Further Optimization Techniques

- CGMA ratio :

  https://www.sciencedirect.com/topics/computer-science/global-memory-access

- Bank conflict :

  https://blog.csdn.net/Bruce_0712/article/details/65447608

- CUDA streams:

  https://developer.download.nvidia.com/CUDA/training/StreamsAndConcurrencyWebinar.pdf