



Vom Videospiegel zum Supercomputer

Informatik. Die Hardware moderner Grafikkarten ist extrem leistungsfähig, aber unflexibel. Mit spezieller Programmierung wollen Wissenschaftler ihr Potenzial für neue Anwendungen nutzen.

VON WOLFGANG DÄUBLE

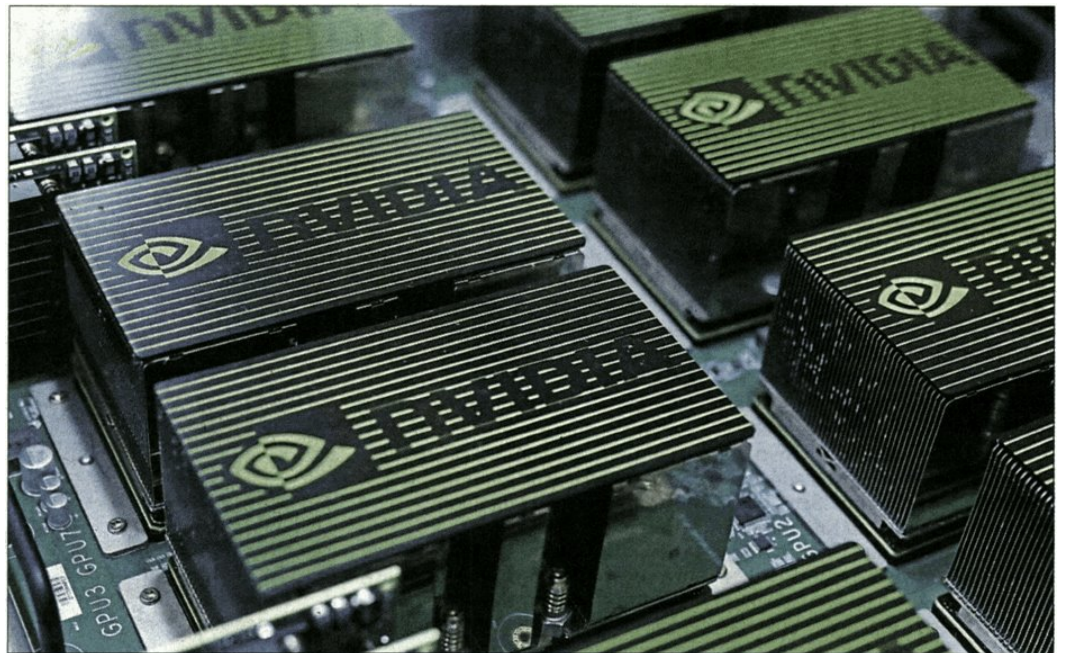
Die Sonne steht tief über dem Horizont, ihre Strahlen brechen sich in einem sanft plätschernden Bach, der sich durch dicht bewaldete Hügel schlängelt. Ein leichter Wind bewegt die tief über das Ufer hängenden Äste und trägt gelegentlich ein Blatt durch die Luft, weiter hinten springt ein kleiner Fisch aus dem glitzernden Wasser.

Wer glaubt, solche Szenen nur auf Wanderausflügen oder in Naturdokumentationen beobachten zu können, hat die Animationen aktueller 3-D-Videospiele wie „Far Cry 5“ oder „Metro Exodus“ noch nicht erlebt – auch wenn es in den Schieß-Spielen meist weniger harmonisch zugeht. Mit unglaublichem Detailreichtum, realistischer Physik und einer filmreifen Beleuchtung übertrumpfen sich die Neuerscheinungen aus der Videospielebranche gegenseitig, die computergenerierten Welten sind von echten Aufnahmen kaum mehr zu unterscheiden. Es scheint nur logisch, dass auch deren Umsatz längst mit der globalen Filmindustrie mithalten kann.

Tausende Kerne, ein Befehl

Um die bildgewaltigen Szenarien voll auszukosten, braucht es entsprechend leistungsfähige Grafikkarten, deren Rechenarbeit von den sogenannten GPUs (siehe Lexikon) geleistet wird. Hier hat sich in den vergangenen Jahren ein rasanter Fortschritt vollzogen, sagt Dieter Schmalstieg, Leiter des Instituts für Maschinelles Sehen und Darstellen der TU Graz und Forscher am Zentrum für Virtual Reality und Visualisierung (VRVis), einem Comet-Projekt des Technologieinstituts.

