

ASC15 Q&A Update

Dear participants of ASC15:

The ASC15 committee has grouped answers for those new questions posted so far. The details are as follows for your reference.

ASC15 Committee

Feb, 2015.

Part1: Gridding optimization

Q1: One case of MKL in the notices doesn't support offload-build.

Answer:

Please remove '-offload-build' option from your compiler directives. That option has been built in the newer compiler, so you do not need to include it manually.

Q2: If I want to do MPI implementation for the Gridding program (Question 4), can I modify the main file tConvolveMPI.cc besides Benchmark.cc? The reason I want to do this is MPI tends to spawn multiple processes to execute the whole Gridding program, so even if I initialize MPI inside the gridKernel() function of the file Benchmark.cc, other stuff (in the main file) still get executed, so I want to initialize MPI inside the main file instead and let the root rank to process trivial jobs (printing messages, initializing convolution function, ...). Is it acceptable? What degree of modification can I make to the Gridding program? Is it just the gridKernel() function or am I allowed to modify something else?

Answer:

As it has been mentioned in the contest notification,

“All the modifications of the program should be within the region that we measure the running time.”

It means that if you want to use MPI to parallelize your code, please be sure to put all the MPI clauses in that region.

Q3: About gridding: should the final results run under 2CPU+MIC? If so, does it mean that there is no need using multi-node methods like MPI in the preliminary round.

Answer:

The following is the answer to your question which has been mentioned in the notification,

“... the final program should be tested with the Workload1 case on one computing node in the CPU+MIC hybrid cluster (the remote testing platform) provided by the organization committee.”

“... You may use some parallel programming methods such as MPI, OpenMP, pThread and OpenCL to write your code, but all these methods should be supported by MIC. ...”

Please also note that the code that is programmed with MPI can also run on a single node.

Q4: For the last question which only counting the gridKernel computing time ,does it mean that we only have to optimize that gridKernel as much as possible?

Answer:

Answer:

Yes, gridKernel function is the only function that you need to optimize.

Q5: The data for the last question is not big enough, could we have a larger dataset?

Answer:

We only provide that series of test data for the preliminary round

Q6: According to the ASC15 Preliminary Contest Notifications, *‘The function named “gridKernel” is the kernel function, which is the only function that is required to be optimized.’* Does it mean that the only part to be optimized is gridKernel of Benchmark.cc, while other parts can be changed as well if needed? And that kind of modification is necessary and allowed.

Answer:

It has been highlighted in the ASC15 Preliminary Contest Notifications that, all the modification for the program should be only conducted in the timing part, anything else could not be changed. In the preliminary program, timing part is only for gridKernel

Part2. NAMD optimization

Q1: Whether the NAMD example file can be modified?

Answer:

NO

Q2: What is your expected wall clock for NAMD application workloads optimization? Besides, whether NAMD can be optimized through deep codes optimization?

Answer:

Different platform has different testing time. We are sorry that the workload running time won't be published. And it's fine to optimize through codes optimization.

Q3: How can I derive "ns/day" from this information? What does "ns/day" exactly mean? I read somewhere and it said that it is the simulation time of protein but am not be able to understand clearly. Do we have to strictly follow the metric "ns/day" in our proposal report?

Answer:

You could get the 'days/ns' informatios from log file, you can use the days/ns value to compute the ns/day value.

And the following figure is a log file's example.

```
Info: Benchmark time: 8 CPUs 0.652145 s/step 7.54797 days/ns 245.386 MB memory:
TIMING: 500 CPU: 334.143, 0.631054/step Wall: 336.824, 0.647479/step, 0 hours remaining, 245.386139 MB of memory in use.
ETITLE: TS BOND ANGLE DIHED IMFRP ELECT VDW BOUNDARY MISC KINETIC
ENERGY: 500 108640.2894 81821.1639 16089.4458 1631.0389 -1336113.9688 112087.8508 0.0000 0.0000 290630.9340
WRITING EXTENDED SYSTEM TO OUTPUT FILE AT STEP 500
WRITING COORDINATES TO OUTPUT FILE AT STEP 500
WRITING VELOCITIES TO OUTPUT FILE AT STEP 500
WallClock: 343.864380 CPUTime: 340.176300 Memory: 259.230255 MB
```

NS(n

anosecond is the system's(protein) simulation time-scale,it is the virtual time for the molecular dynamics. For a fixed system (Protein), the longer ns/day,the higher performance.

You do not have to to strictly follow the metric "ns/day" in your proposal report. But you have to submit your log file of NAMD.

