

30 Jahre Standard Performance Evaluation Corporation

Wer misst ...



Ralph Hülsenbusch

1988 wurde die SPEC gegründet, eine Herstellerorganisation, die Werkzeuge zur Leistungsmessung bereitstellt und die damit gewonnenen Resultate veröffentlicht.

Als Prozessoren nur einen Kern besaßen und ihr Herz noch im Megahertz-Takt schlug, gab es eine Reihe von Progrämmchen, mit denen die Entwickler versuchten, der Leistung von Computern auf den Zahn zu fühlen. Dazu zählen etwa der Dhrystone für den Bereich Festkomma- und der Whetstone fürs Gleitkommarechnen. Mips/Joy entstand auf einem Segelturn und Doduc im Dunstkreis der Kernforschung. Überlebt hat der Linpack bis heute als Basis für die Kür der schnellsten Supercomputer weltweit, die TOP500. Das Lösen linearer Gleichungssysteme kommt halt nicht aus der Mode.

All diese Verfahren begannen früh, Druck auf die Hersteller von Rechnersystemen auszuüben, und der Wildwuchs der Testverfahren erschwerte das Geschäft. Denn Benchmarks erlauben Systemvergleiche und führen zu – mehr oder weniger vernünftigen – Ranglisten, die für Her-

steller geschäftsentscheidend sein können. Fälschungen sind um so eher möglich, je weniger transparent die Verfahren sind. Bestes Beispiel ist das Schicksal des Whetstone. IBMs optimierende Compiler deklarierten Berechnungen, die keine Ausgaben erzeugten, als unproduktiv und übergingen sie einfach.

Transparenz ist Trumpf

Deshalb fanden sich einige Computerbauer zusammen und gründeten ein Konsortium, um gemeinsam eine Testbatterie zusammenzustellen. Sie sollte von allen Mitgliedern kontrolliert werden und vor allem von allen akzeptierte, vergleichbare Resultate liefern. So gründete sich 1988 die System Performance Evaluation Corporation SPEC, die 1989 ihre erste Benchmark-Suite veröffentlichte (heute: Standard Performance Evaluation Corpo-

ration). Die SPEC CPU89 konzentrierte sich auf die Arbeitsgeschwindigkeit der Prozessoren. Sie bestand aus insgesamt 10 Programmen, je fünf für den Festkomma- und den Gleitkommabereich. Sie lieferte als Resultat sogenannte SPECmarks.

Die SPEC entwickelt die Tests nicht selbst, sondern veranstaltet eine Ausschreibung. Die Einreichungen – in der Regel Benchmarks oder Anwendungen aus der Praxis – nehmen die Mitglieder der SPEC unter die Lupe und entscheiden, welche in die Suite einziehen dürfen.

Inzwischen deckt der Komplex über 10 Bereiche ab. Jüngster Zugang ist SPECjEnterprise2018 Web Profile, bei dem es um Java und WebSphere im geschäftlichen Umfeld geht. In anderen Bereichen spannt man andere Werkstücke auf die Bank: Programme zum Ausloten von Anwendungen in der Cloud, computergestützter Grafik, Parallelverarbeitung, Java-Clients und -Servern, Mailservern, Speicher, Stromverbrauch, Virtualisierung und letztlich Webservern.

Zyklen werden länger

Nach wie vor betriebsam geht es auch in der CPU-Suite zu. Allerdings hat die SPEC für die letzte Renovierung über 10 Jahre gebraucht. Dem 2007 erschienenen CPU2006 folgte im Juni 2017 die CPU2017. Für diese sechste Ausgabe haben die Architekten deutlich sichtbar Umbauten vorgenommen. So hat sich die Größe der Testsuite gegenüber der CPU2006 nahezu verdoppelt und sie zieht sich über vier anstatt bisher zwei Etagen hin: Je zwei Abschnitte für Tests im Single-CPU- (speed) und Multiprozessormodus (rate), die wiederum unterteilt nach Fest- und Gleitkomma (int und fp).

