Get Your HPC Cluster Productive Faster

CONTENTS

REVISIONS

Version	Date	Description	Authors
1.0	09/25/2020	First publish	Alex Wu

Introduction

The nature of designing and implementing high-performance computing (HPC) systems can be quite daunting. Rather than just turning on a single computer, HPC systems require tens to hundreds or more systems that must work together to solve complex problems. The largest supercomputers rely on millions of cores working together to simulate climate change, understand cosmic events, or learn more about computational biology. Commercial organizations are incorporating HPC technologies into corporate workflows to produce more optimized products, develop new vaccines, or understand buying behavior. With the addition of artificial intelligence (AI) software into many of these scenarios, implementing an effective and efficient cluster system has become more time consuming, taking away from the time that could be used to solve many significant problems.

In addition to the challenge of creating a working system with physical servers, the loading and testing of the software stack is also critical for bringing up a system for HPC use cases. Required software may come from a variety of sources, including open-sourced software, free binary downloads, commercial applications that require licensing, and home-grown software. Each of these software packages may rely on different versions of underlying libraries, operating system tweaks, and other knowledge to get the most out of the hardware. The system integrator or the customer, in many cases, will have to dedicate several engineers and system administrators to acquire, install, test, and tune many software packages before allowing researchers and engineers to begin using an HPC system.

By simplifying the deployment process from weeks or longer to days and preparing pre-built software packages, organizations can become productive in a much shorter time. Resources can be used to provide more valuable services to enable more research, rather than bringing up an HPC cluster. By using the services that QCT offers, HPC systems can achieve a better Return on Investment (ROI).

User Pain Points

There are several types of roles that are important to the installation and bringing up of an HPC cluster. These can generally be grouped into the following three categories:

- Administrators – System administrators must load software, test, and tune the cluster before turning over to the end-users. Once operational, a system admin must be able to monitor the cluster and adapt to changing use cases. The admins must have the most up-to-date monitoring tools, both before and during the build out of an HPC system in order for it to become efficient.

- Software developers – Software developers tasked with implementing

scientific algorithms on a cluster of servers should not need to spend their time loading underlying libraries, checking version numbers, and doing other tedious tasks. They can become more efficient at developing and implementing new features based on science if they do not have to spend time on creating an efficient computing environment.

- End users – End users expect that when applications are set up, they can seamlessly submit jobs to the HPC cluster. Various existing software systems are able to allocate resources, change user priorities and keep an expensive cluster busy. A simple yet powerful workload manager that is right for the kind of work being performed will allow scientists and researchers to perform simulations without having to be trained on unfamiliar software.

QxSmart Rapid Deployment Service

Quanta Cloud Technology (QCT) provides a rapid deployment service to customers who require experts to assist in the design and deployment of HPC clusters. The QxSmart Rapid Deployment service accelerates all aspects of the preparation and deployment process, significantly reducing the time for HPC clusters to become operational. QCT has demonstrated that the cluster installation process can be reduced from days to hours when using this service. The benefits for the three groups described above include:

- Administrators can get the system up and running faster with the Rapid Deployment Service, which includes system deployment tools, management and monitoring tools, user authentication tools, and more. Processes can be streamlined and integrated with different administration workflows.

- Developers will benefit from the Rapid Deployment Service due to the inclusion of module files which use pre-defined compilers and several MPI Libraries. Also, if they want to change one environment variable, they can leverage module files without login and logout.

- End users will benefit from this service as well, as they can leverage QCT job script templates to prepare their runtime environment. QCT also provides the templates for them if they want to change or use an updated version of the compiler or the drivers.

The Rapid Deployment Service will reduce time to implementation and speed time-to-market for organizations that choose to work with QCT.

QCT

QCT adopts Intel® Xeon® processor scalable processors with high core count, which perfectly fit in HPC. QCT has excellent relationships with a wide range of worldwide vendors, and among them, QCT works closely with Intel® to offer customers industry-leading platforms with Intel® Xeon® Scalable processors, Intel® SSDs, Intel® Optane™ Memory, and network fabrics, etc. QCT offer customers infrastructure with Intel® platform to reduce effort and time spent on system integration and other tasks so that HPC clusters can be used and maintained with less effort, resulting in faster time to results.

Learn more about QCT's HPC and DL solution here

About QCT

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

Product lines include hyper-converged and software-defined datacenter solutions as well as servers, storage, switches, 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.

http://www.QCT.io