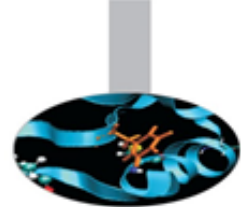


57th Ticket Meeting

HPC User Support @ CINECA
October, 15th 2021

Content



- Status of the clusters - main events affecting production [Sept 14 – October 14]
- Examination of active tickets on HPC-US-SECOND queue
 - escalated to SchedMD support
 - escalated to Intel support
 - escalated to NVIDIA support
 - other active tickets
- Ticket statistics on queue [Sept 10 – October 12]
 - HPC-US-FIRST
 - HPC-US-SECOND
- Module usage on Marconi-SKL

Status of the clusters September 14 – October 14

Main events affecting production



Sept 21st: scheduled maintenance operations on Marconi100 cluster on September 28th

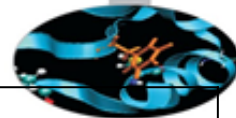
https://www.hpc.cineca.it/center_news/marconi100-scheduled-maintenance-next-tuesday-september-28th

https://www.hpc.cineca.it/center_news/reminder-scheduled-maintenance-marconi100-tomorrow-september-28th

Sept 28th: announced an issue with latest SLURM scheduler version installed on Marconi110 cluster that may cause the failure of jobs with an error log about GPU device unavailability

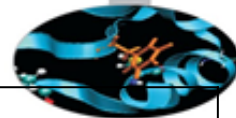
https://www.hpc.cineca.it/center_news/marconi100-back-production-and-possible-killed-jobs

Tickets escalated to SchedMD support



Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
16506	Allocation of GPU jobs in M100	Leo.Ma@ukaea.uk	2021-03-17 16:43:09	M100	2nd SchedMD	<p>The allocation of 1 GPU card job to nodes is pretty random, and not packing tightly; It may be a good idea to pack jobs with some ordering, and occupy less nodes</p> <p>---</p> <p>After a long exchange of messages with SchedMD, we found the way to have four 1-gpu jobs packed on the same node; It is simply needed to discard the directive: <code>#SBATCH --ntasks-per-node=1</code> or modify it as follows: <code>#SBATCH --ntasks=1</code> If not set the ntasks-per-node value will be set to the default (1).</p> <p>A "bug" (or a desired "feature", the SLURM support is still investigating on the matter) is triggered by the <code>--ntasks-per-node</code> directive, forcing the selection of the first free cores per node, and slurm (correctly) reacts choosing a different node to enforce the <code>cpu-gres</code> binding.</p> <p>SchedMD reported: «Fixing this requires a notable change in the select plugin. I am afraid that a proper fix will be available only in slurm 21.08»</p> <p>slurm 21.08 has been released on August 26th. This release has been installed on Marconi100 cluster on September 28th. Tests performed after the slurm upgrade showed the expected behaviour (the bug has been fixed).</p>

Tickets escalated to **SchedMD** support



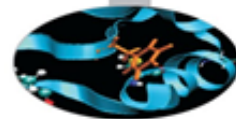
Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
18895	Unclear error message on MARCONI ...	joerg.riemann@ip p.mpg.de	2021-06-24 14:54:01	SKL	2nd SchedMD	<p>The user report randomly observed errors in his jobs as in «srun: error: Unable to create step for job 9552876: Memory required by task is not available».</p> <p>After some extensive checks we made, we got an additional information reported for the second mpirun call: “srun: error: Unable to create step for job 3130006: Memory required by task is not available”</p> <p>It appears that the first mpirun (or, better, the first srun called by the first mpirun) isn't able to free the allocated memory fast enough for give room to the second one.</p> <p>Hence the observed behavior is caused by a slurm bug that was addressed to the slurm support team: “Second mpirun in a job fails after Slurm upgrade”.</p> <p>SchedMD suggested a workaround while waiting for the bug to be fixed in 20.11.8 version. We indicated the user to put a "sleep 5" between the two mpirun commands in the jobscript.</p> <p>Test performed with the latest slurm version, 21.08, installed on Marconi cluster on September 8th showed expected behaviour (the bug has been fixed).</p>

