Christof Windeck

Uhren-Abgleich

Benchmarks zur Einschätzung der Prozessor-Performance

Wer misst, misst Mist: Dieser Spruch von Ingenieuren gilt besonders für den CPU-Leistungsvergleich mit Benchmarks. Nur wer die Eigenheiten der Testsoftware kennt, kann aus den Ergebnissen sinnvolle Schlüsse auf die Arbeitsgeschwindigkeit in der Praxis ziehen.

Seit es CPU-Benchmarks gibt, gibt es auch Streit um sie. Je nach Standpunkt zeigen sie den jeweiligen Prozessor in zu gutem oder zu schlechtem Licht. Doch es gibt keine Alternative für den Performance-Vergleich, wenn man nicht jede in Frage kommende CPU mit der eigenen Software ausprobieren kann.

Das wissen auch die Hersteller von Prozessoren und PCs. Sie wählen mit großem Geschick Messungen aus, bei denen ihre Produkte besonders glänzen. Wer weiß, was die einzelnen Benchmarks messen, kann die Daten richtig interpretieren. Auf zehn aktuellen Prozessoren für Desktop-PCs und Notebooks haben wir elf Benchmarks laufen lassen. Die Ergebnisse zeigen, welcher Benchmark welchen Prozessoren schmeichelt.

Theorie und Praxis

Das Benchmark-Ergebnis eines Prozessors lässt sich nicht 1:1 auf den Einsatz derselben CPU in einem anderen System übertragen. Schon deshalb darf man bei Benchmark-Vergleichen Unterschiede im einstelligen Prozentbereich getrost vernachlässigen; spürbar sind dermaßen winzige Leistungsunterschiede bei der praktischen Arbeit ohnehin nicht. Den Bezug zur Praxis schwächt auch die Forderung, dass ein Benchmark reproduzierbar sein muss, also bei Wiederholung unter gleichen Bedingungen annähernd gleiche Resultate liefert. Dazu muss man bei der Messung störende Einflüsse ausschalten, also parallel laufende Software, Virenscanner und Update-Funktionen stoppen oder den PC vom LAN trennen. Auf dem typischen Windows-Rechner sind jedoch viele (Hintergrund-)Programme gleichzeitig geöffnet, die zufällig Ressourcen belegen. Dadurch schwankt das von einer Anwendung nutzbare Leistungspotenzial ständig.

Außerdem spielen Komponenten wie Grafikkarte, RAM und Festplatte eine Rolle. Sie sollen beim CPU-Vergleich nicht bremsen, wir messen daher stets mit SSD und mindestens 8 GByte RAM. Geht es um den reinen CPU-Vergleich, dann beeinflussen auch andere Eigenschaften des Systems die Messungen: falsche Einstellungen im BIOS-Setup, zu hohe Umgebungstemperatur und mangelhafte Kühlung. Letzteres führt zur Drosselung der Prozessor-Taktfrequenz, also niedrigerer Rechenleistung.

Weil Software verschiedene CPU-Rechenwerke belastet und unterschiedlich ausreizt, kann ein einzelnes Benchmark-Ergebnis keine umfassende Aussage über die Arbeitsgeschwindigkeit eines Computers liefern. Es gibt

© Copyright by Heise Zeitschriften Verlag
Persönliches PDF für alle Leser der c't aus 30625 Hannover

auch ganz verschiedene Benchmark-Konzepte. CoreMark und Linpack liegen etwa im Quelltext vor und dürfen mit hoch optimierten Compilern das Beste aus der jeweiligen CPU herausholen. Andere Benchmarks arbeiten mit fertigen Programmen und führen exemplarisch bestimmte Berechnungen oder Abläufe durch. Dabei prüft etwa der AES-Benchmark von TrueCrypt 7.1a bloß einen sehr speziellen Aspekt der CPU-Performance, nämlich die Geschwindigkeit, mit der diese Verschlüsselung läuft. Schließlich gibt es Benchmarks, die Mittelwerte aus zahlreichen Messungen bilden, um viele Leistungsaspekte in einer einzigen Zahl abzubilden: Etwa der kostenpflichtige BAPCo SYSmark 2014, der mehrere kommerzielle Software-Pakete installiert und teilweise mehrere Programme parallel benutzt. Alle erwähnten Benchmarks finden Sie über den c't-Link am Ende des Artikels und dort auch eine Tabelle mit den Ergebnissen aller von uns durchgeführten Messungen.