Um die Ergebnisse beurteilen zu können, bedarf es eines tieferen Einblicks in die Struktur des Benchmarks und die Besonderheiten der neuen Version. Die Programme der Test-Suite haben entweder ein Update erfahren oder sind neu, somit ist ein Vergleich der CPU2017 mit der CPU2006 nicht möglich, es sei denn, man greift zu rein statistischen Methoden etwa der Art: Ist A immer noch soundsoviel schneller als B? Ein kaum durchführbares Unterfangen, denn sowohl A als auch B haben sich verändert.

Die Mühe muss sich schon lohnen: Nach wie vor kostet die Suite 1000 US-Dollar, jede von der SPEC geprüfte Veröffentlichung von Resultaten weitere 1500 US-Dollar. Gemeinnützige Organisationen erhalten einen Abschlag. SPEC CPU ist relevant für Ausschreibungen.



- 1988 gründeten Systemhersteller die international operierende gemeinnützige Organisation SPEC, um Standards im Bereich der Computerbenchmarks zu schaffen.
- Das erste veröffentlichte Werkzeug zum Messen der Rechnerleistung, die CPU92, bezog sich ausschließlich auf Prozessoren.
- Inzwischen ist die CPU2017-Suite aktuell, die sich im Aufbau von ihren Vorgängern maßgeblich unterscheidet.

Zum Ende des Jahres 2017 hat die ältere Version ihre Gültigkeit verloren. Die Open System Group (OSG), bei der SPEC zuständig für Systemtests, nimmt ab 2018 keine Ergebnisse mehr an, die mit der 2006er-Version entstanden sind. Eines der Handicaps von Testprogrammen: Sie altern und verlieren mit jedem Update meist ihre Gültigkeit. Deswegen führte an einem Neubau der Suite nichts vorbei.

Alternative Linpack

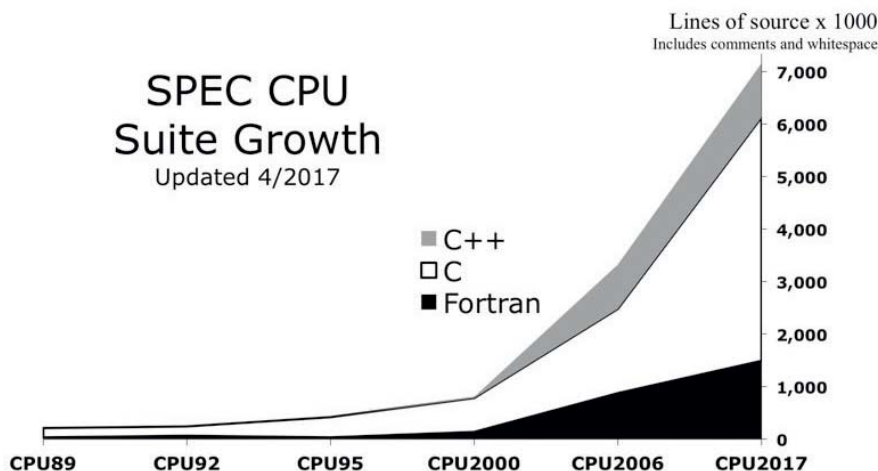
Allerdings gibt es im High-Performance Computing den bereits oben erwähnten Linpack, der immer noch die Grundlage für die Bestenliste der Supercomputer TOP500 liefert. Bei ihm geht es um ein mathematisches Verfahren, die lineare Regression. An dem hat sich im Prinzip nichts geändert. Trotzdem gibt es eine heftige Diskussion darum, wie praxisrelevant der Test noch ist. Inzwischen sind weitere Verfahren für HPC in Arbeit, etwa HPL oder HPGG. Die führen übrigens nicht unbedingt zu denselben TOP500-Listen.

Bei der SPEC gibt es ebenfalls Aktivitäten, für Supercomputer relevante Benchmarks zu liefern. Schon 1996 erschien der HPC96, den die SPEC aber 2003 in Rente geschickt hat. Ein Schicksal, das den meisten Benchmarks blüht. Gründe können sein, dass sie nicht mehr praxisrelevant oder anfällig gegen Manipulationen sind, veraltete Methoden verwenden, nicht genügend Last erzeugen oder dass neue Compiler-Generationen sie schlicht wegoptimieren.