Tickets escalated to Intel support



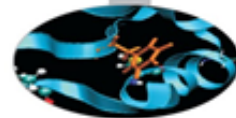
Ticket	Subject	Creation date	Last Updated by Intel	Comments
3932	Pointers to module arrays not working with SIMD nilsm@ipp.mpg.de serhiy.mochalsky@ipp.mpg.de	18/09/2019 17:24:02	31/07/2021 (Intel issue opened by CINECA)	<p>Intel originally found a bug on the compiler front end and solved it. They provided us an archive with the sources, BUILD script and the outputs of the reproducer that did not show the issue. The first bug had "hidden" a second bug on the vectorizer that was triggered by the definition of the macro "USE_ARR_IN_MODULE" in the reproducer. We reported this issue to Intel that confirmed the problem and opened a new bug: Intel bug report CMPLRIL0-33599</p> <p>Intel support has provided a resolution for the problem reported in the second bug: "" We cannot use simd for a loop that has F90 pointer assignment inside. For every iteration of the loop, it is updating the same dope vector for f4a. That means there is a loop carried dependency preventing vectorization. If it is vectorized with veclen =2, for iteration 1 and 2, it is storing into the same location of the dope vector and the address code of f4a is picked up incorrectly. In order to generate the right code, it can run in a non-vector mode. However, attached is fixed.f90 which shows the right way to do it. We need to declare a structure of arrays to store the F90 pointers. In that way, every iteration is storing into different dope vectors. And, of course, the result from fixed.f90 is different because I used different computations in the loop.""</p> <p>We reported to Intel support users' feedback: "" Since I specifically made the f4a pointer private in my omp simd pragmas, it is very unexpected for them to still produce a loop carried dependency. Marking a variable as private is specifically defined as getting around it being a loop carried dependency. The struct of pointers, which fixes the problem, is exactly the kind of solution, the compiler should generate when encountering a pointer variable in a private clause. I hope this is the understanding of the Intel engineers as well, and they'll provide a fix in an upcoming release.""</p> <p>The bug fix will be merged into the compiler, it has been approved its release on Intel oneAPI 2022.0 release</p>

Tickets escalated to NVIDIA support



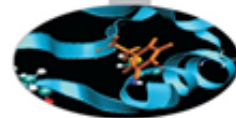
Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
18851	Problem with the nvidia compiler and -lnvc on m100	nilsm@ipp.mpg.de serhiy.mochalsky@ipp.mpg.de	2021-06-23 11:00:02	M100	2nd NVIDIA	The user report a strange behaviour referring to CUDA devices detection when adding (or not) the «-lnvc» flag for the compilation of a test code available on the latest hpc-sdk compiler suite on Marconi100 cluster. We performed some investigations that show that with and without the flag -lnvc the same libraries are linked, the only difference is in the order of the linkage. We are in contact the nvidia support for discussion.
19978	OpenACC compiler problem with function intermediates	nilsm@ipp.mpg.de serhiy.mochalsky@ipp.mpg.de	2021-08-24 17:14:02	M100	2nd NVIDIA	<p>The user provided a test code that is compiled using hpc-sdk/2021--binary module: “The compiler generates an intermediate, called get_arr1 , to store the result of the multiplication in line 26. This intermediate is copied to the device, as seen in the compiler output. Unfortunately, this intermediate needs to be private, as each thread needs its own copy. It is impossible to achieve that, as its name is not known and it doesn't exist during compile time. Array b has the correct result as the result of the function call is saved in the variable c , so no intermediate is necessary. This problem emerges for all function calls to a device function that returns an array.”</p> <p>The issue reported has been escalated to nvidia support for discussion.</p>

Tickets escalated to NVIDIA support



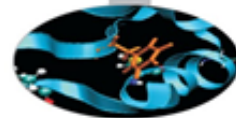
Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
19982	OpenACC compiler problem with writing back from the GPU	nilsm@ipp.mpg.de serhiy.mochalsky@ipp.mpg.de	2021-08-24 17:44:02	M100	2nd NVIDIA	The user provided a test code that is compiled using hpc-sdk/2021--binary module; it do not generate expected results related to a copyout data clause. This problem seems related to the syntax used in the copyout data clause. To complete the analysis the issue has been reported to the nvidia support to ask for further clarifications.
20303	OpenACC compiler problem, unexpected implicit wait	nilsm@ipp.mpg.de serhiy.mochalsky@ipp.mpg.de	2021-09-09 15:58:01	m100	2nd	This issue is still under investigation. We will also report it to nvidia if necessary as soon as preliminary checks will be completed.

