SUCCESS CASE

Speed up Time-to-research on QCT POD HPC & AI Workload-Optimization Ready Solution

Industry

- Medical
- Higher education and research

Location

Taiwan

QCT Solution

QCT POD

Boost Your Digital Transformation Journey with QCT

Nowadays, the implementation of HPC and AI workloads is increasing in higher education & research, and healthcare verticals. Users and organizations are often required to run complicated models and simulations to reach insights in a shorter time frame.

The successful cases will demonstrate how Quanta Cloud Technology (QCT) offers an optimized infrastructure for HPC & AI workload applications with a streamlined HPC & AI development environment that features simplified cluster management and workload package.

QCT Platform on Demand—A Workload Driven Solution

QCT is a leading cloud data center solution provider with extensive experience in developing HPC and AI solutions. Leveraging years of experience and expertise in HPC & AI, QCT developed its Platform on Demand (QCT POD), a workload-driven solution with common building blocks that can be used to design on-premise infrastructure with hardware and software integration, allowing HPC and AI workloads to be run under one system architecture to meet easy management, rapid deployment and shortening the time to value.

QCT Accelerates ASIAA HPC System Implementation Journey

Since its development, QCT POD has been implemented in global institutions to boost their HPC/AI workload performance. Academia Sinica Institute of Astronomy and Astrophysics (ASIAA), a leading institution of astronomy research in Taiwan investigating research topics ranging from the solar system, stars, planet formation and cosmology, is one of our customers.

Challenges and Requirements

While ASIAA kept evolving and devoted itself to more complex planetary science projects, it became increasingly dependent on sophisticated study methods. As an enormous volume of data was being collected, the number of simulation jobs also underwent exponential growth. Managing and fully utilizing the data became a critical mission for IT staff and researchers. Thus, ASIAA was in need of a more powerful HPC system that could accelerate the speed of data analysis.

However, HPC system implementation is a time-consuming journey, starting from designing the complex infrastructure, deploying the environment, all the way to managing the system. It requires lots of effort from the data center administrator and is likely to take away the time that could be used to solve many significant issues. Therefore, ASIAA needed assistance in simplifying their implementation journey to generate more data center resources in time to enable more valuable services and speed up research.

> "QCT supported ASIAA in developing a unified solution for High-Performance Computing for astronomy and astrophysics"
>
> - ***Min-Kai Lin, Associate Researcher, ASIAA***

Main Challenges

- To fully utilize and manage the growing data on simulation jobs.
- To build up a converged environment to run both HPC & AI workloads

QCT Value

- Adaptive HPC & AI infrastructure enhances compute power with advanced technology
- Provides best-fit hardware software configuration to boost workload performance and accelerate time-to-insights

QCT POD Delivers Optimal Performance

QCT implemented QCT POD in ASIAA to help the latter achieve their performance goals. By using their Workload Optimization methodology, QCT analyzed the behavior of key workloads at ASIAA, such as Athena++, FARGO3D BJQ, and PLUTO. QCT then designed the best-fit hardware architecture integrated with optimized software configuration to boost workload performance and accelerate time-to-insights.

From the hardware aspect, the QCT POD implemented at ASIAA was powered by the QuantaGrid D43K-1U, a high compute capacity server powered by two 3rd Gen AMD EPYC processors. It also features 512MB L3 cache and 160 PCIe Gen4 lanes to deliver an incomparable performance boost for workloads that require very low latency. The system design also helped ASIAA meet their budget and energy window.

Figure 1. QuantaGrid D43K-1U

QCT POD Helps Reduce TCO

QCT POD features high serviceability to minimize deployment and system maintenance costs. With automated deployment tools provided by their SI partners, QCT POD reduces the complexity of the deployment process and shortens the deployment time from “months” to “weeks.” QCT also helped to build a customized system monitoring dashboard with Grafana to help ASIAA effectively manage their HPC cluster. As a result, these tools have helped ASIAA to greatly reduce its TCO.

QCT Designs HPC & AI Converged System for NCKU

National Cheng Kung University (NCKU) is a leading higher education institution in Taiwan that has a comprehensive list of academic departments and programs ranging from Engineering, Management, Medicine, Biosciences, and Biotechnology, to the Academy of Innovative Semiconductor and Sustainable Manufacturing and more. In addition to educational institutions, NCKU also established its own comprehensive medical center and AI Research Center to expand its span of research as well as to support all faculty members and students in research, teaching, and training.

Challenges and Requirements

In order to facilitate the research and development of AI applications, NCKU turned to QCT for building an optimized digital environment for HPC and AI, aiming to accelerate the research process in its healthcare center, campus, education system, and other research fields.

Building up a comprehensive HPC & AI system for a diverse number of fields is time-consuming and requires massive effort from IT teams. They need to spend a considerable amount of time to familiarize themselves with the system environment, which may occupy their time that could be used to carry out system administration. Moreover, managing resources in the converged HPC & AI systems is also complex. In turn, these challenges hindered the digital transformation journey of many organizations.

QCT POD Solution Supports NCKU to Realize Digital Transformation

To tackle these challenges that occur during one’s transformation journey, QCT has developed an adaptive HPC & AI infrastructure called QCT POD, which enhances compute power with advanced technology. QCT POD provides best-practice hardware and software integration for computational research. The system design of QCT POD

> "QCT provides a holistic HPC & AI converged solution that not only greatly benefits our research, but also paves the way for our digital transformation."
>
> ***-Chair Prof. Jason Yi-Bing Lin***
> ***Director, Quanta-NCKU Joint AI Research Center***

shortens the deployment journey and simplifies cluster management and monitoring. Furthermore, it also features a comprehensive development environment for users.