Mancher wünscht sich einen Vergleich über viele CPU-Generationen, verschiedene Mikroarchitekturen und Betriebssysteme hinweg. Doch je breiter der Vergleich angelegt ist, desto mehr Einschränkungen schleichen sich ein. Der CoreMark beispielsweise läuft zwar etwa unter Windows, Linux, Android und iOS, liefert jedoch schon bei gleichem Prozessor und Betriebssystem deutlich unterschiedliche Resultate je nach Compiler und Optimierung. Eine uralte, unveränderte Binärversion eines Benchmarks nutzt wiederum viele Funktionen moderner Chips nicht: SSE, AVX, AES-Befehle, OpenCL- und DirectX-Beschleunigung. Allerdings gilt das auch für viele aktuelle Anwendungen, denn bei vielen Programmierern hat Performance-Optimierung nicht die höchste Priorität. Die Software-Entwicklung hinkt der Einführung neuer CPU-Funktionen oft mehrere Jahre hinterher. Noch heute gibt es viel 32-Bit- und Single-Thread-Software. Der Athlon 64 erschien aber schon 2003, der Pentium 4 mit Hyper-Threading im Jahr davor. Es ist aber zur Einschätzung einer CPU ganz entscheidend, ob die von Ihnen häufig verwendete Software bloß einen, mehrere oder alle vorhandenen CPU-Kerne nutzt: Im einen Fall ist die Single-Thread-Performance wichtiger, sonst die Multi-Thread-Rechenleistung. Grob vereinfacht kann man sagen, dass Letztere das Leistungspotenzial aller CPU-Kerne zusammen charakterisiert; die Single-Thread-Rechenleistung ist dagegen für schlecht optimierte oder nicht parallelisierbare Software wichtig.

Ob Ihre Lieblings-Software mehrere Kerne verwendet, können Sie leicht selbst testen: Starten Sie den Task-Manager und dann ein Programm, das länger an einer Aufgabe rechnet. Wenn dabei die gesamte CPU-Auslastung eines Dual-Cores wenig mehr als 50 Prozent beträgt oder bei einem Quad-Core nahe 25 Prozent liegt, ist nur ein Thread aktiv.

Dabei zeigt der Task-Manager häufig schwache Last auf mehreren Cores gleichzeitig. Auch das ist ein Indiz für Single-Thread-Software, die dann nämlich von Kern zu Kern „hüpft“. Das ist normal, weil unter Windows stets Hunderte von Threads gleichzeitig aktiv sind, die aber jeweils nur kurzzeitig Rechenleistung benötigen. Dann verdrängen sie eine andere Anwendung, sobald die einige Taktzyklen auf Daten warten muss. Das Umschalten kostet nur unwesentlich Rechenleistung: Eine moderne CPU erledigt pro Sekunde mehrere Milliarden Operationen – selbst Hunderte von Thread-Wechseln pro Sekunde spielen keine große Rolle. Der Task-Manager kann die tatsächliche Belastung der CPU nicht in Echtzeit ausweisen, sondern zeigt nur einen gleitenden Mittelwert aus den letzten paar Millionen Taktzyklen.

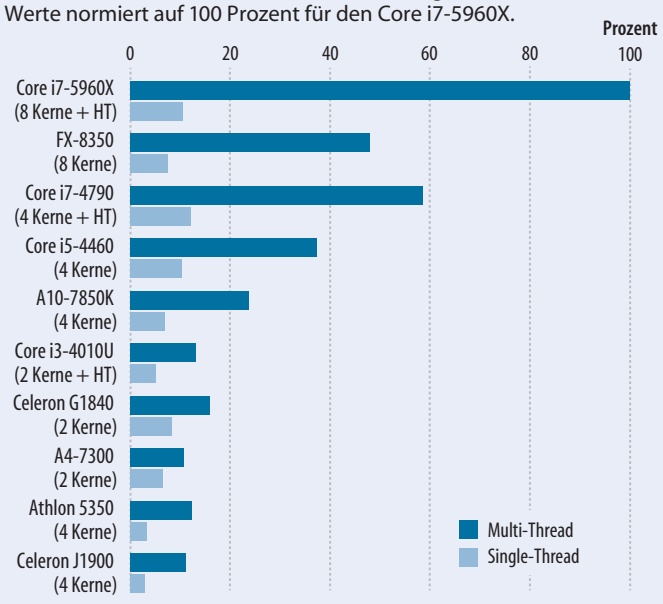
Kurzvergleich

Die Benchmarks Cinebench R15, CoreMark, kcbench und Google Octane sind im Wesentlichen von der CPU abhängig und werden wenig von RAM, Festplatte und GPU beeinflusst. Cinebench simuliert die Arbeit mit dem 3D-Renderer Maxon Cinema 4D. Er führt auch GPU-Tests durch, die wir hier aber nicht berücksichtigen. Octane lässt sich mit dem Webbrowser aufrufen, also auch von vielen Betriebssystemen aus; das Resultat hängt aber stark von der JavaScript-Engine des Browsers ab, schwankt also je nach Browser(-Version). Der kcbench läuft nur unter Linux.

