ICNPG 2023

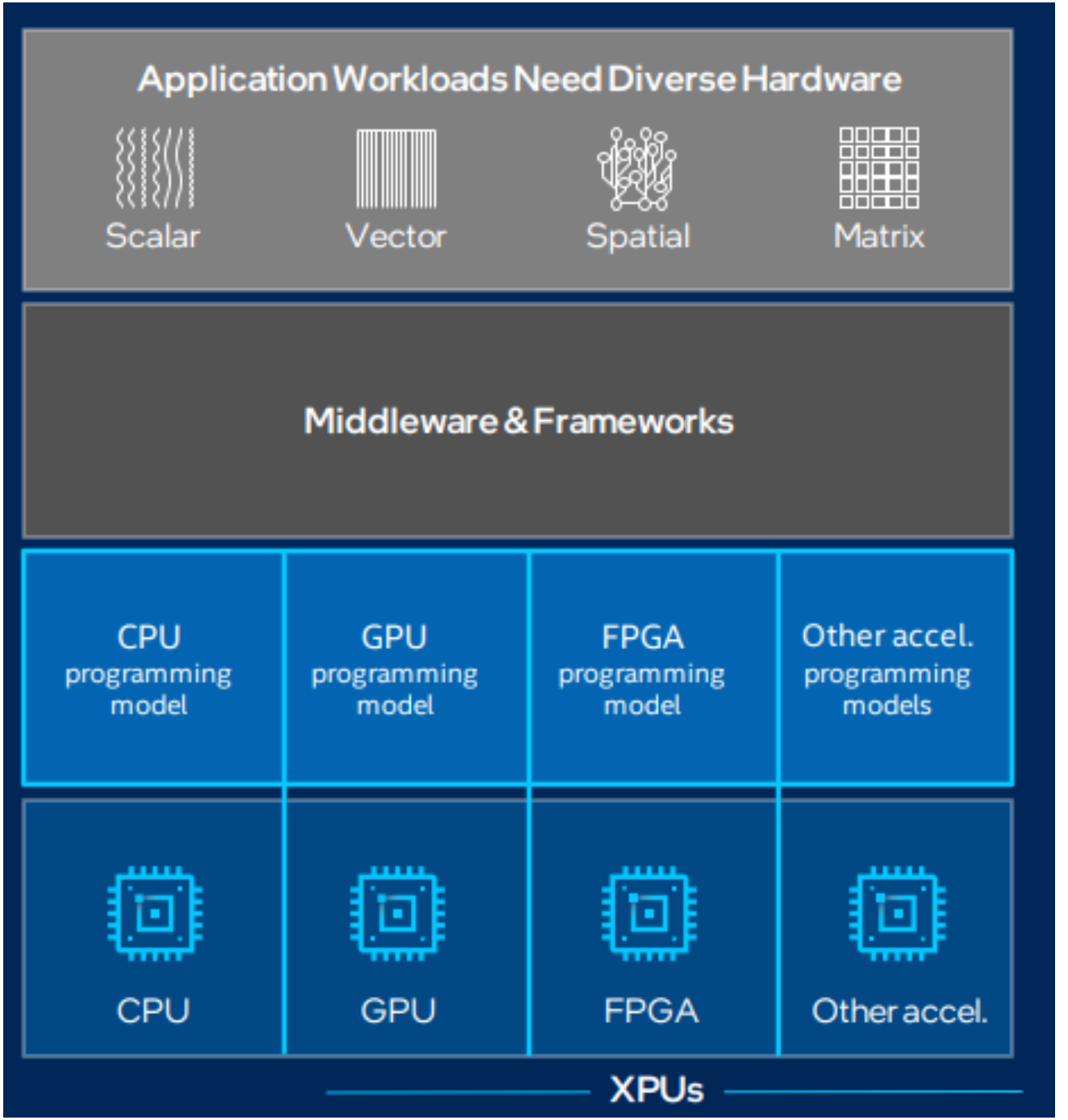
Clase 14: SYCL

Comentarios sobre la clase 14: SYCL
No subieron el power point a tiempo

SYCL es un lenguaje actual que sirve para cualquier placa gráfica. Tiene una estructura distinta a CUDA. Como CUDA C++, se trabaja con SYCL como una librería de C++.