„Traue nur dem Benchmark, den du selbst gefälscht hast“, die Abwandlung eines Spruchs über Statistiken, den man Churchill zugeschrieben hat, beschreibt die prekäre Lage, in der sich Tester von Systemen befinden. Berechtigtes Misstrauen gegenüber Zahlen und Werten, die Hersteller und beteiligte Organisationen veröffentlichen, ist nicht nur aus aktuellem Anlass gegeben. Beste Beispiele liefert hierzu die Autoindustrie – ob mit manipulierten oder praxisfernen Daten.

SPEC CPU Suite Growth

Updated 4/2017



Immenser Zuwachs: Von Version zu Version hat sich die Zahl der Codezeilen vervielfacht.

Hinzu kommt, dass man als Tester leicht in Fallen tappen und Fehler machen kann: „Wer misst, misst Mist“ war schon in *iX* 10/1993 zu lesen. Meist erntet, wer Benchmarks ins Feld führt, skeptische Blicke oder stößt gar auf Ablehnung.

Aber ohne vergleichbare Zahlen geht es eben doch nicht. Die technischen Eckdaten wie Prozessortakt und Zahl der Kerne reichen nicht, wenn Investitionen ins Spiel kommen. So spielen die Resultate von Benchmarks bei Investitionsentscheidungen eine große Rolle. Sie sollen helfen, das angebotene System mit anderen in der Leistung vergleichen zu können.

Benchmarks sollen alltagsnah sein

Das Besondere daran: Das Messverfahren soll produktive Bedingungen simulieren, im Unterschied zu synthetischen Benchmarks wie Whet- und Dhrystone, die mit speziellen Algorithmen Prozessoren an ihre Grenzen bringen wollen. Alltagsnahe Tests zu entwickeln, ist ein hochgradig komplexes Unterfangen, denn es bezieht sich auf komplette Systeme mit unterschiedlichen Hardwarekomponenten wie Hauptspeicher, Massenspeicher und Netzwerk nebst der zugehörigen Software wie Betriebssystem, Compiler und Systemprogrammen. Da die CPU2017-Suite wie ihre Vorgänger im Quelltext vorliegt, müssen die Tester zuerst die Bi-

naries erzeugen, sprich das lauffähige Programm samt seinen Bibliotheken. Somit kommen neben der Qualität der Compiler deren Optionen und Optimierungen hinzu, die auf der einen Seite erlauben, die Leistungen der CPUs auszureizen, auf der anderen aber der Manipulation Tür und Tor öffnen.

Deshalb hat die SPEC von Anfang an Regeln für das Verwenden ihrer Benchmarks festgelegt. Beim CPU unterscheiden sie zwischen zwei Disziplinen: base und peak. Im ersten Fall legen die Run Rules strengere Bedingungen fest. So ist es Pflicht, für eine Suite nur ein und dieselbe Entwicklungsumgebung zu verwenden, und die Wahl der Optimierungsoptionen ist begrenzt. Bei peak hingegen, der Kür sozusagen, dürfen es zum Feintuning verschiedene Versionen und Varianten sein. All das müssen die Benchmarkmarker in einer öffentlich einsehbaren Konfigurationsdatei beschreiben und festlegen. Stimmt etwas nicht oder entdeckt die Software einen Verstoß gegen das Regelwerk der SPEC, bricht der Installationsvorgang der Suite ab und es erscheint in den Ergebnissen der Hinweis „Invalid Run“, teils quer über die Dokumente gedruckt.