Cinebench nutzt viele Kerne gut aus und erlaubt einen einfachen Vergleich zwischen Single-

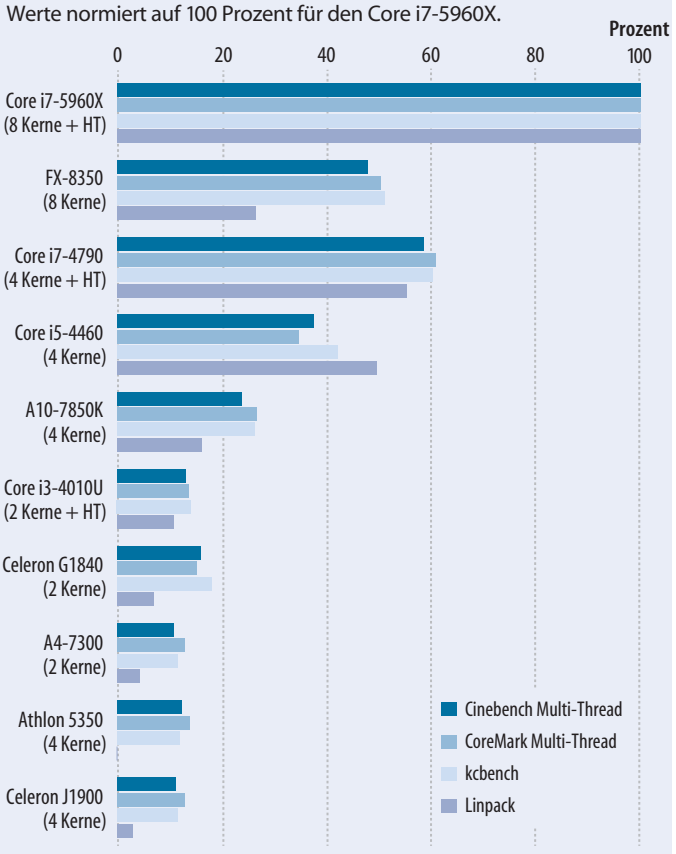
Cinebench R15: Single- vs. Multi-Threading

Der Rendering-Benchmark Cinebench nutzt sämtliche CPU-Kerne und Hyper-Threading gut aus. Im Single-Thread-Lauf zündet bei Turbo-Prozessoren der Beschleuniger.
Werte normiert auf 100 Prozent für den Core i7-5960X.



Multi-Threading-Benchmarks

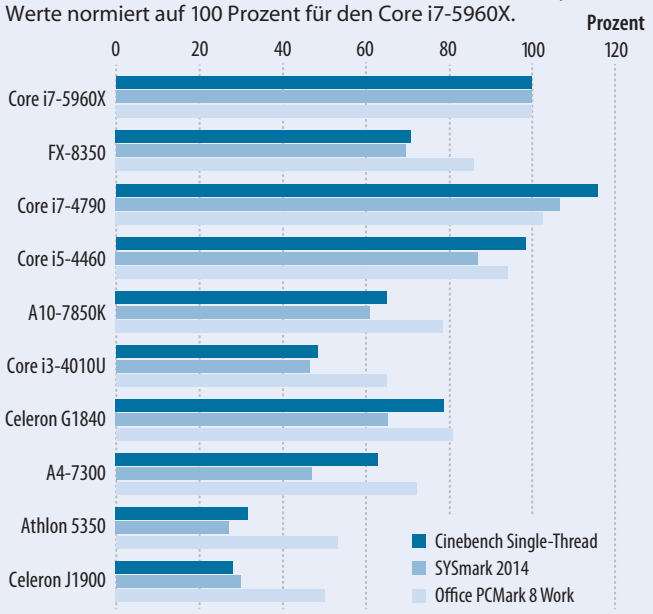
Der Compiler-Benchmark kcbench und der einfach gestrickte CoreMark verhalten sich ähnlich wie der Cinebench; Linpack verlangt hohe Dual-Precision-Gleitkommaleistung.
Werte normiert auf 100 Prozent für den Core i7-5960X.



© Copyright by Heise Zeitschriften Verlag
Persönliches PDF für alle Leser der c't aus 30625 Hannover

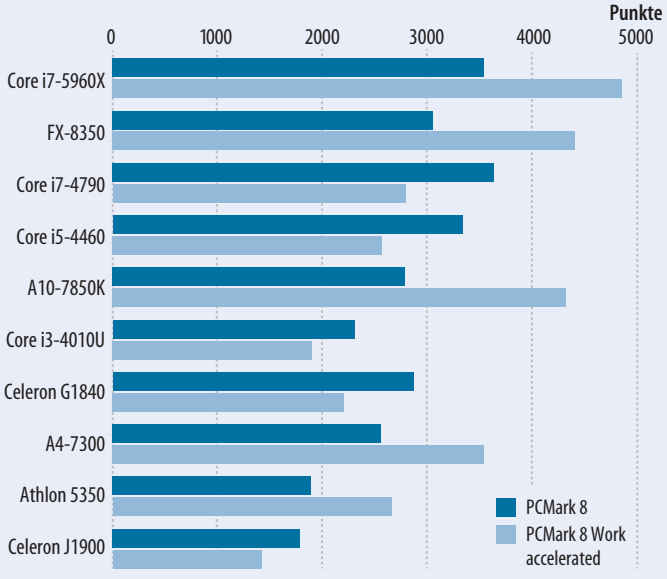
Office-Benchmarks

Bei vielen Prozessoren verhält sich der Office-Teil des SYSmark 2014 ähnlich wie der Single-Thread-Teil des Cinebench R15. Im PCMark 8 Work erreicht der Celeron J1900 die halbe Punktzahl eines 1000-Euro-Chips. Werte normiert auf 100 Prozent für den Core i7-5960X.



