

Practical Exercise #12

GPU Languages: Back to Basics

Exercises

1. Change the host memory allocation in the CUDA streaming triad example to use pinned memory (Listings 12.1–12.6). Did you see any performance improvement?
2. In the sum reduction example, try an array of 18,000 elements, all initially assigned an index value. Run the CUDA source code, then the version in SumReductionRevealed. You may want to adjust the amount of output.
3. Convert the CUDA reduction example to HIP by HIPifying it.
4. For the SYCL example in Listing 12.20, initialize arrays a and b on the GPU device.
5. Convert the two initialization loops in the RAJA example in Listing 12.24 to the Raja:forall syntax. Try running the CUDA example.