Seit Jahren gibt es eine Diskussion zwischen der SPEC und der *iX*-Redaktion betreffs der Festlegungen für die Base Line in den Run Rules. In den Labors des Verlages kommt eine Customer Base Line zum Zuge, die keine zusätzlichen

Quelle: www.spec.org/cpu2017/Docs/overview.html#Q2

Bestandteile der SPEC CPU2017

Kurzbezeichnung	Suite	Inhalt	Maßeinheit	Anzahl paralleler Läufe
intspeed	SPECspeed 2017 Integer	10 Integer-Benchmarks	SPECspeed2017_int_base, SPECspeed2017_int_peak	1
fpspeed	SPECspeed 2017 Floating Point	10 FP-Benchmarks	SPECspeed2017_fp_base, SPECspeed2017_fp_peak	1
intrate	SPECrate 2017 Integer	10 Integer-Benchmarks	SPECspeed2017_int_base, SPECspeed2017_int_peak	1 - n
fpate	SPECrate 2017 Floating Point	13 FP-Benchmarks	SPECspeed2017_fp_base, SPECspeed2017_fp_peak	1 - n

Libraries und undokumentierten Optionen zulässt. Die Resultate liegen daher in der Regel einige Prozentpunkte unter den bei der SPEC veröffentlichten. Es scheint doch recht praxisfern, davon auszugehen, dass jemand seine Anwendungen neu übersetzt und dabei Fremdbibliotheken einbindet. Das mag höchstes für spezialisierte Supercomputer gelten.

Zur Wahl stehen beim CPU-Test zwei Abteilungen der Suite: SPECint für Festkomma- und SPECfp für Gleitkommaberechnungen. Neben C/C++ spricht man noch Fortran, eine 1957 entstandene Programmiersprache, die in wissenschaftlichen Anwendungen immer noch eine Rolle spielt, nicht zuletzt wegen der umfangreichen mathematischen Bibliotheken mit ihren wissenschaftlich geprüften Verfahren und hoher Genauigkeit.

Außerdem gibt es noch die Variante „speed“, bei der jeweils nur eine Kopie des Testprogramms läuft, und „rate“ für den Multiprozessortest mit mehreren Kopien (siehe Tabelle „Bestandteile der SPEC CPU2017“). Hat alles geklappt, erhält man je nach Konfiguration des Testrechners für jede Disziplin nach zig Stunden ein mehrseitiges Dokument – letztlich aber auch jeweils eine Zahl, die sich etwa liest wie: SPECrate2017_int_peak.

Damit alles vergleichbar bleibt, gilt nach wie vor als Referenz ein Server von einst Sun Microsystems, seit 2010 im Besitz von Oracle: Es handelt sich um eine SunFire V490 mit UltraSPARC-CPU vom Februar 2007, derzeit ab 319 Euro im Netz zu haben. Sie legt in allen Disziplinen eine 1 als Bezugsgröße fest. Bei der CPU2006 war es noch eine Ultra Enterprise 2 aus dem Jahr 1997. Die dort verwendete SPARC-Architektur spielt aber kaum noch eine Rolle, heute dominiert x86/x64 das Geschehen.

SPARC-Prozessor als Referenz

Die Suite weiterzuentwickeln, stellt alle Beteiligten vor eine immense Aufgabe. Ursprünglich sollten neue Versionen der CPU-Suite im Jahresrhythmus herauskommen. Aber das klappte selbst bei den früheren Versionen nicht. Ursache der Verzögerung ist unter anderem das bereits erwähnte Auswahlverfahren, dem sich die neuen Versionen bei der SPEC unterziehen müssen. Infrage kommen nur solche, die vom Einsatz befindlichen Programmen abgeleitet sind. In älteren Versionen vorhandene Tests müssen sich vor dem Einzug in die neue CPU-Suite ebenfalls einer Revision unterziehen.

Zu den Kinderkrankheiten von Software zählt der Umgang mit der Speicher-verwaltung. In der CPU89 gab es ein Modul zur Matrizenberechnung. Ein Mitarbeiter von IBM änderte dessen Datenfeld um eine Spalte und erhielt völlig andere Resultate. Die neue Matrize passte bei dem einen System besser in den Hauptspeicher, ein anderes musste laufend Daten nachladen. Aber auch in der bis dato gültigen CPU2006 gibt es für die Datenhaltung im Speicher hochsensible Tests. Das Thema sind vor allem Leerstellen in Matrizen, sogenannte Sparse-Matrizen. Verteilt ein Programm seine Daten in großen Abständen im Speicher, muss die Routine zur Speicher-verwaltung mächtig hin- und herspringen. Clevere Compiler kriegen das mit und räumen im Hintergrund auf. Die Performancegewinne sind so groß, dass die Resultate allen andern davonlaufen. So geschehen beim 462.libquantum in der CPU2006, einer C-Bibliothek zum Simulieren eines Quantencomputers.