OpenCL fürs Office

Der PCMark 8 v2 Work erlaubt „Accelerated"-Durchläufe, die OpenCL etwa in LibreOffice Calc nutzen – bloß werden Intel-Prozessoren dann um bis zu 23 Prozent langsamer. Die AMD-CPUs legen dagegen um 38 bis 55 Prozent zu. Werte für Core i7-5960X und AMD FX jeweils mit GeForce 750 Ti gemessen.



und Multi-Thread-Rechenleistung. Octane liefert auf allen im Test vertretenen Prozessoren Ergebnisse, die annähernd proportional sind zu den Single-Thread-Resultaten im Cinebench – JavaScript läuft vorwiegend auf einem einzigen CPU-Kern.

In der Multi-Threading-Disziplin verhalten sich Cinebench, Coremark und kcbench ganz ähnlich. Das spricht gegen die These, der Cinebench bevorzuge Intel-Prozessoren. Der Luxmark (siehe unten) gewinnt sogar noch stärker durch Hyper-Threading, was nur die Intel-Chips beherrschen. Der Supercomputer-Benchmark Linpack verarbeitet Gleitkommazahlen doppelter Genauigkeit und profitiert deshalb deutlich von der höheren Zahl an AVX-Rechenwerken in den Haswell-Prozessorkernen. Wir nutzen für den Linpack jeweils optimierten Code von AMD oder Intel – so hat jeder die Chance, seine Produkte optimal in Szene zu setzen.

Bleibt man innerhalb einer Prozessorfamilie, lässt sich die Performance ziemlich gut anhand des Cinebench vergleichen. In den Tabellen hier klappt das beispielsweise mit den Haswell-Typen Celeron G1840, Core i3, Core i5 und Core i7. So liefert der 35-Euro-Celeron beispielsweise rund 80 Prozent der Single-Thread-Leistung des Core i7-5960X, der satte 950 Euro kostet. Ein Einzelkern des Core i7-4790 ist dank 4-GHz-Turbo noch schneller.

Vergleiche zwischen unterschiedlichen Mikroarchitekturen sind schwieriger, wie Celeron J1900 und Celeron G1840 illustrieren. Trotz gleichen „Vornamens" steckt in den beiden Chips völlig unterschiedliche Technik – und obwohl der G1840 nur halb so viele Kerne hat, ist er viel schneller. In ihm steckt die potente Haswell-Technik, allerdings blockiert Intel AVX und AES-Befehle. Der J1900 ist dagegen eng verwandt mit dem Atom Z3700 für Tablets: Diese Bay-Trail-Chips sind auf hohe Integration und billige Fertigung getrimmt statt auf maximale Performance. Die einzelnen Kerne haben weniger und einfacher gestrickte Rechenwerke, ein L3-Cache fehlt, es gibt nur einen RAM-Kanal. Die Single-Thread-Performance des Celeron J1900 ist deshalb miserabel. Ähnlich gelagert sind die Unterschiede zwischen den AMD-Mikroarchitekturen Steamroller (Bulldozer) und Jaguar, siehe Tabelle auf Seite 116. Der A4-7300 mit zwei Steamroller-Kernen liefert beispielsweise ähnliche Multi-Threading-Performance wie der Athlon 5350, in dem vier Jaguar-Kerne stecken. Für Single-Thread-Software eignet sich also der A4-7300 besser.

Bei Prozessoren mit kräftiger Turbo-Funktion entspricht der Multi-Thread-Wert im Cinebench nicht genau dem Produkt aus Kern-Anzahl und Single-Thread-Punktzahl: Üblicherweise takten alle Kerne zusammen deutlich niedriger als einer alleine. Der Turbo beschleunigt vor allem Single-Thread-Software.

GPU-Beschleuniger

In vielen aktuellen Prozessoren stecken integrierte Grafikprozessoren (IGPs), also GPU-Kerne. Deren Eigenschaften spielen in der Werbung der jeweiligen Hersteller eine große Rolle. Sie lassen sich für allgemeine Rechenaufgaben nutzen (GPGPU-Computing), etwa mit OpenCL-Code – also Software, die speziell dafür programmiert wurde. OpenCL-Code läuft aber auch auf CPU-Kernen.

LuxMark verwendet die Rendering-Software LuxRender, die mit OpenCL-Code arbeitet. Damit erlaubt LuxMark einen Vergleich der (OpenCL-)Rechenleistung der CPU- und GPU-Teile von Kombiprozessoren. Er kann auch beide gleichzeitig einspannen und funktioniert auf Systemen mit separater Grafikkarte ebenfalls.

Damit OpenCL-Code läuft, muss ein OpenCL-tauglicher Grafiktreiber installiert sein. Solche stellen AMD, Intel und Nvidia bereit. In den OpenCL-Grafiktreibern von AMD und Intel sind dabei auch OpenCL-Treiber für Prozessorkerne enthalten. Notiz am Rande: Intel-CPUs schneiden mit dem OpenCL-CPU-Treiber von AMD im LuxMark etwas besser ab als mit dem Intel-Treiber – aber den AMD-Treiber kann man nur installieren, wenn eine AMD-Grafikkarte im System steckt. LuxMark zeigt offenbar eher die Single-Precision-(SP-)Gleitkommaleistung, was für OpenCL auch typisch ist: Grafikprozessoren, die sich mit OpenCL als Rechenwerke einspannen lassen, liefern meistens sehr viel mehr SP- als DP-Gigaflops (GFlops).

