"How Fugaku became the first Exascale Machine, and how our Journey towards FugakuNeXT will Continue"

Satoshi Matsuoka

Director, RIKEN Center for Computational Science (R-CCS)
7-1-26, Minatojima-Minami-Macchi, Chuo-ku, Kobe, Hyogo 650-0047, Japan
Professor, Dept. Mathematical and Computing Sciences, Tokyo Institute of Technology
2-12-1-W8-33, Ohokayama, Meguro-ku, Tokyo 152-8550, Japan

Fugaku is the first 'exascale' supercomputer of the world, due to its demonstrated performance in real applications that were expected of exascale machines on their conceptions 10 years ago, as well as reaching actual exaflops in new breed of benchmarks such as HPL-AI. But the importance of Fugaku is "applications first" philosophy under which it was developed, and its resulting mission to be the centerpiece for rapid realization of the so-called Japanese 'Society 5.0' as defined by the Japanese S&T national policy, applications that encompasses conversion of HPC & AI & Big Data as well as CPS (Cyber and Physical) systems with immediate societal impact. In fact, Fugaku is already in partial operation a year ahead of schedule, primarily to obtain early Society 5.0 results including combatting COVID-19 as well as resolving other important societal issues. Now that Fugaku is essentially completed next generation machine or "Fugaku NeXT" is being studied at our R-CCS center, with multiple research projects looking at technology options and developments towards achieving another leapfrogging performance over Fugaku around 2028. The journey, however, is not easy due to the ending of the so-called 'Moore's Law' where lithographic shrinking of the semiconductor process will hit its physical limits, thereby ending the exponential improvement in both transistor count and their energy requirements, both of which were the main sources of performance increase over time. For Fugaku NeXT, there is no panacea to the problem, but rather, innovations from devices and hardware architectures, to software, algorithms and applications themselves, needs to be achieved in a holistic fashion, with detailed analysis of the overheads derived from the current systems. As such the development will even more of a 'Moonshot' effort compared to Fugaku, and its resulting technology should not only benefit supercomputers, but general IT sector overall.