Da es sich bei den SPEC-Benchmarks um sogenannte Suites handelt, also um eine Zusammenstellung mehrerer Programme, möchte nicht nur jeder Hersteller seine Favoriten mit unterbringen, auch die Zahl der Anwendungsfälle wächst. Das Konsortium hat die Zahl der Tests in der CPU2017 gegenüber der 2006er-Version etwas abgespeckt.

„Offizielle“ Ergebnisse veröffentlichen dürfen prinzipiell nur die Mitglieder. Die SPEC bittet dafür zur Kasse, denn erst nach einer eingehenden Prüfung publiziert sie selbst. Ein arbeitsaufwendiger Prozess, da die Vertreter aller Mitglieder an dem Verfahren teilnehmen. Trotzdem ist es vorgekommen, dass sie Resultate eines Herstellers zurückziehen musste, weil das angegebene System nie zum Einsatz kam. Andere Fälle waren nicht verfügbare Compiler-Versionen oder praxisferne Konfigurationen. Aber damit steht die SPEC nicht allein. Wer sich die Daten des Benchmarks des Transaction Performance Council (TPC) anschaut, begegnet dort schauerlichen Geisterrechnern.

Wer selbst die CPU2017 einsetzen möchte, ist an die Regeln des Fair Use gebunden. Das Einrichten der Suite und das Durchführen von Tests kann mehrere Tage in Anspruch nehmen. Ergebnisse darf man auch außerhalb der SPEC veröffentlichen und dabei auch nur Teile der Suite verwenden, was unter anderem für Supercomputer sinnvoll sein kann.

Resultate, die nicht auf der Site des Gremiums erscheinen, sind nicht von der SPEC geprüft und gelten als nicht freigegeben.

Hingegen frei öffentlich zugänglich sind die Resultate auf www.spec.org/cpu2017, wie das bei allen anderen Suites ebenso der Fall ist. Deren Umfang nebst zugehörigen Dokumenten und Protokollen liefert eine nahezu unerschöpfliche Menge an Material für nähere Untersuchungen. Ein Resultat umfasst drei und mehr Seiten. Es liegt in vier Formaten vor: HTML, PDF, CSV und TXT.

Multiprozessingfähigkeit integriert

Jahrelange Debatten um die Multiprozessingfähigkeit der SPEC-Suite haben zu einem Kompromiss geführt: Im rate ist nach wie vor kein automatisches Parallelisieren per OPENMP erlaubt und die verwendeten Testprogramme sind nicht multiprozessorfähig. Für Läufe à la SPECrate startet der Tester eine gewünschte Anzahl von Kopien ein und desselben Binary; in der Regel so viele, wie das System an Cores besitzt. Nur bei SPECspeed darf der Tester die Option zum Parallelisieren innerhalb eines Programms einschalten. Ein weiterer Aspekt, der den Vergleich mit früheren CPU2006-Resultaten verbietet.

Was man nach wie vor erhält, ist je Abteilung ein Benchmark, errechnet als geometrisches Mittel der Resultate der einzelnen Tests. Das Verfahren bietet den Vorteil, dass die höheren Ergebnisse nicht so zu Buche schlagen wie beim arithmetischen Mittel. Den Effekt konnte man vor allem bei der CPU2006 im SPECint beim oben erwähnten Ausreißer libquantum beobachten. Während der Test auf einem Server von Dell, Typ R940, einen Wert von 8380 erreichte, lagen die anderen 11 Testprogramme im Bereich zwischen 30 und 111. Heraus kam SPECint_base2006 = 52,5. Beim arithmetischen Mittel wäre rund 752 herausgekommen. Solche Ausreißer gibt es in der CPU2017 nicht mehr.

