 or E1.S drives as hot-tier storage
- Up to (2) dual-width accelerators

QuantaGrid D54U-3U

- Powered by 5th/4th Gen Intel® Xeon® Scalable processors
- PCIe 5.0 & DDR5 platform ready
- Up to (4) DW accelerators or (8) SW accelerators
- Up to (10) SATA/SAS/NVMe 2.5" hot-plug SSDs to speed up data-loading

QuantaGrid D74H-7U

- Powered by 5th/4th Gen Intel® Xeon® Scalable processors
- (8) Hopper H100 SXM5 GPU modules with an HGX baseboard
- (18) SFF All-NVMe drive bays for GPUDirect storage and boot drive

QuantaPlex T43Z -2U

- Powered by 3rd Gen Intel® Xeon® Scalable processors
- 2U4N High density server optimized for space efficiency
- Featuring All-NVMe with high memory footprint and additional expansibility

QuantaVault JB4603-4U

- JBOD, Just a Bunch Of Disks
- Up to (60) HDDs/SSDs in a 4U chassis
- Support up to (4) zoning configurations
- Support multipath for high-availability (HA)

QuantaPlex S24P-5U

- Support up to 1.8 Petabytes with (84) PMR HDDs in a 5U form factor
- Flexible option for single node/dual node MB to meet different workload needs
- Optimized topology with dual SAS cards to achieve enhanced performance

QuantaGrid S54S-1U

- Powered by 5th/4th Gen Intel® Xeon® Scalable processors
- High-density storage server with up to (12) HDDs and (4) 7mm NVMe SSDs in a 1U chassis

QuantaMesh T1048-LB9M

- 48-port 10/100/1000BASE-T and 4-port 10/25GbE SFP28
- Simplified Management
- Software Defined Network (SDN)

QuantaMesh T1048-LY4R

- 48-port 10/100/1000BASE-T and 4-port 1/10GbE SFP+
- ONIE Pre-load
- x86 CPU Design
- Optical OOB Port Support

QuantaMesh T4048-IX8D

- 48-port 10/25GbE SFP28 and 8-port 40/100GbE QSFP28
- x86 CPU Support
- VXLAN
- Multi-Chassis Link Aggregation (MLAG)
- OSPF, BGP4 with ECMP
- Network Automation

About QCT®

Quanta Cloud Technology (QCT®) is a global data center solution provider. We combine the efficiency of hyperscale hardware with infrastructure software from a diversity of industry leaders to solve next-generation data center design and operation challenges. QCT® serves cloud service providers, telecoms and enterprises running public, hybrid and private clouds.

Product lines include hyperconverged and software-defined data center solutions as well as servers, storage, switches and integrated racks with a diverse ecosystem of hardware component and software partners. QCT® designs, manufactures, integrates and services cutting-edge offerings via its own global network. The parent of QCT® is Quanta Computer, Inc., a Fortune Global 500 corporation.

For more information, visit the QCT® website at www.QCT.io.

QCT® Authorized Partner

QCT® POD

www.QCT.io

Contact Us

ver Jun., 2024