2 Statistics (School Year 2013)

2.1 System and User Statistics

In the following, we present statistics for operation time taken in the period from April 2013 to March 2014 (SY 2013). In Table 2, we show general statistics of the supercomputer system in SY 2013. The total number of CPUs in System A, B, and C is 64, 3840, and 384 respectively. Consumed disk points amount to about 4%, 5%, and 1% of the total consumed points in System A, B, and C respectively.

In the left column of Fig. 2, availabilities, utilization rates, and consumed points in each system are plotted for each month. Throughout the school year, the utilization rates were high enough. Especially in System B, they were exceeding 90% throughout most of the year. In System C, roughly half of the total utilized resources were used by CMSI projects. This amounts to about 20% of the total usage of the computational resources in this school year. The user statistics are shown in the right column of Fig. 2. The horizontal axis shows the rank of the user/group arranged in the descending order of the execution time (hour×CPU). The execution time of the user/group of the first rank is the longest. The vertical axis shows the sum of the execution time up to the rank. From the saturation points of the graphs, the number of "active" users of each system is around 50 and 220 for System A and B, respectively. The maximum ranks in the graphs correspond to the number of the users/groups that submitted at least one job.

2.2 Queue and Job Statistics

Queue structures of System A, B, and C in SY 2013 are shown in Table 3. In each system, the queues are classified by the number of CPUs the user can use and the maximum duration of each submitted job. In System A, in addition to the usual P class jobs, there is a queue "D1" for debugging, and "L1" for jobs which require only one CPU but quite a long time. Parallel jobs are executed with "P4" and "P16", 16 CPUs being available at maximum with one job using "P16".

In System B, a highly detailed classification is adopted. The biggest portion (20 racks out of 30 in total) of the resources is allotted for "F256", which mainly uses 128 or 256 CPUs at once. "F16", "F32", and "F64" are for smaller-scale jobs using 16, 32, and 64 CPUs respectively. The elapsed-time limit of the above queues is 24 hours for one job, while it is set smaller for smaller-scale queues ("F4" and "F8") to speed up their rotation. For time-demanding jobs, L-type queues are also introduced, whose time limit is set longer than F-type queues. "P64" queue is set up to accept jobs which require any number of CPUs more than 1 and not exceeding 64. "i32" is a queue for debugging, which corresponds to interactive mode in the previous system. In "i32", users can execute their jobs using up to 16 nodes at once from the command line, as if they were logging into the calculation node.

In System C, the "F" and "L" queues are set up similarly to System B. In addition, a debug queue is set up for short debugging jobs utilizing 1 to 4 CPUs, and an interactive queue that can use 1 to 4 CPUs is also available.

