* Zufällig gefüllte Matrizen

**Benchmark-Vergleiche CPU-GPU:**

Steigende Matrizenzahlen

* Ab welcher Matrizengröße wird GPU schneller als CPU?

Quadratische Matrizen gegeneinander vergleichen

Nicht-quadratische Matrizen gegeneinander vergleichen

* Gibt es Unterschiede zwischen Benchmarks von quadratischen und nichtquadratischen Matrizen oder ist Speedup der GPU bei beiden Varianten gleich?

32,64,128,256,1024,4096,… (durch Warp-Größe/32 teilbare Größen der Matrizen) gegeneinander vergleichen

Nicht durch 32 teilbare Matrizen gegeneinander vergleichen

* Ist der SpeedUp bei Matrizen, deren Größe durch 32 teilbar ist, größer als bei Matrizen, deren Größe nicht durch 32 teilbar ist?

Rechte Matrix transponieren und Zeile mit Zeile statt Zeile mit Spalte multiplizieren.

* Wie hoch ist der SpeedUp?

In einem Thread nur eine Matrixstelle ausrechnen lassen gegenüber mehreren (8,16,32,64,128,…)

Die Multiplikation sollte zunächst schneller werden, aber irgendwann wieder abnehmen, wenn es zu seriell wird.

* Welche Anzahl an Stellen, die im selben Thread ausgerechnet werden, sind optimal? Wie hoch ist der SpeedUp?

Vergleich mit Amdahl’s Law und mit dem WorkSpan-Model

* Inwiefern treffen die Aussagen dieser Modelle zu?

Shared Memory(SM) als L1 konfiguriert nutzen und vergleichen mit der Nutzung manuell programmiert (GPU interner Vergleich)

* Wie viel schneller ist der manuell programmierte SM?