

Prozessorgeflüster

Von first und second

Intel first – so lautet die Maxime von Intel-Chef Brian Krzanich. Beim Trendthema Deep Learning will Intel gleich mit fünf Architekturen an die Spitze. Bei den Desktop-PCs gibt es aber erst einmal wieder erstarkte Konkurrenz.

Von **Andreas Stiller**

Als einer der ersten bedeutenden Konzernlenker reiste Brian Krzanich ins Weiße Haus, um dem neuen Präsidenten zu huldigen und milliardenschwere Investitionen anzukündigen, natürlich innerhalb der USA. Genau das hatte allerdings schon vor einiger Zeit sein Vorgänger Otelini dem Trump-Vorgänger Obama versprochen, dann aber das weitgehend fertiggestellte Werk in Arizona mangels Absatz erst einmal auf Eis gelegt – was im heißen Arizona ja nicht so ganz einfach ist.

Nun aber will Intel die 7 Milliarden Dollar teure Fab 42 für 10- und 7-nm-Prozesstechnik wirklich fertigstellen und in Betrieb nehmen. 3000 High-Tech- und etwa 10.000 Support-Arbeitsplätze sollen hier entstehen – nachdem Intel vor nicht allzu langer Zeit angekündigt hatte, 12.000 Arbeitsplätze im Rahmen der Restrukturierung abzubauen.

Sorgen machen sich nun die großen Standorte außerhalb der USA, im irischen Leixlip und im israelischen Kiryat Gat, wo gerade erst in November ein Investitionsprogramm in Höhe von 6 Milliarden US-Dollar gestartet wurde, intensiv gefördert von der israelischen Regierung.

Ein bisschen chaotisch siehts bei Intels Roadmap aus. Noch in diesem Jahr soll laut einer auf einem Investor Meeting getroffenen Aussage der Nachfolger von Kaby Lake (7th Core Generation) erscheinen, die achte Generation mit, na ja, nicht gerade fulminanter Performance-Steigerung beim Babco Sysmark 1.5 um 15 Prozent. Aber was ist die achte Generation? Der eingeschobene Coffee Lake in alter 14-nm-Tech-

nik (inzwischen 14 nm+) mit bis zu sechs Kernen oder die ebenfalls für dieses Jahr angekündigten ersten Intel-Chips in 10 nm namens Cannon Lake für die U- und Y-Mobile-Klasse, die vielleicht eher kommen?

Richtig „first“ war Intel bei Trendthema Machine Learning, auch AI genannt, bislang gar nicht, hat das eher verschlafen. Doch nun will die Firma mit Macht diesen Markt erobern und gleich von allen Seiten mit fünf Architekturen angreifen. Die Programmierer legten sich dazu mächtig ins Zeug und konnten mit der neuen Mathematik-Kernel-Bibliothek für Deep Neural Networks (MKL-DNN) auch den Xeon mit seinen AVX-Einheiten als Basis weit besser ins Spiel bringen. Via OpenCL mitrechnen kann dabei außerdem in der Xeon-E3-Familie und bei den Desktop-Prozessoren auch die HD-Grafik.

Neudefinierte AI-Performance

Beim Xeon Phi haben die Programmierer gar die Performance von Caffe/Alexnet von „out-of-Box“ bis zu liebevoll handoptimiert um Faktor 400 beschleunigt. Und gegen Ende des Jahres soll dann noch ein speziell für AI optimierter Xeon Phi Knights Mill ins Rennen gehen, der bei halber Genauigkeit (Fp16) bis zu viermal so schnell wie die normale Ausführung sein soll.

Ebenfalls in diesem Jahr sollen zwei der in letzter Zeit vollzogenen Einkäufe

die AI-Leistung ihrer Produkte unter Beweis stellen. Alteras Arria 10 will dabei nicht nur als FPGA-Coprozessorkarte DLIA (Deep Learning Inference Accelerator), sondern auch integriert im Gehäuse mit Xeon Broadwell punkten. Das FPGA kann man für alle möglichen Aufgaben konfigurieren und unter OpenCL betreiben, eben auch als Single-Precision-AI-Rechenmaschine mit 1,5 TFlops.

Spezifisch auf den Deep Learning-Job maßgeschneidert ist hingegen die mit dem Start-up Nervana eingekaufte Nervana Engine, jetzt auf Lake Crest umgetauft. Auf einem AI-Day in München stellte ihn der Nervana-Mitgründer Naveen Rao nun auch hierzulande vor. Lake Crest kommt demnach als ein von TSMC gefertigtes ASIC in 28-nm-Technik. Es thront zusammen mit vier Stapeln à 8 GByte High Bandwidth Memory 2 (mit insgesamt 1 TByte/s Bandbreite) auf einem Interposer-Board, das auf einer PCIe-Karte steckt, ohne weiteren Hauptspeicher. Der Prozessor hat 12 Hochgeschwindigkeit-Links, ein jeder mit 100 GBit/s bidirektional, mit denen sich vielfältige Topologien aufbauen lassen. Das Konzept erinnert irgendwie an die Transputer aus den 80er Jahren.

