

# INTRODUCTION TO OPENACC COURSE

## Lecture #2 Q&A



#	Question	Answer
1	How do vector/worker/gang map to the CUDA concepts of thread/block/grid?	Gang => block, Vector => thread, Worker => thread
2	num_workers*vector_length = block size?	Yes
3	What if a data structure changes size such as size grows in the CPU, can you still use structured data region directives? Does the data size change accordingly in the device?	This is not done automatically unless you are using managed memory. If you change your data structure on the host, you will need make the same changes on the device.
4	Do you mean the acc update device could only update the value not the data structure?	The data clauses perform a shallow copy of the structure. If the structure contains pointers, the pointer itself will be copied but not the object it points to.
6	Why OpenACC is not freely available for windows as cuda, openmp, mpi and will it be freely available anywhere in future?	I can only speak to PGI's availability. The problem there is we have to distribute some Microsoft components that are not free to us. We're trying with Microsoft to find a way to make these bit available for free as well. No promises, but hopefully.
7	What's difference between #pragma acc copyin(...) versus #pragma acc data copyin(...)	pragma acc copyin"" isn't a valid directive. Did you mean ""pragma acc kernels copyin""? If so, they both copy the data to the device, but the lifetime of the data will be different. Paired with a compute region, the data lifetime is only within that
8	I found and watched first lecture video from 2015 year. Is it similar to this year first video?	Lectures from the last year are similar, but we have updated the content for 2016
9	Is gcc compiler supports openacc directives?	GNU 6.x does have limited support for OpenACC with full support still in development
10	Does OpenACC use CUDA streams - when NVIDIA GPU is accelerator - to asynchronously copy? Furthermore can it split up a big copy automatically to start working quickly on large data?	Yes. This is the "async" clause. It will be covered next week.
11	Is there a way to print out an overview of your hardware with information like # of SMX's and max vector size per SMX?	PGI has a utility called "pgaccelinfo" which will show this.
12	What if size of memory we are attempting to copy in is larger than GPU memory?	The program will get a runtime error stating that you are out of memory.
13	Can I try OpenACC compilation free on Windows?	The PGI compilers for Windows does come with a 15 day free trial, but does require a for fee license to continue after that.
14	What is the difference between OpenACC, OpenCL, OpenMP?	OpenACC and OpenMP are directives-based models. OpenCL is a low level programming language. OpenACC has been designed for parallel programming and works across multiple architectures.

# INTRODUCTION TO OPENACC COURSE

## Lecture #2 Q&A



15	What is NVLINK?	NVLINK is a high-speed interconnect for data transfers between GPUs and GPUs and CPUs
16	In the future, OpenACC could have the same performance than normal cuda-c, cuda-fortran ?	The exact performance will depend upon the particulars of your program. On average we find that OpenACC achieves about 80-90% performance of a well-tuned CUDA program.
17	If the compiler reuses the data that exist on GPU, it can cause trouble for future. How can I errase the data existed on GPU before running the code?	This question has been answered verbally.
18	Is it possible to run OpenACC on 32 bit NVidia Tegra TK1?	Support for ARM processors, including NVIDIA's Tegra product is on PGI's roadmap but won't be available for awhile
19	Does OpenACC automatically parallelize over GPU and CPU at the same time, or the programmer must choose where to run the code? I.e. a dual XEON Workstation with an Nvidia Quadro, Can OpenACC heterogeneous parallelize the code?	No, not at this time. The issue is the data and how to distribute it across discrete memories.
20	How does non-coalesced memory access pattern affect performance? ex. $a[i] = b[f(i)]$ where f is mapping function	Poorly. The PGI compiler does attempt to use texture memory when possible to help mitigate this.
21	Is it possible to optimize the code, if the code does not have a loop?	You can offload serial code to the GPU, but you wont be taking advantage of all the parallel processing available.
22	How is ""granularity"" and the level of control allowed with OpenACC ? OpenCL/CUDA can tune the groupe size/ and the size of execution domain, group memory (shared memory) is automatically managed? Is there a similar memory model in OpenACC as in OpenCL?	You have some control, but OpenACC targets a generic accelerator device so some device specific features are not available. To tune the group size, you can set the widths of the "vector"
23	How can I allocate on the device a 3 dimensional (dynamically allocated) array? $A[i][j][k]$	You can add rectangular multi-dimensional arrays directly in a data clause. For example <code>"#pragma acc data copy(A[0:n][0:m][0:p])"</code>
24	Is there a good tutorial for compiling OpenACC with gfortran or g++?	You're taking it! While Jeff is using PGI in his examples, they apply to all compilers that support OpenACC
25	Does Qualcomm Snapdragon GPU support OpenACC?	Not that I'm aware of.
26	How much does the PGProf cost?	For all OpenACC licensing questions - could you send email to <a href="mailto:openacc@nvidia.com">openacc@nvidia.com</a> , we will provide the answer there. Thanks!
27	Is it necessary use to PGI compiler in fortran applications or gfortran its possible?	GNU 6.x, including gfortran, has limited support for OpenACC with full support still in development.