Other active tickets on HPC-US-SECOND



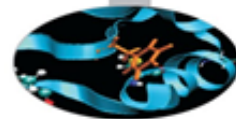
Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
18205	paraview on M100	mattwi@fysik.dtu.dk	2021-05-28 11:46:01	M100	Int.	The user reported issues when running paraview version available as module through a RCM session. To fix this is required to recompile the paraview version (in progress), we will include the installation of the nvidia plugin index required also by this user.
19157	Error on marconi	axel.koenies@ipp.mp.g.de	2021-07-07 11:32:02	SKL	2°	The user reported two jobs running euterpe code that wrote no output and asked for a budget refund of the lost computing resources. We performed checks for those jobs and we did not find any filesystem issue involved. We informed the user about this behaviour previously observed with euterpe code that has been fixed in other situations by recompiling the code; also, about our policy that do not foresees the budget refund due to system issues etc.
18908	Paraview not running with SLURM	louis.stenger@epfl.ch	2021-06-25 07:32:02	SKL	2nd INT	The user reported errors when trying to use paraview in parallel mode. We informed the user that the present installation of this module is a serial implementation to be used through RCM sessions.

Other active tickets on HPC-US-SECOND



Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
20015	Device to device MPI communication with OpenMP4.5	asahi.yuichi@jaea.go.jp	2021-08-26 08:08:03	M100	2nd	The user report errors when trying to compile his code. This first error refers to a XLC++ version of the code that could compile and run correctly before the major update of the cluster. We helped the user into the compilation of the code by including a missing linking flag. The user confirmed that he can go on with the work. The second error reported refer to the usage of hpc-sdk compiler suite for the compilation of the same code (FATAL ERROR: "data in use_device clause was not found on device 1: host:0x2140dbf02), this is still under investigation.
20019	Segfault before code launch with srun	Patrick.TAMAIN@cea.fr	2021-08-26 09:28:03	SKL	2nd	The user is experiencing issues when trying to run his code. For the code compilation are used both modules available on the system and custom libraries available at the user personal areas. The error seems related to the petsc libraries (module) used for the compilation of the code. Investigations are underway.
20676	Qdstrmlmporter	huw.leggate@dcu.ie	2021-09-30 13:04:01	M100	2nd	The user reports that using nsys the Qdstrmlmporter step fails for the qdstrm files. He tried also the execution of Qdstrmlmporter from the command line obtaining the same result. He is attempting to use openmp offloading with the default gcc 8.4.0, that appears to have been built without offloading enabled, and he suspects that may be causing the crash.

Other active tickets on HPC-US-SECOND

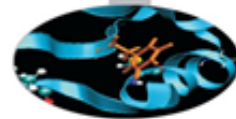


Ticket	Subject	Requestors	Created	Host	Supp. Level	Comments
20832	GNU/8.4.0 Offloading support	huw.leggate@dcu.ie	30/09/2021 13:04:01	M100	2nd	The user reports that the default gnu version 8.4.0 on Marconi100 appears to have been compiled without offloading support, and asks for the possibility to provide an 8.4 build with offloading enabled as this is the required version to build other libraries.
21125	Prioritization of jobs	mhoelzl@ipp.mpg.de	2021-10-11 15:14:01	SKL		The user asked for the spamela0 user's jobs prioritization for a period of 3 weeks to match an important deadline. This user would run jobs requiring 400 nodes at most for a max. time of 1:30 hours, to be submitted sequentially.
21153	Gitlab access and ssh keys	hm1234@york.ac.uk	2021-10-12 12:56:01	SKL	2nd	The user manages to login to Marconi cluster via ssh but he is experiencing issues when trying to access using ssh-keys, that prevents him the local execution of jupyter kernel running on Marconi. We are supporting the user into the generation of the keys that seem not to work anymore when using rsa (this works with ecdsa).