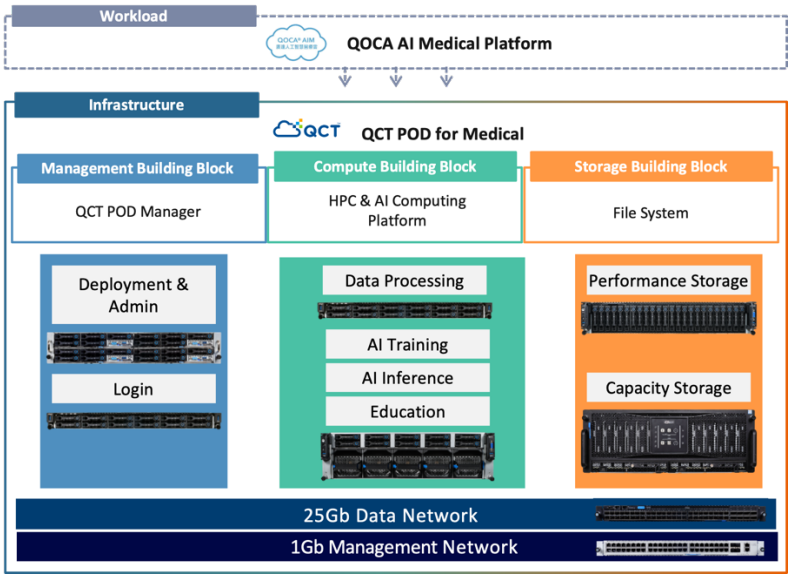

Figure 2. NCKU QCT POD architecture design

Improved Productivity and Simplified Management Journey

To reach better time-to-value, QCT POD provides a Jupyterhub web-based dashboard, which integrates tools, frameworks, and applications required during the development process to allow users to start their jobs easily. It also supports multiple users to simultaneously develop, test, and execute containerized HPC & AI applications with automated resource management and job scheduling.

To solve administration challenges, QCT POD also offers the Grafana admin dashboard, which enables monitoring system resources and statuses in real time. It also consolidates user account and group management functions to allow system administrators to take full control of account authentication. Furthermore, QCT POD also delivers rapid deployment tools for IT teams that shorten system implementation.

Figure 3. Monitoring dashboard

In terms of hardware infrastructure, QCT POD has a pre-defined hardware portfolio that offers a best-fit combination of servers to optimize execution and performance. For example, the QuantaGrid D43K-1U server enables accelerated computing performance with dual AMD EPYC™ processors, up to 4TB memory capacity and up to 128 cores within a 1U chassis. The QuantaGrid D43N-3U is a GPU server that supports graphics-intensive AI inference and training tasks.

QCT POD integrates software and hardware to shorten the system deployment process while alleviating the stress of management for IT staff. More importantly, it allows researchers and users to achieve “simplicity” and “efficiency” all the way from system planning to preparation, accelerating time-to-value as a result.

Figure 4. NCKU AI data center powered by QCT

Looking Ahead

To make new discoveries and precise analysis, an enormous amount of data must be processed and managed to perform more HPC simulation and AI computation work. QCT POD solution effectively enhances the process and, at the heart of it, enables professionals in their fields to focus on their job to drive industry breakthroughs and accelerate innovation.

With expertise in HPC & AI technology and the rich project experiences that QCT has had, QCT provides solutions that fully meet customer demands from both a performance and cost perspective. In this way, QCT assists their customers to start the digital transformation journey and paves the way to achieve more.

Access the following link to learn more about QCT POD solution:

https://go.qct.io/qct

United States
QCT LLC., Silicon Valley Office
1010 Rincon Circle, San Jose, CA 95131
TOLL-FREE: 1-855-QCT-MUST
TEL: +1-510-270-6111
FAX: +1-510-270-6161
Support: +1-510-270-6216

China
云达科技,北京办公室（Quanta Cloud Technology）
北京市朝阳区东大桥路 12 号润诚中心 2 号 楼
TEL: +86-10-5920-7600
FAX: +86-10-5981-7958

云达科技,杭州办公室（Quanta Cloud Technology）
浙江省杭州市西湖区古墩路浙商财富中心 4 号 楼 303 室
TEL: +86-571-2819-8650

Japan
Quanta Cloud Technology Japan 株式會社
日本国東京都港区芝大門二丁目五番八号牧田ビル 3 階
TEL: +81-3-5777-0818
FAX: +81-3-5777-0819

Germany
Quanta Cloud Technology Germany GmbH
Rurbenden 48, 52353 Düren
TEL:+49-2421-3863400

Korea
QCT Korea, Inc. (주식회사 큐씨티코리아)
서울특별시 영등포구 의사당대로 97
교보증권빌딩 10 층, 07327
TEL: +82-10-5397-1412
FAX: +82-2-6336-6710

Other regions
Quanta Cloud Technology
No. 211 Wenhua 2nd Rd., Guishan Dist.,
Taoyuan City 33377, Taiwan
TEL: +886-3-327-2345
FAX: +886-3-397-4770

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 components and software partners. QCT designs, manufactures, integrates and services cutting-edge offerings via its own global network. The parent of QCT is Quanta