Guang-wen YANG, Hao-huan FU, 2018. Application software beyond exascale: challenges and possible trends. *Frontiers of Information Technology & Electronic Engineering*, 19(10):1267-1272. https://doi.org/10.1631/FITEE.1800459

Application software beyond exascale: challenges and possible trends

Key words: Supercomputing; Exascale; Application

Corresponding author: Guang-wen YANG
E-mail: ygw@tsinghua.edu.cn
ORCID: Guang-wen YANG, https://orcid.org/0000-0002-8673-8254

Motivations

1. According to the current research progress in the US and China, exascale systems are projected to be usable by scientists by approximately 2021.

2. Such an enormous increase in computing power would surely involve hardware innovations in many different aspects. As a result, the radical changes in both architecture and scale would bring completely different challenges to application software developers.

Main idea

In this study we will start our discussion with the current 125-Pflops Sunway TaihuLight system in China and its related application challenges and solutions. Based on our current experience from Sunway TaihuLight, we provide our projection into the next decade and discuss about the potential challenges and possible trends we would probably observe for the future HPC application software.

1. The Sunway TaihuLight Supercomputer: application software challenges and solutions

(1) The first design challenge is to derive the right parallelization scheme maping our target application into the processes and threads using the over 10 million cores of the system.

(2) The second challenge is the memory wall of this system. With a byte-to-flop ratio that is 5 to 10 times lower than other top 5 systems, we need extraordinary memory-related innovations to scale the simulation capability with the 125 Pflops computing performance.

(3) The third challenge is the migration of software to such an architecture with radical changes in both compute and memory hierarchy.

2. Beyond exascale: what to expect for application software

(1) Possible trend 1: programming efforts shifting from compute to data. For current supercomputers, the computing capability is clearly going beyond the data movement capability at different levels. For the exascale and beyond exascale supercomputers in the next decade, unless more revolutionary innovations go to the memory part, we would expect more of such a shifting of programming efforts from the compute part to the data movement part.

2. Beyond exascale: what to expect for application software

(2) Possible trend 2: precision optimization. While most scientific applications are still conservative about sacrificing any bits from double precision, we think the precision optimization strategy which widely used in deep neural networks can also be applied to many simulation applications.

(3) Possible trend 3: programming hardware instead of software. While reconfigurable FPGAs are not well accepted by the HPC community yet, its potential to improve the system efficiency in a significant manner might make it a promising candidate for the next decade.

Conclusions

For the exascale systems and beyond, the continuous demand from scientific simulation and big data analytics will hopefully lead to revolutionary developments in both the way that we build the system and the way that we utilize the system. Unsurprisingly, for the possible trends that we discuss in this study, two are related to the programming part, which solves the problem of mapping science to the underlying hardware system. The other one focuses on precision, which has long been a missing element in the design process of both hardware and software.