Part 3: the architecture and HPCC

Q1: I have the following questions:

- 1) Can I install 2 NVIDIA GPGPU cards (K80) inside the Inspur NF5280M4 server?
- 2) Can I install one NVIDIA GPGPU (K80) + One Intel XEON PHI-31S1P inside Inspur NF5280M4 server?
- 3) If your answer is yes, please let me know the memory requirements for each case.

Your prompt response will be highly appreciated since we are in a process to set up our lab.

Answer:

You can install 2 NVIDIA GPGPU cards (K80) and one NVIDIA GPGPU (K80) + One Intel XEON PHI-31S1P inside Inspur NF5280M4 server

Q2: How to download the entire Gridding code from the remote server (located at `ascxxx@asc15.inspur.com:/home/Gridding`), as we want to try the code on our local machines first? We have tried to use `rsync`, `scp` but were not able to pull it to our server machines.

Answer:

The remote platform locates in the wide area network, while your local machines maybe locate in the local area network, so it can't directly pull the file to your local in the remote server. But you can download the file from remote server just login in your local machines using the command: `scp -r ascxxx@asc15.inspur.com:/home/Gridding <your local directory path>`.

Q3: Should the power supplying of 6pin and 8pin in the accerate card be connected ? What is the set power of the computer?

Answer:

Yes , 6 pin and 8pin should be plugged in. The set power of the 31S1P Mic card is 270w. Be advised, that there may need to have over 300w left for MIC besides the power for other part.

Q4 : Whether the HPCC code can be changed?

Answer :

The efforts are welcomed for HPCC optimization in many aspects, such as code modification, compilation and environment variable etc. The modification of the HPCC code has to follow certain rules. If you wish to optimize HPCC through codes optimization, you have to explain what do you do and how it works in detail, which is important for your scores.

Q5 : In mic-cm

After compilation through `icc -std=c99 -openmp -mkl -mmic -o nbody nbody.c`

Ran: `./nbody`

The output is: `-bash: ./nbody: cannot execute binary file`

How to solve this problem?

Answer :

We don't provide the nbody in our workloads due to the platform limitation. Generally speaking, we don't provide technical support to other workloads. It was not suggested that running other workloads in our platform. Thank you for your understanding.

Two suggestions for your problems

✚ delete `-mmic` in your compilation

✚ run executable codes in MIC

Q6 : When is the final?

Answer :

Pay attention to the notice in <http://www.asc-events.org/> to get more timely details for ASC15. It will be in late May.

Q7 : CPU+GPU is accepted in the HPL test last year, while there is no such kind of codes for heterogeneous platform in the HPCC test this year. It hasn't been made clear whether we can use GPU for computing in the Preliminary Contest Notifications, so if possible, can we use the modified HPCC codes to make GPU part of the computing?

Answer :

The efforts are welcomed for HPCC optimization in many aspects, such as code modification, compilation and environment variable etc. The modification of the HPCC code has to follow certain rules. If you wish to optimize HPCC through codes optimization, you have to explain what you do and how it works in detail, which is important for your scores.

Q8 : Could we do more optimization than only modifying the functions according to the rules posted in the HPCC official website? Will it be acceptable to make it run in the heterogeneous platform with changes more than those functions but no changes in algorithm?

Answer:

Please follow the instructions in the official website.

Q9 : Could you please help us clarify on the following questions?

- ✚ For Question1 HPCC benchmark, can we use different accelerator to achieve better performance for different measures? For example, we want to use GPU to get HPL measure; and we want to use maybe Xeon Phi to run other measures. Is this ok?
- ✚ For Question 2, is power consumption value (e.g. 270W) the average measure or upper/lower limit measure? Are we supposed to compute using the values provided directly?

Hope you could help us better understand the question.

Answer:

- ✚ HPCC is an integral part of the benchmark, Can't split . All results must be from the same results file .
- ✚ 2.270W is the rated power of 31s1p , So you need to evaluate the power consumption in specific environment.

Q10: The asc15.inspur.com is not available. You may visit by using ssh.

Answer:

Please try IP 111.207.123.4 or login through remote access platform

Q11: The hosts file may be required besides Parallel.pbs or mic.pbs when submit parallel jobs or MIC jobs. Probably, our hosts file is not correct, the job cannot finish and only can be qdeled. So how to compse the host file, what are the specific names of each nodes?

Answer:

In the remost test platform, jobs are managed by job secheduling system. When submit jobs, no host files is needed. Please just modify the #PBS -l nodes=nodes number:ppn=cores to submit your jobs. After successful submitssion, system shall automatically regenerate the host file per your parameters and assign the nodes accordingly.

Q12: What is the node for MIC compilation? How can we submit the compilation jobs?

Answer:

The compilation node for MIC in the ASC15 remote test platform is MIC-CM. You can login into MIC node through 'ssh mic-cm' to have your compilation. Please use script of mic.pbs for submission.

