NVIDIA Visual Profiler

File View Window Run Help

Analysis

Export PDF Report

1. CUDA Application Analysis
2. Performance-Critical Kernels
3. Compute, Bandw... or Latency Bound

The first step in analyzing an individual kernel is to determine if the performance of the kernel is bounded by computation, memory bandwidth, or instruction/memory latency. The results at right indicate that the performance of kernel "cuContactsSMEM" is most likely limited by compute.

Perform Compute Analysis

The most likely bottleneck to performance for this kernel is compute so you should first perform compute analysis to determine how it is limiting performance.

Perform Latency Analysis

Perform Memory Bandwidth Analysi

Instruction and memory latency and memory bandwidth are likely not the primary performance bottlenecks for this kernel, but you may still want to perform those analyses.

Rerun Analysis

If you modify the kernel you need to rerun your

Results

Kernel Performance Is Bound By Compute

For device "GRID K520" the kernel's memory utilization is significantly lower than its compute utilization. These utilization lev
indicate that the performance of the kernel is most likely being limited by computation on the SMs.