Die eigentliche Domäne von Grafikprozessoren sind Spiele, doch dafür sind IPGs oft zu schlapp. Ob eine GPU in einem Benchmark 4 oder 8 Bilder pro Sekunde (fps) liefert, ist unerheblich: Beides ist viel zu wenig. Wichtig ist, dass das Lieblingsspiel in der gewünschten Auflösung und Detailstufe im Durchschnitt mit mindestens 30 fps läuft und auch komplexe Szenen nicht mit 15 fps stottern. Unter dieser Voraussetzung muss man auch die häufig genannten Ergebnisse der 3DMark-Benchmarks der Firma Futuremark betrachten. Im 3DMark FireStrike

© Copyright by Heise Zeitschriften Verlag
Persönliches PDF für alle Leser der c't aus 30625 Hannover

beispielsweise sind Werte unter 1500 Punkten indiskutabel, denn dabei läuft die erste der dabei vermessenen DirectX-11-Spielszenen mit weniger als 8 fps – unspielbar. Ein Vergleich auf Kellerniveau von beispielsweise Core i5-4460 (740 Punkte) mit Athlon 5350 (384) ist sinnlos, denn die Aussage lautet stets: zu lahm.

Doch die integrierte GPU kommt durchaus sinnvoll zum Einsatz, und zwar jenseits von OpenCL und 3D-Beschleunigung: Aktuelle Web-Browser, für manche Funktionen auch PowerPoint oder Adobe Reader verwenden OpenGL oder DirectX, um Grafiken und Text-Fonts anzuzeigen. Viele Videoplayer, etwa Adobe Flash oder der HTML5-Player in Google Chrome, überlassen große Teile des Video-Decodings der GPU. Bei derartigen Aufgaben gilt oft: Schneller ist nicht unbedingt besser, weil ein Video nicht flüssiger als flüssig laufen kann. Ein Vorteil ergibt sich erst in besonderen Nutzungsfällen, etwa bei Videokonferenzen mit mehreren HD-Streams oder beim Videoschnitt. Dann aber sind Unterschiede von 10 oder auch 30 Prozent in der Videobeschleunigung irrelevant, erst Faktor 2 oder 4 würden einen anderen Chip attraktiver machen.

Ein Schlaglicht auf den Praxisbezug wirft auch die AES-Verschlüsselung mit TrueCrypt 7.1a. Der TrueCrypt-Benchmark zeigt große Unterschiede und Datentransferraten bis fast 8 GByte/s – lässt aber außer acht, dass ein real existierender PC die Daten auch auf SSD oder Festplatte schreiben oder von dort lesen muss. Jenseits von 500 MByte/s bringt AES-Beschleunigung folglich nur in speziellen Nutzungsfällen deutlich spürbare Vorteile.

Voll ausgereizt

Benchmarks wie Linpack und LuxMark lassen Rückschlüsse darauf zu, wie gut sich theoretische Rechenleistung praktisch nutzen lässt. Wenn schon ein hoch optimierter Benchmark die beworbene Rechenkraft nicht ausreizt, dann sieht es in der Praxis eher noch trüber aus.

LuxMark entlarvt das simple Aufaddieren der einzelnen Rechenleistungen von CPU-Kernen und GPU bei Kombiprozessoren als Beschönigung: Lässt man ihn zunächst jeweils nur auf den CPU-Kernen und danach auf der GPU laufen und addiert die Ergebnisse, kommt üblicherweise ein höherer Wert heraus, als wenn LuxMark sämtliche Rechenwerke gleichzeitig belastet. Der AMD A10-7850K liefert dann nur 82 Prozent der Summe der Einzelwertungen, der Core i7-4790 macht es mit 85 Prozent nicht viel besser. Der schwächere Core i5-4460 bringt jedoch 91 Prozent, noch darüber liegen die schlappen Mobilprozessoren.

Das lässt sich leicht erklären, denn die Rechenwerke der Kombiprozessoren sitzen jeweils auf demselben Chip und konkurrieren um die Datentransferleistung des RAM sowie um Energie beziehungsweise Kühlung: Schließlich ist jeder Prozessor für eine gewisse Thermal Design Power (TDP) spezifiziert und drosselt sich irgendwann, um Kühlsystem und Stromversorgung nicht zu überfordern. Die Prozessoren mit schwächeren Rechenwerken stoßen dabei später an die Grenzen.