Für den Vergleich von Systemen spielt zuerst einmal die resultierende Zahl eine Rolle. Auf der Site der SPEC kann, wer die Vergleichbarkeit und Abhängigkeiten unter die Lupe nehmen möchte, dies mit konfigurierbaren Abfragen tun. Die SPEC bietet den Zugriff auf eine große Menge an Parametern an, wie CPU-Typ, Zahl der Prozessoren oder Kerne, Compiler, Betriebssystem und Optionen zum Sortieren – um nur einige zu nennen. In vielen Fällen dürfte es aber um die Eigenschaft der untersuchten Systeme für spezielle Aufgaben gehen. Da lohnt sich der Blick auf die einzelnen Tests.

Die Benchmarks in der SPEC CPU 2017

SPECrate 2017 Integer	SPECspeed 2017 Integer	Programmiersprache ¹	KLOC ²	Anwendung
500.perlbench_r	600.perlbench_s	C	362	Perl
502.gcc_r	602.gcc_s	C	1,304	GNU C
505.mcf_r	605.mcf_s	C	3	Routenplanung
520.omnetpp_r	620.omnetpp_s	C++	134	ereignisorientierte Simulation
523.xalancbmk_r	623.xalancbmk_s	C++	520	XML-HTML-Wandlung via XSLT
525.x264_r	625.x264_s	C	96	Videokompression
531.deepsjeng_r	631.deepsjeng_s	C++	10	AI: Alpha-Beta-Baumsuche (Schach)
541.leela_r	641.leela_s	C++	21	AI: Monte-Carlo-Baum (Go)
548.exchange2_r	648.exchange2_s	Fortran	1	AI: rekursive Lösungssuche (Sudoku)
557.xz_r	657.xz_s	C	33	Datenkompression
SPECrate 2017 Floating Point	SPECspeed 2017 Floating Point	Programmiersprache ¹	KLOC ²	Anwendung
503.bwaves_r	603.bwaves_s	Fortran	1	Explosionsmodelle
507.cactuBSSN_r	607.cactuBSSN_s	C++, C, Fortran	257	Relativitätstheorie
508.namd_r		C++	8	Molekulardynamik-Simulation
510.parest_r		C++	427	biomedizinische Bildgebung
511.povray_r		C++, C	170	Ray Tracing
519.lbm_r	619.lbm_s	C	1	Fluidodynamik
521.wrf_r	621.wrf_s	Fortran, C	991	Wettervoraussage
526.blender_r		C++, C	1,577	3D-Rendering und -Animation
527.cam4_r	627.cam4_s	Fortran, C	407	Atmosphärenmodellierung
	628.pop2_s	Fortran, C	338	weiträumige Ozeanmodellierung
538.imagick_r	638.imagick_s	C	259	Bildbearbeitung
544.nab_r	644.nab_s	C	24	Molekulardynamik-Simulation
549.fotonik3d_r	649.fotonik3d_s	Fortran	14	Berechnung elektromagnetischer Felder
554.roms_r	654.roms_s	Fortran	210	regionale Ozeanmodellierung

¹bei mehreren Sprachen die erste zum Linken aufgeführte; ²Lines of Code in Tausend

Von den insgesamt 29 Testprogrammen der CPU2006 – 12 in der Integer- und 17 in der Floating-Point-Gruppe – schrumpfte die 2017er-Fassung auf 10 für SPECint und 14 für SPECfp zusammen. Im erstgenannten Bereich bleiben fünf Kandidaten erhalten, beim Gleitkomma-rechnen sechs. Die einzelnen Methoden sind auf www.spec.org, der Site der SPEC, im Detail beschrieben.

Im Festkommabereich blieben dabei: der Perl-Interpreter-Test (*perlbench*), GCC *gcc*, der Routenplaner (*mcf*), die Simulation eines Computernetzes (*omnetppd*) und der Konverter für XML nach HTML (*xalancbmk*). Bei den neu hinzugekommenen liegt ein deutlicher Schwerpunkt auf der künstlichen Intelligenz, die sich im Schach (*deepsjeng*), beim Go (*leela*) und im Sudoku (*exchange*) bewähren darf. Hinzugekommen sind außerdem zwei Kompressionsverfahren für Video (*x264*) und Daten generell (*xz*). Bis auf eine Ausnahme sind alle Programme wie bisher in C oder C++ geschrieben, für 2017 haben die Entwickler bei der Sudoku-Simulation im SPECint Fortran gewählt.