„GPUs sind die derzeit billigste und leistungsfähigste Hardware der Welt. Zum Vergleich: Der Hauptprozessor eines typischen PC hat normalerweise zwei, vier oder bei den teuren Geräten acht Rechenkerne. Eine GPU hat aber zwischen 2000 und 4000 Rechenkerne. Also um drei Zehnerpotenzen mehr“, rechnet der Informatiker vor. „Zumindest theoretisch



Die aufgemotzten Grafikkarten bringen für unterschiedlichste Anwendungen mehr Rechenleistung.

[Reuters]

hat man damit eine viel höhere Leistungsfähigkeit – bei vergleichbarem Stromverbrauch, der mittlerweile ein wesentlicher Kostenfaktor ist.“

Seit einiger Zeit werden die Grafikkarten daher zweckentfremdet, zum „Schürfen“ der digitalen Währung Bitcoin etwa, oder um die Leistung von Supercomputern zu steigern. Doch die Grafikkarten haben ein entscheidendes Manko: Ihre volle Rechenkraft können sie nur für ganz bestimmte Anwendungen entfalten. „Was GPUs normalerweise tun, ist, einen einzigen Befehl auf riesige Mengen an Daten anzuwenden“, erklärt Schmalstieg. Die vielen Tausend Rechenkerne werden dabei zusammenschaltet und arbeiten parallel an derselben Aufgabe. Jede Veränderung dieses Prinzips geht bisher auf Kosten der Effizienz und macht die Vorteile der hohen Rechenleistung schnell zunichte.

In einem vom Wissenschaftsfonds FWF finanzierten Projekt arbeitet Schmalstieg mit seinem Team daran, diese Grenzen zu durchbrechen. „Wir geben jetzt

eine Software-Schicht hinzu, die auch eine Ablaufplanung für komplexe Probleme gestattet. Dabei werden die Prozessorkerne in Gruppen zusammengefasst, die alle etwas anderes machen. Geschickt angeordnet, erhält man eine Verarbeitungskette, in der die Daten, die eine Gruppe erzeugt, genau im richtigen Moment an die nächste Gruppe weitergegeben werden – wie bei einem Fließband.“ Zwar verwendet er die vom Hersteller Nvidia entwickelte Programmiersprache Cuda, um die Prozessorkerne zu steuern, der Konzern gebe

jedoch nur ungenügend Einblick in die Details seiner Produkte. Ähnlich wie beim Hacken von Hardware muss Schmalstieg Team daher eigene Lösungen finden.

Unendliche Detailstufe

Die potenziellen Anwendungsmöglichkeiten der Grafikkarten seien aber vielfältig, ist sich der Wissenschaftler sicher. Einerseits könne man sie natürlich für grafische Software nutzen, wie etwa Animationen für Architekten, die ganze Stadtteile in praktisch unendlicher Detailstufe darstellen. Doch auch für Simulationen, die kein Bild benötigen, wie sie in der Materialfertigung oder für die Wettervorhersage üblich sind, sei seine Arbeit gedacht.

Schmalstieg gelang es auch, die GPUs in Supercomputern besser auszureizen. In diesem Bereich sieht er besonderes Potenzial: „Hier wird weniger Standard-Software benutzt, die Hemmschwelle, alternative Entwicklungsumgebungen zu verwenden, könnte daher geringer sein. Wir haben jedenfalls nachgewiesen, dass sich hier viel herausholen lässt.“

LEXIKON

Grafikkarten, nach der englischen Bezeichnung „Graphics Processing Units“, als GPUs abgekürzt, übernehmen rechenintensive Aufgaben der 2-D- und 3-D-Computergrafik und entlasten dadurch den Hauptprozessor (CPU). Sie sind nach der SIMD-Architektur (Single Instruction, Multiple Data) gebaut, die für die parallele Bearbeitung einzelner Befehle für große Datenmengen konzipiert ist. Algorithmen mit wenig Datenparallelität können GPUs daher normalerweise nicht auslasten.