Q13: I failed to copy files by scp from the remote test platform, after login to asc15.inspur.com. How to download the files from asc15.inspur.com?

Answer:

Login to the local host and use command-- scpascxxx@asc15.inspur.com:/home/ascxxx/ (original files) you want to copy local directory.

If it still doesn't work, you can try software like xftp to do so.

Q14: The first question in the preliminary round requests refers to the construction of a hardware platform under 3000w budget. But we don't have any of that hardware. Then, how can we set up the platform after we identify the hardware we need.

Answer:

For the first question we need no more than the construction proposal.

Q15: If so the first question only requires a proposal, then the HPCC Test and the NAMD Test can not run in practice, which means we can only have results through analysis.

Answer:

HPCC and NAMD can be partly tested through the remote test platform or your local owned platform. You can predict the results by small size test or and theological analysis. We pay more attention to the theological analysis and predicting.

Q16: Our team has a problem submitting jobs via the .pbs scripts. We tried running the original tConvolve with a slightly modified .pbs script (just changed the path to the executable)

Answer:

There are something wrong with the path of your execut file, you should copy the Parallel.pbs to the corrent path.

Q17: I submitted a test mpi program to joblist but the job are submitted to the master node. and it's always queueing .. How can i submit a job to cpu nodes and mic nodes and why the job always queueing

Answer:

The resource of the system was running at full capacity when you qsub the job, you can try to reduce the request resource of the job or just wait a moment, and now you can qsub job using the corresponding pbs script.

Q18: How much resource can one account take up when testing parallel codes in the ASC15 platform?

Answer:

2 computational nodes, 48 kernels at most.

Q19: Can we use 2 MIC cards for one job?

Answer:

Yes

Q20: We are wondering whether we should run our HPCC & NAMD test on the hardware configuration from the first question

Answer:

There is no need to run the HPCC & NAMD on the configuration from the first question. You can test the HPCC & NAMD on the remote platform or your local platform. We pay more attention to the analysis method and process than the final results.

Q21: How to use the VNC in the remote test platform? (We are trying to do ide debugging.)

Answer:

We have limited the usage of vnc in the remote test platform you can do ide debugging in your local platform

Q22: In the problem of Design a HPC system, you states that inspur NF5280M4 server's power consumption estimation is : E5-2680v3 TDP 115W, memory nominal 7.5W, hard disk nominal 10W. I want to know 115W means only 1 CPU or 2 CPUs and 7.5W means 16G memory or 16G*8 memory

Answer:

The consumption of CPU & MEM means only one of them, not the sum of the entire server.

Q23: We used mkl in our testing program. But it informs us of lacking libmkl_intel_lp64.

What should we do? Our account doesn't have the modification permission for /lib64.

Answer:

Please check your route for mkl in your compilation. And the correct one is /home/opt/intel/mkl/lib/intel64.

Q24:When running NAMD in the remote test platform (CPU+MIC), it informs us *offload error:Cannot load library to the device 0*

Answer:

Please check

- ✚ whether the test workloads and related libraries have been copied to the MIC
- ✚ whether the settings of compilation applied is correct

Q25: What is your expected wall clock for NAMD application workloads one and two after the optimization? Besides, whether NAMD can be optimized through deep codes optimization?

Answer:

Different platform has different testing time. We are sorry that the workload running time won't be published. And it's fine to optimize through codes optimization.

Q26: When reading some documents regarding to measure the performance of NAMD, I found that they usually refer to a metric called "nanoseconds per day" or ns/day.

How can I derive "ns/day" from this information? What does "ns/day" exactly mean? I read somewhere and it said that it is the simulation time of protein but am not be able to understand clearly. Do we have to strictly follow the metric "ns/day" in our proposal report?

Answer:

- ✚ You could get the 'days/ns' informatios from log file, you can use the days/ns value to compute the ns/day value.
- ✚ ns(nanosecond) is the system's(protein) simulation time-scale,it is the virtual time for the molecular dynamics. For a fixed system (Protein),the longer ns/day,the higher performance.
- ✚ You do not have to to strictly follow the metric "ns/day" in your proposal report. But you have to submit your log file of NAMD.

Q26: I have submitted a job by qsub. In my source code, I have used the printf function, but the execution did not print result to my terminal

Answer:

The job is submitted to the compute nodes where you can not get the output directly, your printf fuction was executed on the compute nodes in fact. But you can get the result indirectly using the commands below:

- ✚ Using the command qpeekOut jobid to check the output or qpeekErr jobid to check the error output.
- ✚ Specifing the output to your local file in the .pbs script ,and check the file later to get the output.