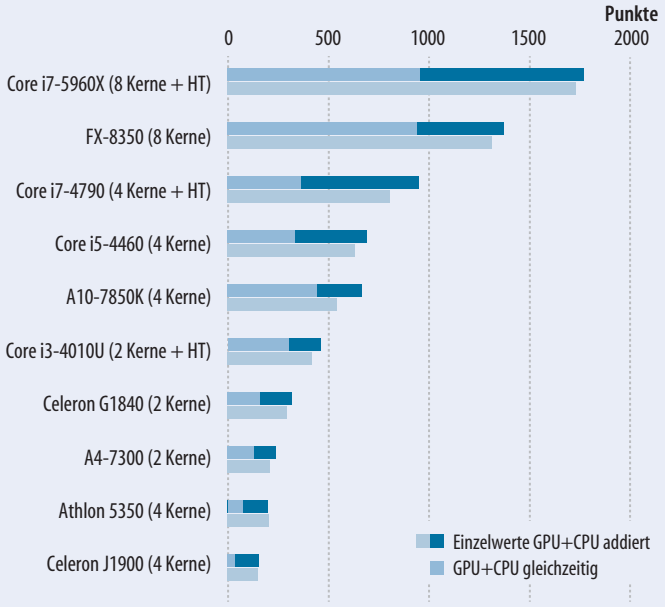
Der Supercomputer-Benchmark Linpack führt Berechnungen durch, die bei typischer PC-Software selten vorkommen. Daher liefert der Linpack für sich genommen wenig relevante Aussagen für Desktop-PCs und Notebooks. Interessant wird der Linpack dadurch, das er als Ergebnis die Rechenleistung in GFlops anzeigt. Dieser Wert lässt sich leicht mit dem theoretischen Leistungspotenzial des Prozessors vergleichen. Je mehr sich von dieser Performance mit hoch optimiertem Code nutzen lässt, desto besser hat der CPU-Hersteller die Caches und internen Datenpfade des Prozessors optimiert. Die theoretische Rechenleistung errechnet sich als Produkt aus Taktfrequenz, Anzahl der CPU-Kerne und der Zahl der pro Taktzyklus möglichen Operationen pro CPU-Kern. Dabei gibt es gewaltige Unterschiede: Pro Kern und Taktschritt kann Intels Haswell beispielsweise 16 doppelt genaue Gleitkommazahlen verarbeiten, ein AMD Steamroller bloß vier und der Atom-Celeron kümmerliche 1,5.

Beim Verhältnis aus gemessener und theoretischer Linpack-Leistung liegt der Celeron G1840 weit vorne: Er besitzt die für AVX2 optimierten Haswell-Kerne und schnellen L3-Cache, darf aber nur mit halb so leistungsfähiger SSE-Technik rechnen. Seine theoretische Rechenleistung liegt also zwar nicht besonders hoch, lässt sich aber gut ausnutzen. Der Mobil-Haswell Core i3-4010U schneidet vermutlich deshalb schlechter ab, weil wir ihn in einem passiv gekühlten PC vermessen haben. Der im Linpack absolut führende Core i7-5960X reizt sein Potenzial besser aus als etwa der Core i5-4460; vermutlich tragen der größere L3-Cache und der schnellere DDR4-Speicher dazu bei.

Auf AMD-Seite kann Linpack die Leistung des A10-7850K besonders gut nutzen. Er hat doppelt so große L2-Caches wie sein billigerer Dual-Core-Bruder A4-7300. Im FX-8350 steckt ältere Piledriver-Technik, außerdem konkurrieren hier doppelt so viele CPU-Kerne um die Datentransferrate zweier RAM-Kanäle – und der L3-Cache hilft wohl zu wenig. Auf dem Athlon 5350 mit Jaguar-Cores startete die

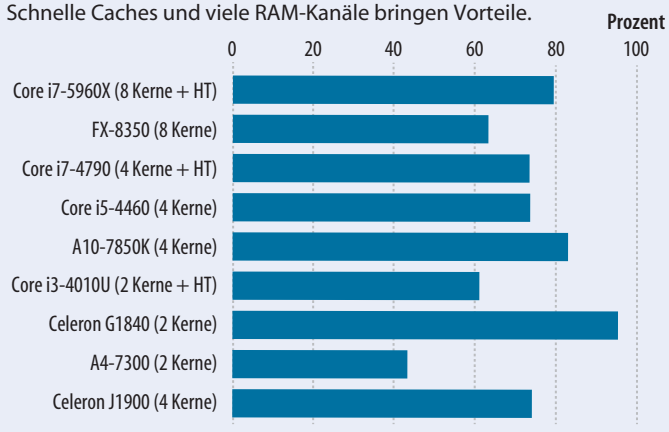
LuxMark: OpenCL

LuxMark berechnet die Szene Sala per OpenCL nach Wunsch auf CPU- oder GPU-Kernen – oder auf allen gleichzeitig. Die OpenCL-Performance der GPU-Kerne ist bei Kaveri deutlich höher als die der CPU-Kerne, bei den Billigheimern und manchen Haswells aber niedriger. Sind alle Kerne gleichzeitig belastet, sinkt die Rechenleistung. Werte für Core i7-5960X und AMD FX jeweils mit GeForce 750 Ti gemessen.



Linpack-„Wirkungsgrad“

Das Verhältnis des Linpack-Ergebnisses zur theoretischen DP-Gleitkomma-Performance zeigt, wie gut hoch optimierter Code die Rechenwerke ausreizen kann. Schnelle Caches und viele RAM-Kanäle bringen Vorteile.