Hay un cluster Intel DevCloud para programar. Solo hay que registrarse. Pero al usar las placas Intel no podemos usar CUDA C++, sino que tenemos que usar SYCL. Tmb hay otras librerías de alto nivel y librerías de CUDA C++ porteadas para que funcionen con Intel.

Tmb estpa Saturn Cloud que al inicio es gratis pero luego hay que pagar


Application Workloads Need Diverse Hardware
Scalar
Vector
Spatial
Matrix
Middleware & Frameworks
CPU programming model
GPU programming model
FPGA programming model
Other accel. programming models
CPU
GPU
FPGA
Other accel.
XPUs

SYCL – Khronos standard for heterogeneous computing

Template library specification

C++ with SYCL:

- Pick a device
  - Binds a queue
- Share data
  - Unified shared memory (USM) or buffers
  - Implicit and explicit data transfers
- Offload computation
  - Submit command groups to the queue
  - Inorder and out-of-order (DAG) scheduling

CUDA to SYCL dictionary

CUDA	SYCL
Block	Work group
Thread	Work item
Grid	ND-range
Kernel	Command group
CUDA Stream	Queue
Shared memory	Local memory
Cooperative groups	Subgroups
Unified memory	Unified shared memory(USM)
Graphs	tf::syclflow in Taskflow