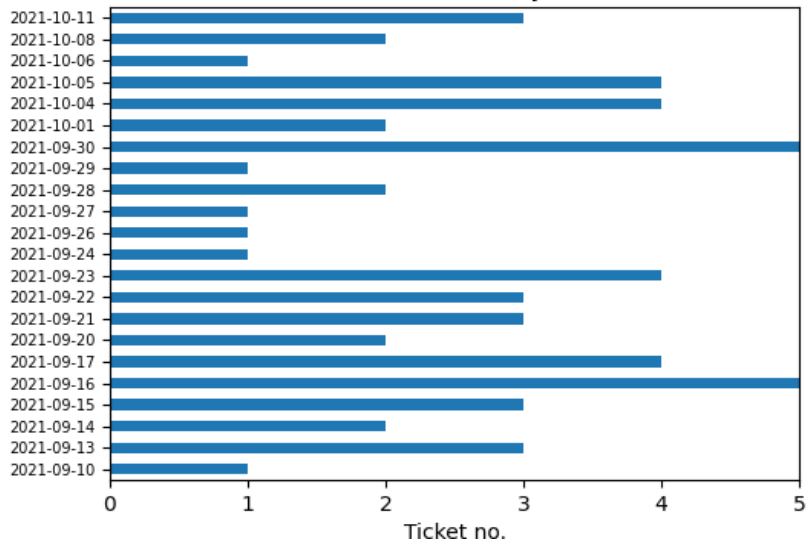
Ticket Statistics

Sept 10 – October 12

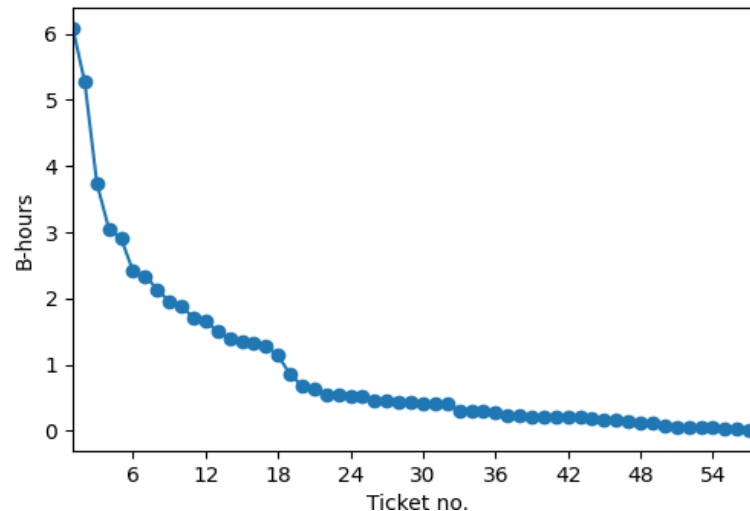
HPC-US-FIRST & HPC-US-SECOND



Received tickets by date



Business hours from ticket creation to initial address
HPC US FIRST & HPC US SECOND