© Copyright by Heise Zeitschriften Verlag
Persönliches PDF für alle Leser der c't aus 30625 Hannover

AMD-Version des Linpack leider nicht.

Fürs Büro

Sowohl der teure SYSmark 2014 des Industriegremiums BAPCo als auch der PCMark 8 der für den 3DMark bekannten Firma Futuremark bewerten Bürorechner. Ganz bewusst spielen bei diesen Benchmarks außer der reinen CPU-Rechenleistung etwa auch Festplatte (SSD) und GPU eine gewisse Rolle.

Beim SYSmark-Teil „Office Productivity“ liefern Adobe Acrobat XI, Excel und Word 2013 zusammen 78 Prozent der Punktzahl. Sie verhält sich bei vielen CPU-Typen ziemlich proportional zur Single-Thread-Performance – typisch für Office-Software. Bei der Gesamtwertung des SYSmark bringen mehr Kerne auch deutlich mehr Punkte: Hier spielen die Videoschnitt-Software Premiere Pro CS6 und Photoshop CS6 gewichtige Rollen, zusammen mit Excel machen sie rund 79 Prozent der SYSmark-2014-Gesamtnote aus.

PCMark 8 soll außer Multi-Threading auch Hardware-Beschleuniger gut nutzen, in den optionalen „Accelerated“-Durchläufen auch OpenCL. Dabei setzt Futuremark nur zum kleineren Teil auf real existierende Fertigprogramme wie LibreOffice Calc, das manche Rechenaufgaben mit OpenCL-Code löst. Stattdessen kommt etwa für den „Writing“-Ablauf, der 20 beziehungsweise 17 Prozent der Punktzahlen der Szenarien „Work“ und „Home“ ausmacht, eine in .NET geschriebene Mini-Textverarbeitung zum Einsatz, die typische Funktionen simuliert. Sie verwendet Standard-Routinen, von denen Windows einige mit DirectWrite beschleunigt. Die Komponente Bildbearbeitung im „Home“-Teil des PCMark 8 verwendet das Kommandozeilenprogramm ImageMagick, das OpenCL nutzen kann. Vom PCMark 8 werden häufig die Ergebnisse der Szenarien Work und Home veröffentlicht; für das Szenario Applications braucht man kostenpflichtige Software.

Im PCMark 8 Work schneiden die meisten AMD-Prozessoren vergleichsweise besser ab als im SYSmark 2014. Doch PCMark 8 Work misst auch nur vergleichsweise kleine Unterschiede zwischen sehr verschiedenen Prozessoren: Der langsamste Billigheimer Celeron J1900 erreicht etwa halb so viele Punkte wie der Core i7-5960X. Das deutet darauf hin, dass die per Hardware beschleunigten Teile des PCMark 8 auf vielen Prozessoren ähnlich schnell laufen. Im SYSmark 2014 Office ist der Abstand erheblich größer.

Außerdem verwirren beim PCMark 8 Work seltsame Streuungen der OpenCL-Punktzahlen: Im vermeintlichen „Accelerated“-Durchlauf des Szenarios Work sinken die Ergebnisse auf allen Intel-Prozessoren deutlich. Nur bei manchen ist das verständlich, weil deren GPUs auch der LuxMark niedrigere OpenCL-Performance als den CPU-Kernen bescheinigt. Beim Celeron G1840 sieht LuxMark jedoch CPU und GPU ähnlich schnell, trotzdem fällt er im PCMark 8 Work Accelerated um 23 Prozent zurück. Von OpenCL-„Beschleunigung“ würde man außerdem erwarten, dass sie intelligenterweise die schnellsten zur Verfügung stehenden Rechenwerke verwendet und nicht einfach irgendwelche. Futuremark weist außerdem darauf hin, dass die Werte der normalen und „Accelerated“-Messungen nicht untereinander vergleichbar sind.

Prozessoren: technische Daten (Papierform)

Prozessor	Kerne/ Threads	CPU-Typ	Fassung	Takt (Turbo)	CPU-Rechenleistung Double/Single Precision	GPU: GFlops Single Precision	TDP	Preis
Core i7-5960X	8/16	Haswell-E	LGA2011 v3	3,0 (3,5) GHz	384/768 GFlops	–	140 Watt	950 €
AMD FX-8350	8/8	Vishera	AM3+	4,0 (4,2) GHz	128/256 GFlops	–	125 Watt	155 €
Core i7-4790	4/8	Haswell	LGA1150	3,6 (4,0) GHz	230/461 GFlops	384 GFlops	84 Watt	275 €
Core i5-4460	4/4	Haswell	LGA1150	3,2 (3,4) GHz	205/410 GFlops	352 GFlops	84 Watt	160 €
AMD 10-7850K	4/4	Kaveri	FM2+	3,7 (4,0) GHz	59/118 GFlops	737 GFlops	95 Watt	140 €
Core i3-4010U	2/4	Haswell	–[1]	1,7 (–) GHz	54/109 GFlops	320 GFlops	15 Watt	281 US-$[2]
Celeron G1840	2/2	Haswell	LGA1150	2,8 (–) GHz	22/45 GFlops	168 GFlops	53 Watt	35 €
AMD A4-7300	2/2	Kaveri	FM2+	3,8 (4,0) GHz	30/61 GFlops	197 GFlops	65 Watt	35 €
AMD Athlon 5350	4/4	Jaguar	AM1	2,05 (–) GHz	25/66 GFlops	154 GFlops	25 Watt	50 €
Celeron J1900	4/4	Bay Trail-D	–[1]	2,0 (2,42) GHz	12/48 GFlops	43 GFlops	10 Watt	82 US-$[2]

