Training «Massively parallel computations, architecture and CUDA programming model + OpenACC»

Riga Technical University, March 25-28

March 25 (Mon)

10:00-11:30	Architecture and programming of massively parallel systems: performance and parallelism. GPU evolution. SIMD and SIMT
11:30-12:00	Coffee break
12:00-13:30	CUDA programming model: CUDA 'Hello world' example. Main principles. Blocks and threads. Data exchange between GPU and host. Error handling
13:30-14:30	Lunch break
14.30-16.00	Hands-on: configuring the system, introduction to CUDA programming
16:00-17:00	Q&A

March 26 (Tue)

10:00-11:30	Overview of memory hierarchy in CUDA. Register file, constant memory. Global memory. Shared memory. Texture memory. Standard algorithm implementation on CUDA: matrix multiplication, reduction
11:30-12:00	Coffee break
12:00-13:30	Thrust library: Linear transformations and functors. Placeholders and tuples. Performance. CUDA/C interoperation. CUSP library
13:30-14:30	Lunch break
14.30-16.00	Hands-on: memory hierarchy, Thrust, CUSP.
16:00-17:00	Q&A

March 27 (Wen)

10:00-11:30	GPU-accelerated libraries: CURAND, CUBLAS, CUSPARSE, CUFFT, MAGMA
11:30-12:00	Coffee break
12:00-13:30	Hands-on: libraries
13:30-14:30	Lunch break
14.30-16.00	Hands-on: Debugging and profiling: usage, key principles; gdb and cuda-gdb, cuda-memcheck; CUDA-profiler. session
16:00-17:00	Q&A

March 28 (Thu)

10:00-11:30	Multi-GPU systems: programming and debugging. Hybrid systems, NUMA-systems.Device context. MPI. POSIX-threads. OpenMP. CUDA Events
11:30-12:00	Coffee break
12:00-13:30	Fast development on GPU using directives. OpenACC and PGI compiler. Basic directives and examples, data localization. Kernel configuration and parallelization of loops. Profiler and collecting of execution characteristics
13:30-14:30	Lunch break
14.30-16.00	Hands-on: Multi-GPU, OpenACC
16:00-17:00	Q&A, Summing up