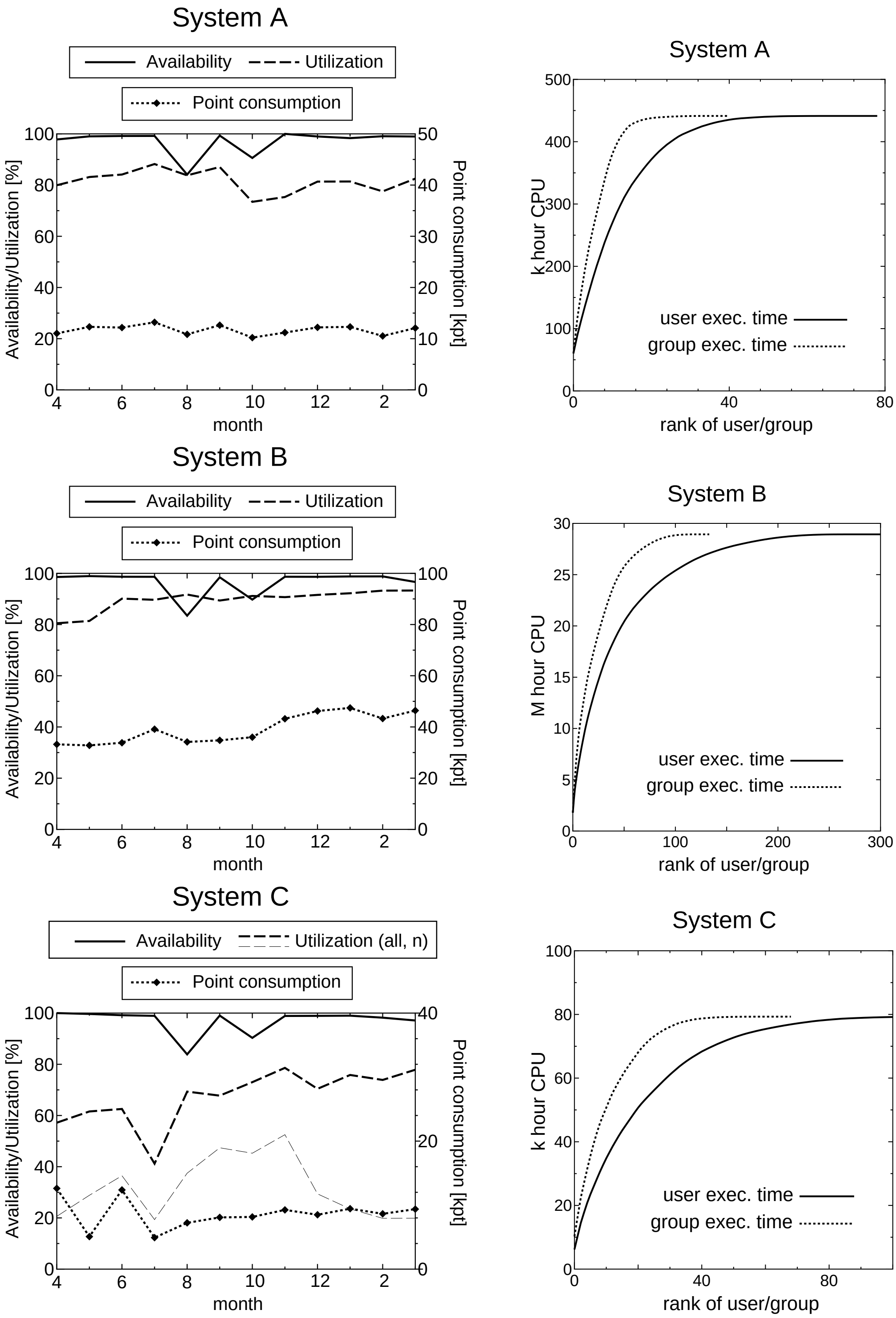

Figure 2: Left: Availabilities, utilization rates and point consumptions of each month during SY 2013. For System C, the utilization by CMSI projects (denoted by “n”) is plotted in addition to the total utilization. Right: User statistics. The horizontal axis shows the rank of the user/group arranged in the descending order of the execution time (hour×CPU). The vertical axis shows the sum of the execution time up to the rank.

Table 2: Overall statistics of SY 2013

	System-A	System-B	System-C
total service time (k hour × CPU)	541.0	32301.1	3243.4
number of executed jobs	27196	189954	29182
total consumed points (k point)	140.7	470.5	85.9
CPU points (k point)	134.8	445.8	84.8
disk points (k point)	5.9	24.7	1.1
points consumed by CMSI (k point)	–	–	42.9
total exec. time (k hour × CPU)	438.8	28931.4	2184.1
availability	97.0 %	96.5%	96.9%
utilization rate	81.1 %	89.6%	67.3%

The CPU points are set smaller for larger-scale queues for System B as shown in Table 3, while it is more uniform in System A. To prevent overuse of the storage, points are charged also for usage of disk quota in the three systems, as shown in Table 4. Disk points are revised often for optimal usage of the resources by examining usage tendencies each year.

Although we do not mention here in detail, to promote utilization of the massively parallel supercomputer, background queues ("B16", "B32", "B64", and "B256"), which charge no CPU points for the jobs, have also been open in System B.

The ISSP Supercomputer also supports large-scale jobs, which use tens of thousands of cores at once by exclusively using the necessary number of CPUs. In-advance application is necessary to execute this type of job. Large-scale jobs can be executed in queues "P512", "P1024", "P2048", and "P3840" just after the scheduled monthly maintenance. However, since such large-scale jobs are now covered by the K Computer, no jobs were executed in these queues in SY2013.

The number of jobs, average waiting time, and total execution time in each queue are shown in Table 5. In System A, the average waiting times of P4 and P16 are a bit long compared with the elapsed-time limit (24 hours). This is because a few active users tend to submit many jobs at once. Because fair-share scheduling is adopted, the waiting time is considered to be appropriate for fair distribution of computational resources. We will continue to look for more appropriate queue settings also in the next school year to meet the user's tendency of resource usage.

In System B, a large portion of jobs have been executed in queues "F16", "F32", "F64", and "F256". As we intended, most of the execution time has been consumed in "F256" and "L256". In all of these queues, the queue settings meet the user's tendencies in that the waiting times are on the order of the elapsed-time limit.