TOTAL	→ 57 tickets
- HPC-US-FIRST	→ 27
- escalated to HPC-US-SECOND	→ 30

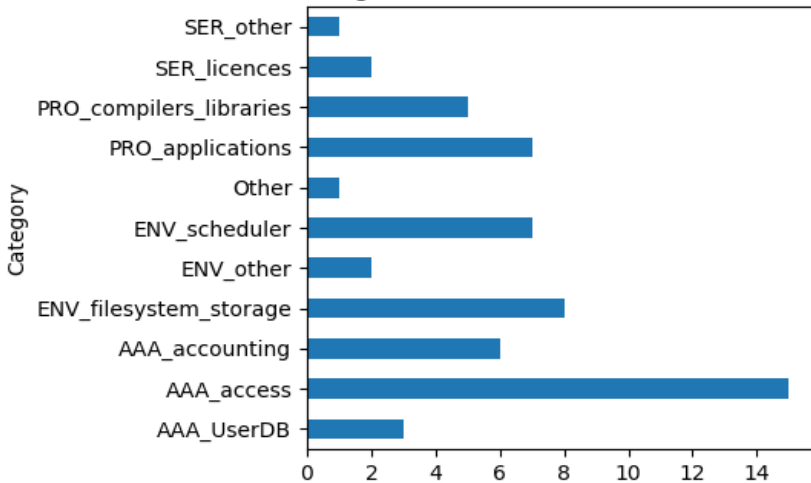
Ticket statistics

Sept 10 – October 12

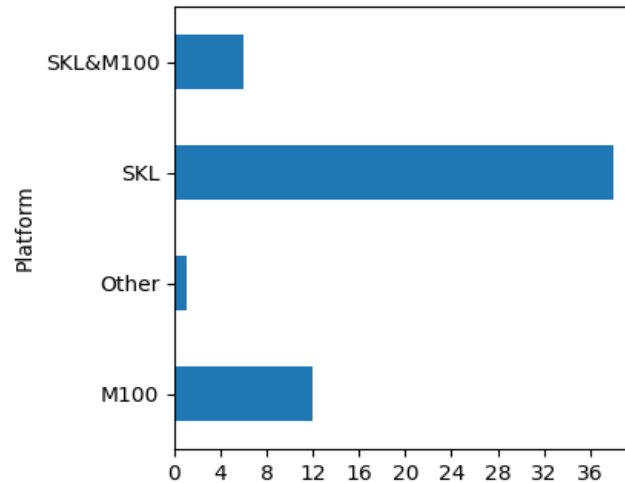
HPC-US-FIRST & HPC-US-SECOND



Ticket categories on HPC-US-FIRST & HPC-US-SECOND

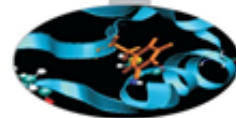


Classification of tickets by platform
HPC-US-FIRST & HPC-US-SECOND

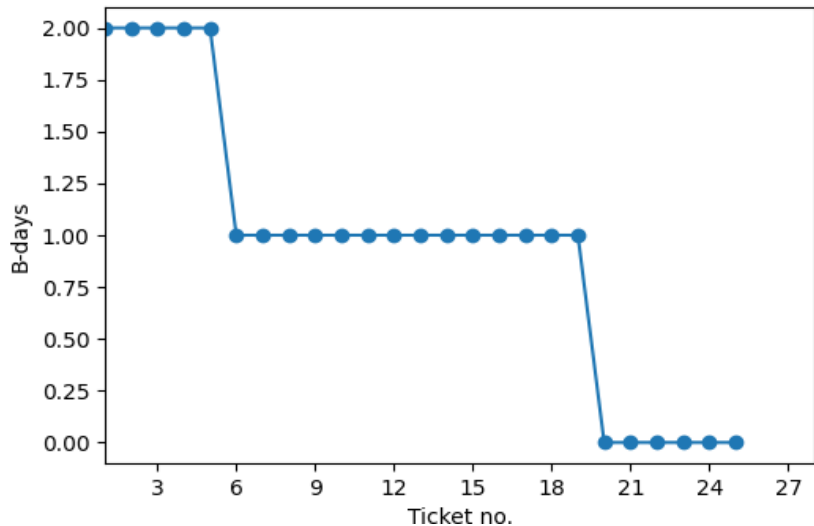


Ticket statistics Sept 10 – October 12

HPC-US-FIRST & HPC-US-SECOND

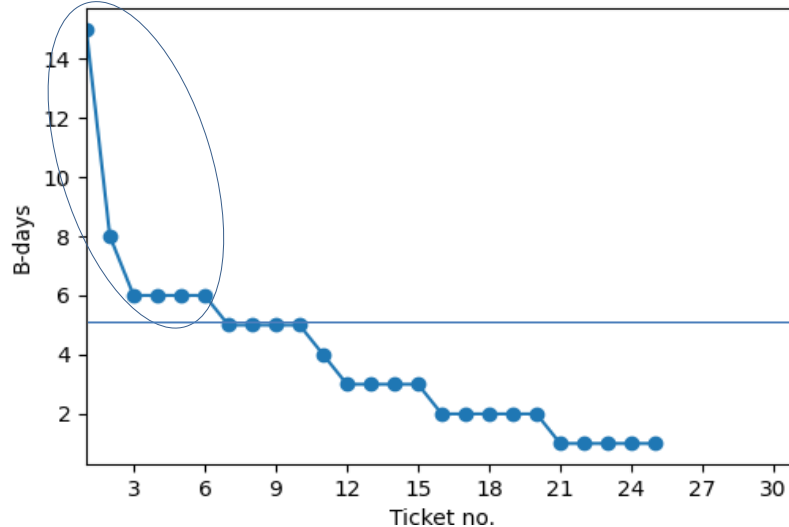


time from ticket creation to ticket resolution
HPC US FIRST



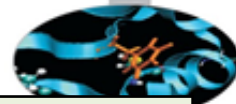
25 / 27 resolved tickets

time from ticket creation to ticket resolution
HPC-US-SECOND



25 / 30 resolved tickets

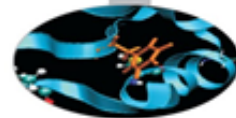
Ticket statistics Sept 10 – October 12



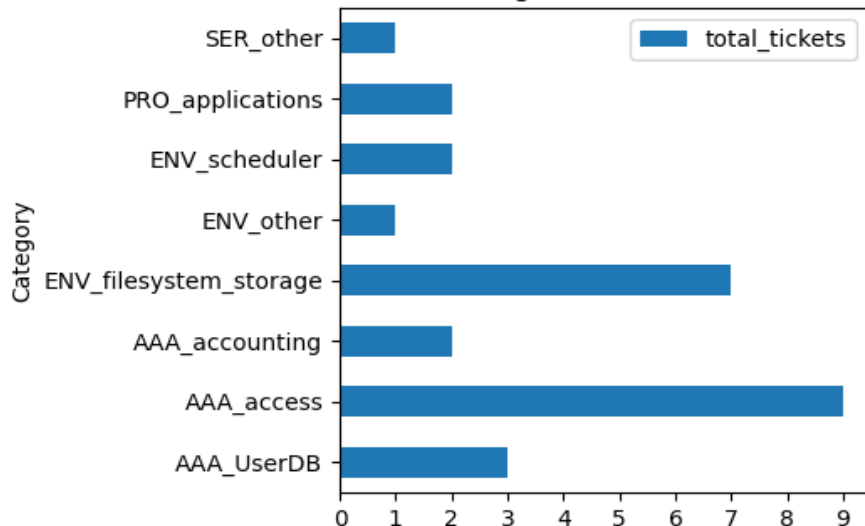
Ticket SECOND	Days	Creation time	Requestor	Subject	Notes
20407	15	2021-09-14 11:44:02	antonio.froio@polito.it	Problemi di connettività verso server licenze da MARCONI	The user reports connectivity issues when trying to contact the Siemens licese server, necessary to run his starccm jobs. Together with Siemens support we managed to resolve user's issues.
20517	8	2021-09-16 18:02:01	anna.perona@polito.it	Richiesta accesso su Marconi	We provided to reset the password for this user that contacted us back to get a copy of this.
20554	6	2021-09-17 18:28:01	zhangyujiascott@gmail.com	connection refused	The user reported access issues and asked for a password reset. This was non changed as the user managed to login before the reset request. This ticket was closed after some days.
20633	6	2021-09-21 16:20:02	markus.held@chalmers.se	Killed postprocessing on marconi m100	The user reported some interactive jobs killed after a certain time. We reminded him about the cpu time limit imposed on the login nodes and suggested him to use the bdw_all_serial partition. This ticket was closed after some days waiting for user's feedback.
20678	6	2021-09-23 09:58:01	leonhard.leppin@ipp.mpg.de	Marconi Bus error	The user reported some GENE jobs failing with Bus error. Results from checks performed for those jobs suggested an out of memory related error.
20835	6	2021-09-30 14:02:01	lah583@york.ac.uk	marconi password	We provided to reset the password for this user. The ticket was closed with some delay.

