e The aggregate peak bandwidth of the main memory access in each node must be 500GByte／sec or higher.

f Certain number of nodes has the additional non-volatile memory 8TiByte or larger per node and the sum of them in these nodes of the Subsystem must be 128TiByte or larger.

g The theoretical peak injection bandwidth of each node to inter-node interconnect must be 24GByte／sec or higher, and the bi-section bandwidth of the Subsystem-wide interconnect must be at least half of the product of the injection bandwidth and the number of nodes in the Subsystem. The interconnect must have the capability of adaptive routing to choose appropriate links according to data traffic.

h Each node must be equipped with one or more SSDs of 1TByte or larger in total.

i The Subsystem must operate with peak load consuming 600kW electric power or less.

j The Subsystem must have the software stack shown in Item D.

C Subsystem C

a The Subsystem C must be a highly-parallel supercomputer composed of multiple computer nodes. Hereafter, the computer node, or the node in short, is defined as a computer system composed of 36 or more CPU cores, a main memory shared by them, and 4 or more hardware accelerators.

b The total theoretical peak performance of CPU cores on double-precision floating point of each node must be 2.3TFLOPS or higher.

c The total theoretical peak performance of hardware accelerators on half-precision floating point of each node must be 600TFLOPS or higher, and the total sum of them in all nodes of the Subsystem must be 10PFLOPS or higher.

d The CPU core should have the x86—64 instruction set architecture.

e Each node must have the main memory of 128GiByte or larger, and the memory on the accelerators of 128-GiByte or larger.

f The aggregate peak bandwidth of the main memory access and the accelerator' memory access in each node must be 300GByte／sec or higher, and 4,000GByte／sec or higher, respectively.

g The theoretical peak injection bandwidth of each node to inter-node interconnect must be 24GByte／sec or higher, and the bi-section bandwidth of the Subsystem-wide interconnect must be the product of the injection bandwidth and the number of nodes in the subsystem. The interconnect must have the capability of adaptive routing to choose appropriate links according to data traffic.

h The Subsystem must operate with peak load consuming 50kW electric power or less.

i The Subsystem must have the software stack shown in Item D.

D Software stack

a Automatic parallelizing compilers for Fortran2008, C11, and C++14, or their successors, must be provided. These compilers must support OpenMP version 4.5 or later.

b Java SE 13 platform or its successor must be provided.

c The communication library MPI version 3.1 or later must be provided.

d A Linux operating system of kernel 4.12 or later must be provided.

e A job scheduler with the following functionality must be provided.

① Each user group has its own job queue.

② Each user group has three attributes of usable resource amount namely minimum amount, standard amount (upper limit usable for a job) and maximum amount (upper limit usable for all jobs belonging to the user group in total).

③ A job must be judged executable only if the following conditions, where used amount for a group is the total amount of resources used by all executing jobs belonging to the group included that to be used by the job in question for the group to which the job belongs, reserved amount for a group is the margin from its used amount to its minimum amount, and available amount is the difference between the total resource amount of the subsystem and the sum of used amounts for all groups.

・If the available amount is less than the sum of reserved amounts for all groups, the job is judged executable only if the used amount of its group does not exceed the minimum amount.

・If the available amount is not less than the sum of reserved amounts for all groups, the job is judged executable only if the used amount of its group does not exceed the maximum amount of the group. If two or more jobs satisfy this condition but all of them do not at the same time, arbitration of jobs from different groups must be performed based on fair-share scheduling.

④ Jobs in a queue must be prioritized by a combined metric of the submission time and fair-share among users, and the weights of two factors must be able to be adjustable.

⑤ The following three policies to choose a job to run from the prioritized waiting jobs defined 4) above must be available so that one of them can be chosen for a job queue.

・FCFS policy with which the job of highest priority among all waiting jobs is chosen.

・Passable policy with which the job of highest priority among those which are runnable with respect to the resource they need is chosen.

・Backfill policy with which the job of highest priority among those which will not prevent another job, whose priority is highest among all jobs but whose required resource is not available, from starting on its estimated time of start.

⑥ The job scheduler must be capable of interactive run and interactive debug of jobs.

f A wide spectrum of application software and numerical／scientific libraries such as those included in the current supercomputer system must be provided.

g Machine learning frameworks such as TensorFlow, Keras, PyTorch and Apache MXNet must be provided.

E Storage Subsystem

a The total capacity of disk arrays must be 32PByte or larger.

b The aggregate data transfer bandwidth between the disk arrays and the Subsystems A, B and C must be 300G-Byte／sec or higher.

c The total capacity of high-speed storage must be 8Pbyte or larger.

d The aggregate data transfer bandwidth between the high-speed storage and the Subsystems A, B and C must be 1,000GByte／sec or higher.