zsolt wrote on Sat, 15 February 2025 23:08I'm absolutely not an expert, but AFAIK
nvidia-cuda-toolkit on Ubuntu can be used in Clang or G++ projects.

In Linux yes. In Win32, no way - you need to use Visual C++. At least that is Nvidia's position.
https://stackoverflow.com/questions/68726422/is-it-possible-
to-integrate-nvcc-with-gcc-g-in-windows

Now nvcc seems to work by separating cuda code and host code, then compiles cuda code and
creates a new c++ file which combines host code and compiled cuda code (as arrays like

unsigned long long fatbinData[]= {0x00100001ba55ed50ull ....
)

Then this file is compiled by host compiler. There are explicit checks in CUDA headers for MSC++
compiler (#ifdef MSC_VER). There might also be a problem with CUDA runtime library, but I
doubt that...

I believe that if I take clang-cl and give that to nvcc, with some additional trucks like -D MSC_VER
it should work. Do not have time right nor energy now to test right now...