Ticket statistics Sept 10 – October 12

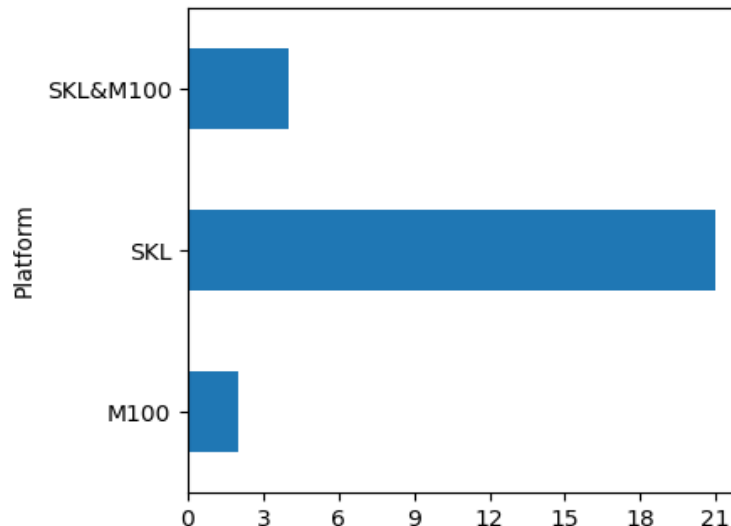
HPC-US-FIRST



Ticket categories on HPC-US-FIRST

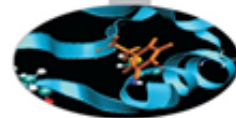


Classification of tickets by platform
HPC-US-FIRST



Ticket statistics Sept 10 – October 12

HPC-US-FIRST

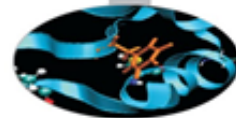


HPC-US-FIRST		By status			By platform		
		open	resolved		M100	SKL	SKL&M100
Information	AAA_accounting		1				1
	ENV_filesystem_storage		1			1	
	ENV_other		1			1	
	SER_other		1			1	
Problem	AAA_access	1	3		1	3	
	PRO_applications		2		1	1	
Service Request	AAA_UserDB		3			2	1
	AAA_access		5			3	2
	AAA_accounting		1			1	
	ENV_filesystem_storage		6			6	
	ENV_scheduler	1	1			2	
total	27	2	25		2	21	4