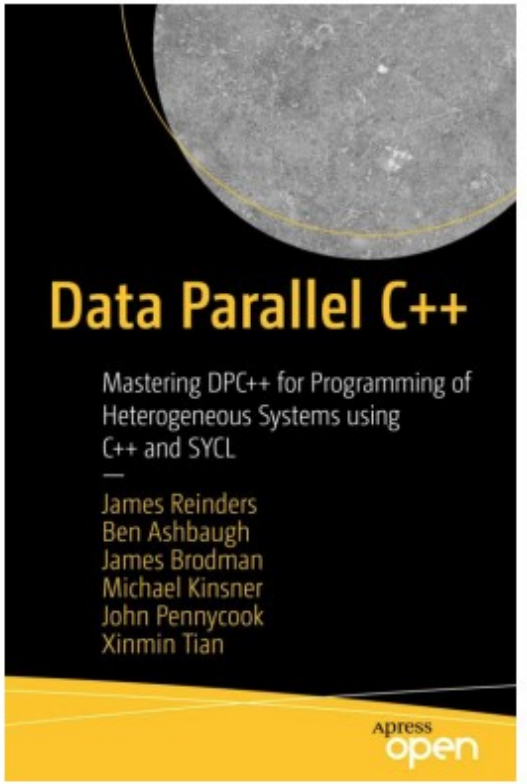
Simple SYCL program

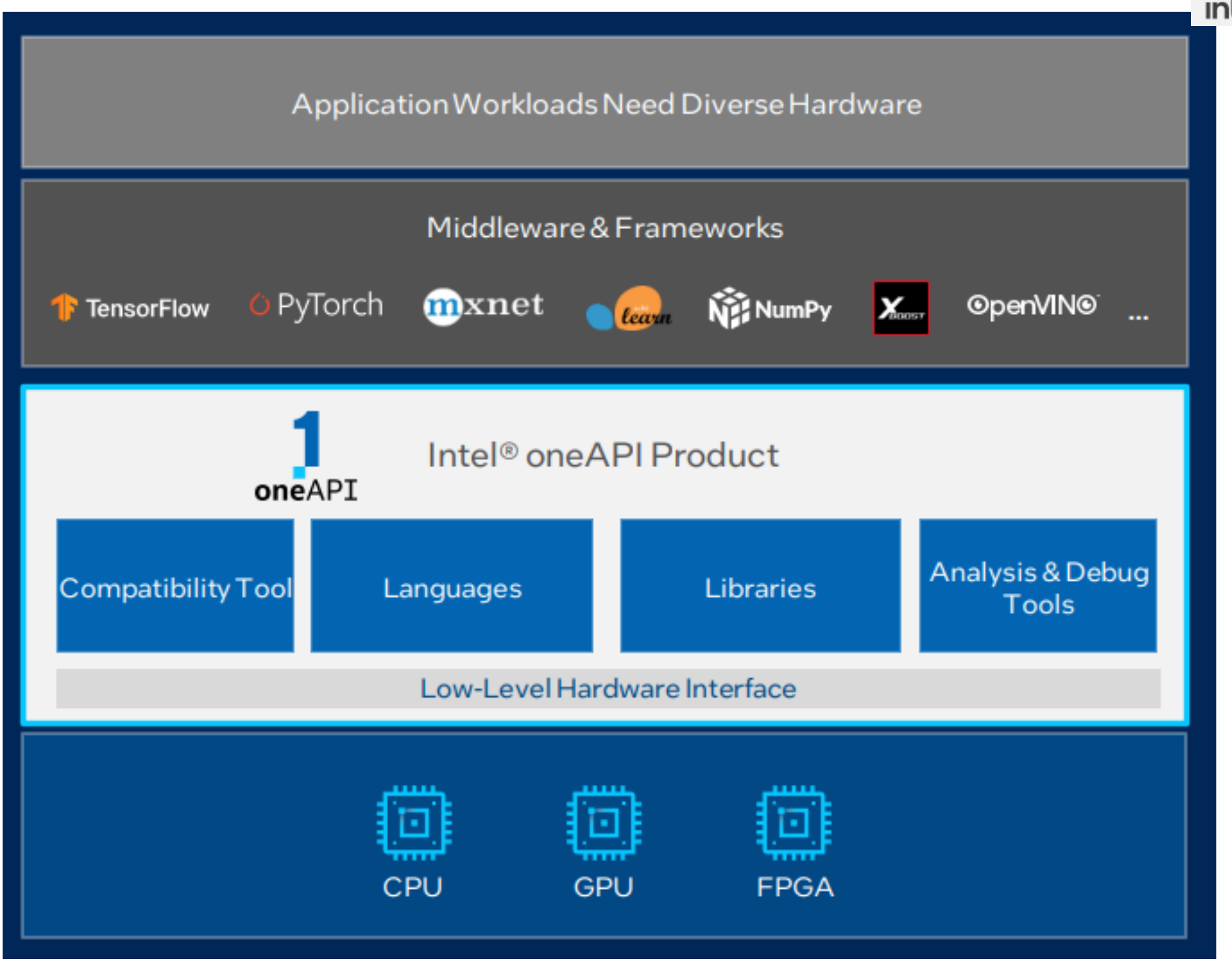
```
#include <CL/sycl.hpp>
#include <array>
#include <iostream>
using namespace sycl;

int main() {
  constexpr int size=16;
  std::array<int, size> data;

  // Create queue on implementation-chosen default device
  queue Q;

  // Create buffer using host allocated "data" array
  buffer B { data };

  Q.submit([&](handler& h) {
    accessor A{B, h};
    h.parallel_for(size , [=](auto& idx) {
      A[idx] = idx;
    });
  });

  // Obtain access to buffer on the host
  // Will wait for device kernel to execute to generate data
  host_accessor A{B};
  for (int i = 0; i < size; i++)
    std::cout << "data[" << i << "] = " << A[i] << "\n";

  return 0;
}
```