[1] CPU zum Auflöten [2] Listenpreise laut Intel

Prozessor-Architekturen

Codename	CPU-Fassung	Typenbeispiele	Flops pro Taktschritt und Kern	Eigenschaften
Intel Haswell-E	LGA2011 v3	Core i7-5900, Core i7-5800	16 DP/32 SP	2 AVX2-Einheiten pro Kern, L3-Cache, Vierkanal-RAM, ohne GPU
Intel Haswell	LGA1150	Core i3/i5/i7-4000, Pentium, Celeron	16 DP/32 SP	2 AVX2-Einheiten pro Kern (außer Pentium/Celeron), L3-Cache, Zweikanal-RAM, mit GPU
AMD Vishera	AM3+	FX-8350, FX-6300, FX-4100	4 DP/8 SP (8/16 pro Modul)	Piledriver = 2. Generation Bulldozer; pro Modul 2 Kerne und 1 AVX-FPU, L3-Cache, Zweikanal-RAM, keine GPU
AMD Kaveri	FM2+	A10-7850K, A8-7600, A6-7400K, A4-7300	4 DP/8 SP (8/16 pro Modul)	Steamroller = 3. Generation Bulldozer; pro Modul 2 Kerne und 1 AVX-FPU, kein L3-Cache, Zweikanal-RAM, mit GPU
AMD Kabini	AM1/BGA[1]	Athlon 5350, A4-5200, E2-1800	3 DP/8 SP	Jaguar = 2. Generation Bobcat; AVX in mehreren Zyklen, kein L3-Cache, Einkanal-RAM, mit GPU
Intel Bay Trail	BGA[1]	Atom Z3000, Celeron J1900, Celeron N2930	1,5 DP/6 SP	Silvermont: kein AVX, nur SSE4, kein L3-Cache, Einkanal-RAM, mit GPU

[1] Ball Grid Array: Chip zum Auflöten

Fazit

Ein Benchmark misst bloß das, was er selbst tut. Eine einzelne Zahl reicht nur aus, wenn man sie anhand von Erfahrungswerten einordnen kann. Aber auch wer dutzende Benchmarks vergleicht, muss weitere Informationen einbeziehen, etwa zur sonstigen Hardware-Ausstattung des Rechners. In flachen Notebooks oder bei lüfterloser Kühlung lässt sich das volle Leistungspotenzial oft nur für Minuten nutzen, weil die Chiptemperatur so schnell ansteigt, dass sich die CPU drosseln muss. Für Aufgaben, bei denen stundenlang Volldampf nötig ist, eignen sich solche Rechner weniger. Trotzdem profitieren sie von schnellen Prozessoren: Ihre Leistung wird im typischen PC-Alltag nur kurzzeitig gefordert und macht den Rechner spritziger.

Geht es um einen Office-PC, ist meistens der Prozessor mit den kräftigeren Einzelkernen die bessere Wahl – ebenso wie für einen Surf-PC, weil JavaScript auch bloß einen Kern nutzt. Die Multi-Threading-Teile von Cinebench, CoreMark, kcbench und der CPU-Teil des LuxMark ermöglichen eine grobe Einschätzung des Leistungspotenzials sowie treffende Vergleiche innerhalb von Prozessorfamilien. Das Verhältnis aus Linpack-Messwert und theoretischer Maximalleistung liefert Hinweise darauf, wie weit sich die Rechenwerke ausreizen lassen.

Kompliziert ist der Vergleich von Prozessoren, die sehr unterschiedliche Rechenwerke und (3D-)Beschleuniger enthalten. Dann kommt es darauf an, wozu man seinen Computer vorwiegend benutzen möchte. Besonders schwierig ist eine Bewertung der per OpenCL nutzbaren GPU-Rechenleistung. LuxMark und PCMark liefern zu manchen Prozessoren gegensätzliche Aussagen. PCMark 8 Work wendet die Bedeutung „Accelerated“ auf Intel-CPUs ins Gegenteil. Die Aussagekraft solcher Messungen für die Praxis ist fragwürdig: Wer würde „Beschleuniger“ einschalten, die sein System bremsen? Im Zweifel geht kein Weg daran vorbei, die zur Auswahl stehenden Prozessoren mit der gewünschten Software selbst zu vergleichen – zur Not mit der Stoppuhr. (ciw)

c't Benchmarks: ct.de/yknf

© Copyright by Heise Zeitschriften Verlag
Persönliches PDF für alle Leser der c't aus 30625 Hannover