SPECfp nutzt vorwiegend wissenschaftliche Anwendungen als Vorlage. Erhalten blieben aus der Flüssigkeitsdynamik das spezielle Explosion Modelling *bwaves* und das allgemeine *lbm*, aus der Physik Berechnungen zur Relativitätstheorie per *cactuBSSN*, die Simulation biomolekularer Modelle *namd*, das aus der Computervisualisierung stammende Ray-Tracing-Tool *povray* und schließlich steuert die Klimaforschung eine Wettervorhersage namens *wrf* bei. Neu eingezo-

gen sind aus der Tomografie *parest*, der 3D-Visualisierung *blender*, der Klimaforschung *cam4*, der Bildverarbeitung *imagick*, der Molekulardynamik *nab*, der Berechnung elektromagnetischer Felder *fotonik3d* und der regionalen Erforschung von Ozeanen *roms*. Mit den Weltmeeren beschäftigt sich *pop2*, das nur im Single-Test zum Zuge kommt.

Während beim Integer-Test die Einrichtung für *speed* und *rate* auf derselben Grundlage beruht, unterscheidet sie sich bei der fp-Suite: *pop2* gibt es nur im Single-Test *speed*, *namd*, *parest*, *povray* und *blender* einzig im *rate*. Diese Unterscheidung gab es bei der CPU-Suite bis dato nicht.

Und immer noch kein Windows

Alle bisher veröffentlichten Ergebnisse stammen von Systemen, die unter Linux liefen und Intels Compiler für C/C++ und Fortran verwendeten. Einzig die Sun Fire als Referenzsystem und eine 1-Chip-VM mit SPARC M7 liefen unter Solaris und nutzten die Compiler von Oracle. Ein weiterer neuer Aspekt ist der Stromverbrauch der Systeme, den Anwender zusätzlich messen können. Allerdings ist das noch nicht zwingend.

Im November 2018 lagen rund 7500 erste Ergebnisse vor, davon gut 3000 der Kategorie *speed*, also Single-CPU, und etwa 4500 Multi-CPU-Ergebnisse. Das wirft ein erstes Schlaglicht auf die Akzeptanz der neuen Testbatterie: Einige

Systemhersteller beschränken sich auf den Multiprozessor-Test der klassischen Art, obwohl es genügend Anwendungen gibt, die singular auf einem Rechner laufen, und zunehmend mehr, die multiprozessorfähig sind.

Das Herstellerprofil enthält neben den erwähnten SPARC-Referenzen Fujitsu als einzigen Anbieter mit SPARC-M12-CPU, Solaris als Betriebssystem und Compilern von Oracle. Daneben veröffentlicht der Hersteller aber auch Resultate für seine Intel-Plattform Primequest.

Vom erwarteten Rennen zwischen Intel und AMD bei den Prozessoren ist bisher wenig zu sehen, die 303 Systeme mit AMDs Epyc CPU erscheinen gegenüber 7197 mit Intels Xeon unter ferner liefen.

Auffällig ist der hohe Anteil im SPECfp, bei dem es vorwiegend um wissenschaftliche Anwendungen geht. Die Rechner, von denen die Hersteller Benchmarks eingereicht haben, sind in der Regel im RZ zu finden. Da kommen bei der Relevanz der reinen CPU-Leistung Zweifel auf, denn über die Performance von Netz und Speicherverwaltung sagt die SPECcpu nichts aus. Wer nach neuen Plattformen für Server sucht, muss also weitere Testverfahren einsetzen.

(js@ix.de)

Ralph Hülsenbusch

war leitender Redakteur bei iX und ist heute gelegentlich als freier Autor tätig.

Alle Links: ix.de/ix1901096