In System C, the waiting times for the "F" queue jobs are less than twelve hours. The "L96" queue has a waiting time of nearly five days, owing to the large amount of resources the jobs occupy when run in this queue.

Table 3: Queue structures in SY 2013

System–A

queue name	Elapsed time limit (min)	# of CPU /Job (n)	# of CPU /queue (p)	memory size	CPU points / (CPU·day)
D1	15	1	2	60GB	7.776
L1	7200	1	4	60GB	7.776
P1	1440	1	10-30	60GB	7.776
P4	1440	4	16-32	240GB	7.776
P16	1440	16	16	960GB	6.048

System–B

queue name	Elapsed time limit (min)	# of CPU /Job (n)	# of CPU /queue (p)	CPU points /(CPU·day)
P1	720	1	32	0.690
P64	720	2-64	64	0.518
F4	720	4	96	0.518
F8	720	8	96	0.518
F16	1440	16	1024	0.518
F32	1440	32	1024	0.518
F64	1440	64	1024	0.518
L16	7200	16	64	0.518
L32	7200	32	64	0.518
L64	7200	64	64	0.518
i32	20	1-32	64	0.518
F256	1440	65-256	2560	0.358
L256	7200	65-256	512	0.358
P512	–	128-512	512 or 1024	0.358
P1024	7200	384-1024	3072	0.358
P2048	–	128-2048	2048	0.358
P3840	1440	1024-3840	3840	0.358

* The available memory size is limited to 21 GB per one node.

* P queues require in-advance application (see main text). The elapsed-time limit for P512 and P2048 queues is determined on a per-application basis.

System–C

queue name	Elapsed time limit (min)	# of CPU /Job (n)	# of CPU /queue (p)	CPU points /(CPU·day)
debug	30	1-4	24	1
interactive	30	1-4	24	1
F12	1440	2-12	60	1
F96	1440	2-12	288	1
L12	7200	24-96	24	1
L96	7200	24-96	192	1

* The available memory size is limited to 28 GB per one CPU.

Table 4: Disk points of System A, B, and C

		point/day
System A	`/home`	$0.0125 \times \theta(q-2)$
	`/work`	$0.005 \times \theta(q-30)$
System B	`/home`	$0.05 \times \theta(q-2)$
	`/work`	$0.005 \times \theta(q-30)$
System C	`/home`	$0.05 \times \theta(q-5)$
	`/work`	$0.005 \times \theta(q-60)$

* q is denoted in unit of GB.
* $\theta(x)$ is equal to the Heaviside step function $H(x)$ multiplied by x, i.e., $xH(x)$.

Acknowledgments

The staffs would like to thank Prof. Takafumi Suzuki (now at University of Hyogo) for developing WWW-based system (SCM: SuperComputer Management System) for management of project proposals, peer-review reports by the SPAC committee, and user accounts. We also thank Ms. Reiko Iwafune for creating and maintaining a new WWW page of the ISSP Supercomputer Center.

Table 5: Number of jobs, average waiting time, total execution time, and average number of used CPU's per job in each queue.

System-A

queue	# of Jobs	Waiting Time (hour)	Exec. Time (k CPU × hour)	# of CPU
D1	6103	0.02	0.30	1.00
P1	15131	29.56	177.65	1.00
L1	1042	7.50	16.90	1.00
P4	3898	26.62	134.10	3.97
P16	978	39.60	109.89	16.00

System-B

queue	# of Jobs	Waiting Time (hour)	Exec. Time (k CPU × hour)	# of CPU
P1	71590	14.56	199.37	1.0
P64	20873	24.07	346.00	4.7
F4	16364	21.36	287.23	4.0
F8	11125	9.18	376.86	8.0
F16	20226	13.98	2575.80	16.0
F32	8990	12.66	2099.62	32.0
F64	5044	29.79	2561.01	64.0
L16	302	44.99	135.22	16.0
L32	85	94.65	160.05	32.0
L64	21	123.45	32.07	64.0
i32	20178	0.02	42.05	21.8
F256	10712	20.09	15405.86	198.3
L256	210	31.35	1737.60	209.9
P512	0	0	0.00	0
P1024	0	0	0.00	0
P2048	0	0	0.00	0
P3840	0	0	0.00	0

System-C

queue	# of Jobs	Waiting Time (hour)	Exec. Time (k CPU × hour)	# of CPU
F12	9084	5.08	309.69	6.4
F96	6900	11.97	1592.51	40.1
L12	408	9.74	49.80	4.9
L96	146	64.67	222.04	30.8
debug	8887	0.04	1.53	2.0
interactive	3165	0.00	0.82	1.2