Der Tensor-Chip besitzt intern 12 sogenannte Cluster-Prozessoren, ein jeder bestehend aus zwei MPUs mit 1024 Bit Recheneinheiten. 55 Tera-Ops/s soll der Chip laut Rao leisten können, wobei er aber nicht mit IEEE Half Float (Fp16), sondern mit nicht genauer dokumentiertem „Flexpoint“ rechnet. Mit welchen magischen Rechenricks man damit letztlich auf 55 TOPs/s kommt, bleibt offen. Intels vielsagende Antwort: „The Intel Nervana architecture will redefine the way AI workload performance will be calculated with its proprietary tensor core technology that concentrates high density compute onto Intel architecture at unprecedented levels of efficiency.“

Dieses Flexpoint-Format arbeitet intern mit Integer und soll eine variable Ge-

AMD Ryzen, Modelle und vermutliche Preise			
Modell	Kerne/Threads	Basis/Turbo	Preis
Ryzen 7 1800X	8/16	3,6/4,0 GHz	600 €
Ryzen 7 1700X	8/16	3,4/3,8 GHz	440 €
Ryzen 7 1700	8/16	3,0/3,7 GHz	355 €
Ryzen 5 1600X	6/12	3,3/3,7 GHz	275 €
Ryzen 5 1500	6/12	3,2/3,4 GHz	245 €
Ryzen 5 1400X	4/8	3,5/3,9 GHz	220 €
Ryzen 5 1300	4/8	3,3/3,6 GHz	195 €
Ryzen 3 1200X	4/4	3,4/3,8 GHz	165 €
Ryzen 3 1100	4/4	3,2/3,5 GHz	140 €

naugigkeit bieten. Letztlich, so Rao, soll es gar hinunter bis zu Single-Bit-Precision gehen können. Das ist aber erst für den für 2020 vorgesehenen Knights Crest geplant, bei dem Nervana's Tensor-Hardware auf einen Xeon-Chip wandern soll.

Das im letzten Prozessorgeflüster angesprochene ominöse Product Quality Issue, das Intels Finanzchef Swan für größere Verluste verantwortlich machte, hat sich inzwischen geklärt – nicht der Xeon Phi, auch nicht der Broadwell – nein, der Atom C2000 ist der Schuldige. Dessen dicker Bug im Low-Pin-Count-Interface war deshalb so teuer, weil so wichtige Kunden wie Cisco davon erheblich betroffen waren (siehe S. 33).

Ryzen-Preise

Und nun bekommt Intel endlich wieder echte Konkurrenz, denn der Start der Ryzen-Prozessoren von AMD steht kurz bevor. Auf ersten Händlerlisten tauchen sie schon mit ihren offiziellen Namen sowie mit Preisen in den erwarteten Bereichen

Workload optimized Hardware? Lake Crest? – Von wegen: Der Neuintelrianer Naveen Rao zeigte stattdessen lieber das Die-Plot eines achtkernigen Haswell-E.

bis maximal 600 Euro auf. Man findet die Versionen zudem mit X oder mit Pro verziert, was bei gleicher TDP auf andere Gehäusetemperaturen und Kühlerklassen hinweist. So lässt das Achtkern-Spitzenmodell Ryzen7 1800x mit 95 Watt TDP nur maximal 60 °C Gehäusetemperatur zu, wofür es eine anspruchsvollere Kühlerklasse HS81 benötigt. Demgegenüber ist der Ryzen 7 1800 Pro bei ebenfalls 95 Watt TDP für eine Gehäusetemperatur von bis zu 72,4 °C ausgelegt und erlaubt einfacher gestrickte Kühler.

Ein ziemliches Hin und Her gabs beim Thema Windows 7. Erst sicherte AMD

volle Unterstützung zu, um dann kurz danach wieder zurückzurudern: „only support and drivers for Windows 10 will be provided in AMD Ryzen desktop processor production parts“. Auch zahlreiche neue Benchmark-Ergebnisse machten die Runde. Ein Ryzen 1700X mit acht Kernen liegt danach im CPU Mark etwa auf dem Level eines Core i7 6850 mit sechs Kernen, mal besser, mal schlechter. Den Vierkerner Kaby Lake Core i7 7700K hängt er zumeist klar ab – nur bei Single-Thread-Benchmarks liegt jener um 20 Prozent voraus. Aber all das ist alles noch mit Vorsicht zu genießen. (as@ct.de)

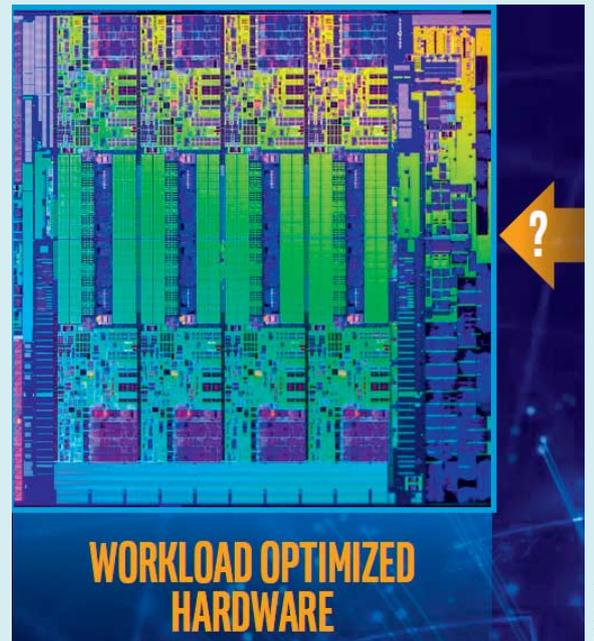


Bild: Intel

Anzeige