Host code

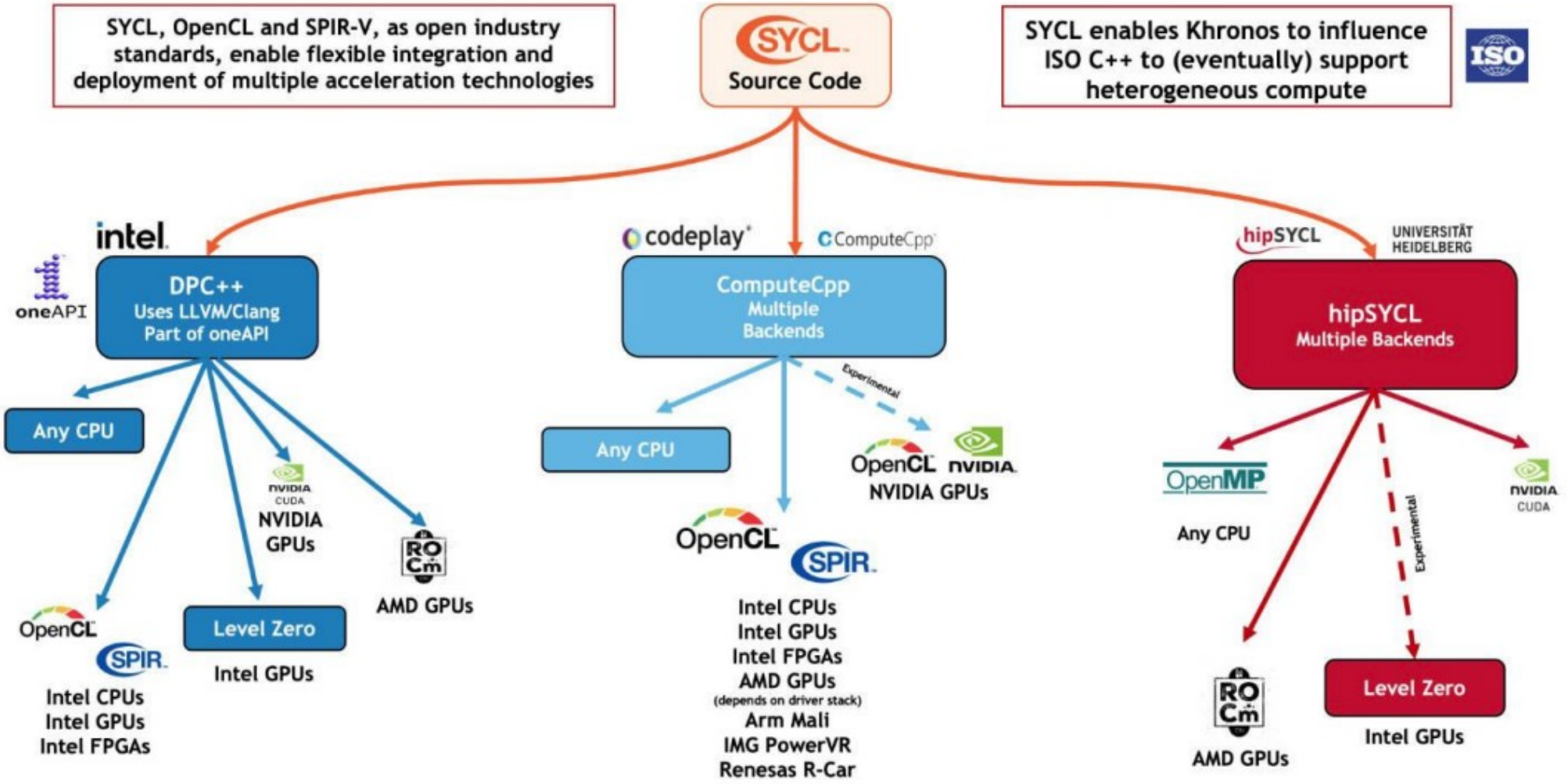
Device code

Host code




Application Workloads Need Diverse Hardware
Middleware & Frameworks
TensorFlow
PyTorch
mxnet
learn
NumPy
XGBOOST
OpenVINO
...
oneAPI
Intel® oneAPI Product
Compatibility Tool
Languages
Libraries
Analysis & Debug Tools
Low-Level Hardware Interface
CPU
GPU
FPGA

SYCL is gaining traction



https://www.khronos.org/sycl/