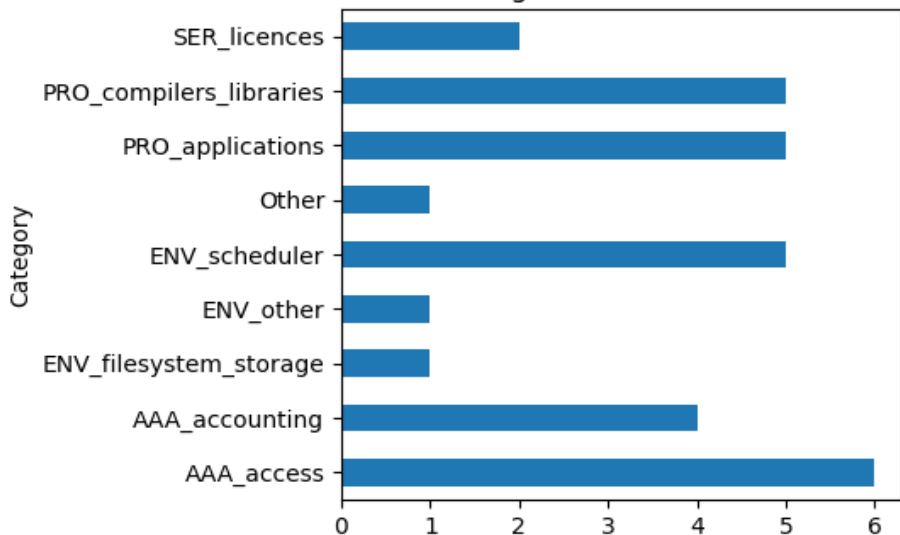
Ticket statistics

Sept 10 – October 12

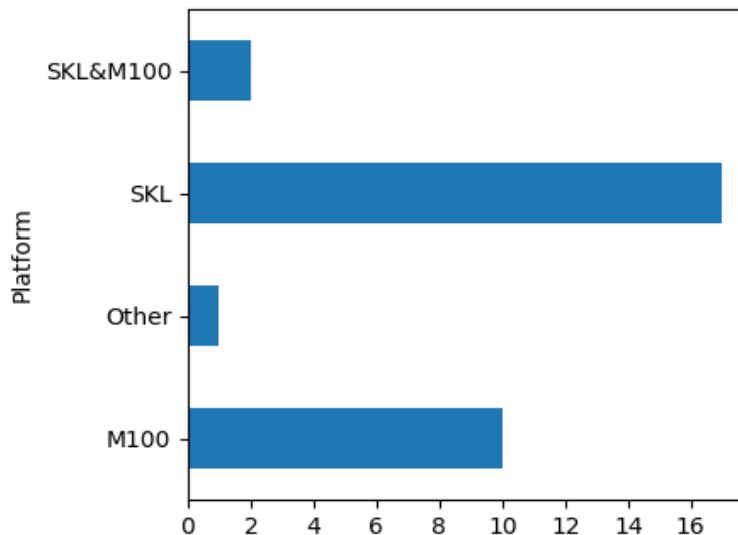
HPC-US-SECOND



Ticket categories on HPC-US-SECOND

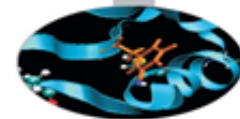


Classification of tickets by platform
HPC-US-SECOND



Ticket statistics Sept 10 – October 12

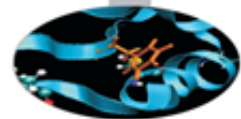
HPC-US-SECOND



HPC-US-SECOND		By status		By platform			
		open	resolved	M100	SKL	M100&SKL	Other
Information	AAA_accounting		2		1	1	
	ENV_scheduler		2	2			
	PRO_applications		2	2			
	PRO_compilers_libraries		2	2			
Problem	AAA_access		4		3	1	
	AAA_accounting		1		1		
	ENV_other	1			1		
	ENV_scheduler		2	1	1		
	PRO_applications		3		3		
	PRO_compilers_libraries	3		2	1		
	SER_licences		2		2		
Service Request	AAA_access		2		2		
	AAA_accounting		1		1		
	ENV_filesystem_storage		1		1		
	ENV_scheduler	1		1			
	Other		1				1
total	30	5	25	10	17	2	1

Module usage on Marconi-SKL

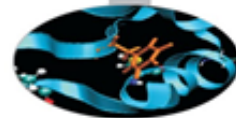
Batch jobs on September 2021



category	modulename	partition	qos	module_load_count	elapsed_timelimit_sum
compiler	intel	skl_fua_dbg	normal	38	3,54
		skl_fua_prod	normal	5	1,98
	intelmpi	skl_fua_dbg	normal	38	3,54
		skl_fua_prod	normal	5	1,98
	python	skl_fua_prod	normal	6	2,27
			skl_qos_fuabprod	6	2,53
tool	cubegui	bdw_all_serial	normal	18	11,79
	matlab	skl_fua_prod	normal	3	0,22
	scalasca	bdw_all_serial	normal	18	11,79
	scorep	bdw_all_serial	normal	18	11,78
	vtune	bdw_all_serial	normal	3	1,12

Module usage on Marconi-SKL

Batch jobs on September 2021



category	modulename	partition	qos	module_load_count	elapsed_timelimit_sum	
library	cubelib	bdw_all_serial	normal	17	10,77	
	fftw	skl_fua_prod	normal	6	2,27	
			skl_qos_fuabprod	6	2,53	
	hdf5	skl_fua_prod	bdw_all_serial	normal	45	33,78
			normal	6	2,27	
			skl_qos_fuabprod	6	2,53	
	mk1	skl_fua_prod	skl_fua_dbg	normal	38	3,54
			normal	5	1,98	
	mumps	bdw_all_serial	normal	1	1	
	netcdf	skl_fua_prod	normal	6	2,27	
			skl_qos_fuabprod	6	2,53	
	netcdf	skl_fua_prod	normal	6	2,27	
			skl_qos_fuabprod	6	2,53	
	petsc	skl_fua_prod	bdw_all_serial	normal	47	35,77
			normal	6	2,27	
skl_qos_fuabprod			6	2,53		
qt	bdw_all_serial	normal	18	11,79		