Figure 6.3: An example of the grid/block hierarchy used to describe a set of threads that will be spawned by CUDA. The figure illustrates a 4x3x2 grid made by 3x2 blocks. The grid connections are there only for illustrative purposes.
Each grid point = thread

OpenCL
Thread
Geometry

getGlobalSize(dim) = gridDim x blockDim

globalSize(dim) = blockDim

Each grid point = thread

⇒ Work Group