# INTRODUCTION TO OPENACC COURSE

## Lecture #2 Q&A



28	It would be nice if the compiler, given the data size of the arrays in the loop and a user-specified accelerator model, would optimize the gangs/workers etc.	By default, the compiler does optimize the loop schedule (gang,workers,vector) for the particular target device.
29	Can PGI support an OpenACC pragma for Fortran to parallelize a vector operation? I.e. let's say I have two arrays a and b and I have a fortran code of <code>""c=a+b""</code> , can I put an Openacc comment to parallelize that operation or do I need to re-write it into a loop?	You can put array-syntax within a <code>!"\$ACC KERNELS"</code> region and the PGI compiler will automatically offload the implicit loop to the target device.
30	What compilers can be used to compile an OpenACC code?	PGI, Pathscale, Cray, GNU 6.x (though GNU support is still limited and in development)
31	Will we have to install any program or complete any exercises until next week?	You can either take a Qwicklab ( <a href="https://nvidia.qwicklab.com/">https://nvidia.qwicklab.com/</a> ) which does not require additional software
32	What exactly is quicklab?	qwiklabs offer access to OpenACC code samples for you to practice. qwikLAB instances run on Amazon AWS
33	Can I do deep copy of a nested structure? For example, a SOA inside of a structure.	You can but it's not done automatically.
34	How to mix OpenACC and CUDA? An example?	OpenACC and CUDA are interoperable. There are examples online and included with the PGI compilers.
35	Is it possible to query device info and use that to parameterize worker num and vector length dynamically?	You can query the memory size, but there not a way to get the max vector or worker size. Though, dynamically setting the worker or vector size probably have limited value. The compiler usually does a good job on selecting the optimal schedule,
36	Is there is Java library for OpenACC?	No. Currently only Fortran, C, and C++ are supported.
37	So is Java is not a better language when coming to program GPU?	Java requires runtime JIT compilation while GPU code requires the compiler to create heterogeneous code for both the host and device. As far as I am aware of, there's no JIT which can automatically divide the code between host and device.
38	What hardware is OpenACC optimized for?	OpenACC runs on multiple architectures - x86 and POWER CPUs, NVIDIA and AMD GPUs
39	Is it possible to set up OpenACC with Qt C/C++ IDE on Linux?	OpenACC is a feature of the compiler. Hence if you can configure Qt to use a compiler that supports OpenACC, then you should be able use OpenACC from within Qt. However, Qt won't have any knowledge of OpenACC.

# INTRODUCTION TO OPENACC COURSE

## Lecture #2 Q&A



40	Are there other IDEs with knowledge of OpenAcc and by knowledge do you mean some integrated features or syntax completion?	None that I'm aware of.
41	Qt ide has GUI libs for mingw, msvc etc , But if custom compiler is it likely that rest of the dll files cannot work together?	Possible. However, I'm not a Qt user so don't know what issue you may encounter.
42	Is there loop or code that is too simple and thus unefficient to run via OpenAcc on gpu or too complex and long to run on gpu or inefficient?	Sure. You want enough compute to keep the accelerator busy. For bigger code segments, you may run into register pressure or other hardware constraints. The specifics will depend on the target device and code.
43	For matrix to sum values it makes sense to run parallel on GPU, but if longer or more complex loop would there be some limit where it becomes inefficient to run on gpu?	The short answer is yes, but I've parallelized loops with thousands of lines and still achieved efficient code.
44	Is a non trail version of pgprof obtainable for personal use?	For license questions please email to <a href="mailto:openacc@nvidia.com">openacc@nvidia.com</a> and we will be able to give a detailed answer. Regarding the profiler - you can also use the NVPROF which is a part of the CUDA Toolkit.
45	Is there a reason to use pgprof over nvprof ?	No. As of the PGI 2016 release, pgprof uses nvprof under the hood.
46	When is finalize needed?	There is a reference counter for each unstructured data region. Each time the data region is entered the counter is incremented, and decremented each time it's exited. The data is not deleted until there are no more references.
47	What is the difference between the `acc parallel loop` and `acc loop` pragmas? Are `parallel` and `loop` separate pragmas that can be stacked?	Correct. They are separate directives that can be combined.
48	Is NVIDIA planning to "extend" the OpenACC to c# language in the future?	not in the near future
49	Will Intel compiler ever have openacc support? How do we compare/evaluate when we should use OpenACC vs. OpenMP vs. OpenCL etc.?	I do not believe Intel has plans to support OpenACC, however please ask them directly.
50	If I have 3 Tesla-X's, will all be used automatically be OpenACC? Or do I have to put specific code for each?	Jonathan, not sure I am familiar with Tesla-X, but OpenACC should work on pretty much all NVIDIA GPUs
51	(Tesla-X -> Titan X)	Titan board is a GeForce card, Tesla is a different family of products designed for data centers
52	Does any of the optimizations seen today yield any improvement when the compiler targets multi CPU?	Yes. The optimization will apply to all accelerator targets including multi-core CPUs.
53	When running an OpenACC code on Tegra X1 in which both CPU and GPU share the same memory, do you still need to copy data?	PGI OpenACC is not available on ARM yet

# INTRODUCTION TO OPENACC COURSE

## Lecture #2 Q&A



54	What if a data structure changes size such as size grows in the cpu, can you still use structured data region directives? Does the data size change accordingly in the device?	This is not done automatically unless you are using managed memory. If you change your data structure on the host, you will need make the same changes on the device.
55	Do you mean the acc update device could only update the value not the data structure?	The data clauses perform a shallow copy of the structure. If the stucture contains pointers, the pointer itself will be copied but not the object it points to.
56	Why OpenACC is not freely available for windows as cuda, openmp, mpi and will it be freely available anywhere in future?	I can only speak to PGI's availability. The problem there is we have to distribute some Microsoft components that are not free to us. We're trying with Microsoft to find a way to make these bit available for free as well. No promises, but hopefully.
57	What's difference between #pragma acc copyin(...) versus #pragma acc data copyin(...)	pragma acc copyin"" isn't a valid directive. Did you mean ""pragma acc kernels copyin""? If so, they both copy the data to the device, but the lifetime of the data will be different. Paired with a compute region, the data lifetime is only within that
58	I found and watched first lecture video from 2015 year. Is it similar to this year first video?	Lectures from the last year are similar, but we have updated the content for 2016
59	Is gcc compiler supports openacc directives?	GNU 6.x does have limited support for OpenACC with full support still in development
60	Does openacc use cuda streams - when nvidia gpu is accelerator - to asynchronously copy? Furthermore can it split up a big copy automatically to start working quickly on large data?	Yes. This is the "async" clause. It will be covered next week.
61	Is there a way to print out an overview of your hardware with information like # of SMX's and max vector size per SMX?	PGI has a utility called "pgaccelinfo" which will show this.
62	What if size of memory we are attempting to copy in is larger than GPU memory?	The program will get a runtime error stating that you are out of memory.
63	Can I try OpenACC compilation free on Windows?	The PGI compilers for Windows does come with a 15 day free trial, but does require a for fee license to continue after that.
64	What is the difference between OpenACC, OpenCL, OpenMP?	OpenACC and OpenMP are directives-based models. OpenCL is a low level programming language. OpenACC has been designed for paralel programming and works across multiple architectures.