

Heiße Rechner

Die International Supercomputing Conference 2017 in Frankfurt am Main

Mit über 3200 Teilnehmern legte die ISC17 abermals zu. Sie konnten auf der Konferenz zu aktuellen Themen des High Performance Computings vor allem auch die begleitende Ausstellung genießen – rund um Intel Xeon, IBM Power9, Cavium ThunderX2, AMD Epyc, NEC Aurora ...

Von Andreas Stiller

Traditionell wird zur Eröffnung der ISC die neue Top500-Liste der Supercomputer vorgestellt. Hier ragte, wie schon in der letzten Ausgabe berichtet, Europas schnellster Supercomputer Piz Daint vom schweizerischen Supercomputerzentrum (CSCS) in Lugano heraus – nicht nur in der Performance, sondern vor allem auch in

der Energieeffizienz. Mit 10,4 GFlops/Watt kam Piz Daint auf Platz 6. Das CSCS war bislang allerdings auch der einzige Betreiber, der einen nach den Regeln der Green500-Liste erlaubten, speziell auf Energieeffizienz optimierten Lauf eingereicht hatte. Dieser war zwar um 13 Prozent langsamer, verbrauchte dafür aber um 28 Prozent weniger Energie. Bei voller Linpack-Performance wären es 9,2 GFlops/Watt gewesen, damit hätte Piz Daint Platz 12 auf der Green500-Liste erreicht.

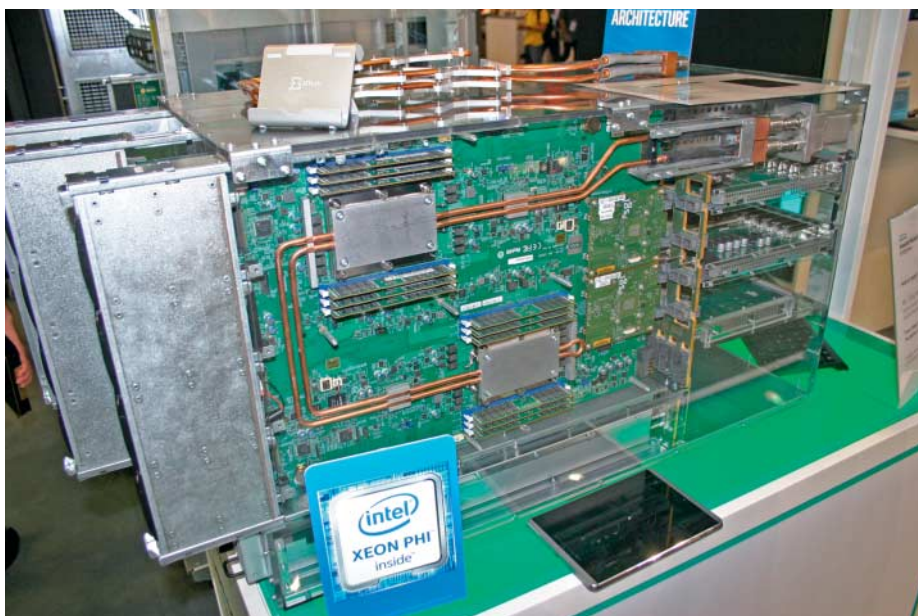
Green500-Sieger wurde mit 14,1 GFlops/Watt der mit Nvidia P100 bestückte Tsubame 3.0 am GCIC-Center in Tokio. Das Siegersystem HPE SGI 8600 mit SGIs ICE-XA-Architektur stellte HPE stolz am Stand aus. Nach Kauf von SGI entschied sich HPE für diese wassergekühlte SGI-Linie und opferte dafür die hauseigene Apollo 8000. Das Besondere der Kühltechnik der ICE-XA-Architektur ist die Durchflutung der Kühlkörper mit

eng aneinanderliegenden Zu- und Rückläufen. Dieser Trick soll für etwa gleiche Temperaturen der zu kühlenden Chips sorgen. Bei kleineren Servern setzt HPE auf die mit Kunststoffverbindungen arbeitende Kühltechnik von CoolIT.

Heißwassergekühlte Systeme sah man allüberall, von Lenovo, IBM, Dell, Bull ... Auch das Chemnitzer HPC-Haus Megware ist mit seiner Kühltechnik Cold-Con gut dabei. Es hat im Rahmen der Slice-LC-Serie zudem wassergekühlte Netzteile und als Erster auch heißwassergekühlte Omni-Path-Switches im Programm. Aus dieser Serie wird derzeit der CoolMUC3 am Leibniz-Rechenzentrum mit 148 Knoten aufgebaut.

Thunder-Metamorphose

Bull hat das heißwassergekühlte Blade-System Sequana X1310 vorgestellt, das mit Cavium Thunder X2 laufen soll. Auch Penguin hat ein Thunder-X2-System im Angebot, jedenfalls kann man es schon bestellen. Die Auslieferung ist fürs dritte Quartal geplant. Ob das bis dahin klappt, sei dahingestellt. Thunder X2 ist nämlich nicht mehr der Thunder X2, den Cavium im letzten Jahr angekündigt hatte. Anfang des Jahres hatte Cavium nämlich von Avago/Broadcom das Vulcan-Design aufgekauft. Wie es auf der ISC offiziell hieß, habe man inzwischen die Designs „ge-

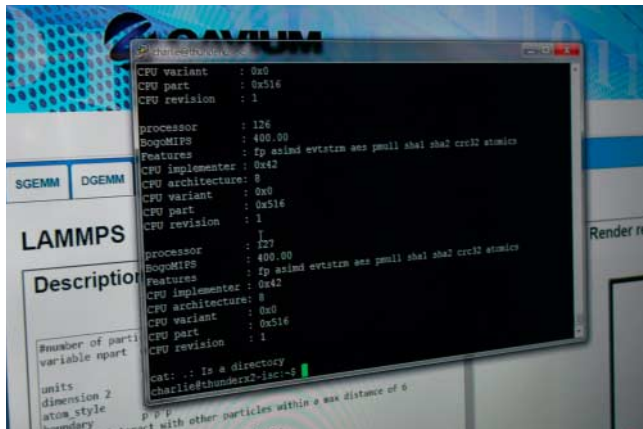


Nein, kein schneller „Kupfer-Interconnect“ zwischen den beiden Xeon-Phi-Knoten, sondern die spezielle SGI-Kühltechnik mit dicht aneinanderliegenden Zu- und Rückläufen, so wie sie in der HPE SGI 8600 zum Einsatz kommt



Hier zum Vergleich das wassergekühlte Xeon-Phi-Serverboard von Megware

Anzeige



Cavium ThunderX2 wurde intern mit dem zugekauften Vulcan von Broadcom „gemergt“. Noch sah man keinen lauffähigen Rechner, aber wir konnten remote auf ein Entwicklungssystem mit 128 Kernen zugreifen.

mergt“. Insider berichteten, dass wesentliche Kernteile aus dem Vulcan eingeflossen seien. Als Chefarchitekt wirkt inzwischen Avinash Sodani, der im letzten Jahr noch Intels Chefarchitekt für Xeon Phi war. Man sah schon etliche riesige ThunderX2-Chips, allerdings kein lauffähiges System.

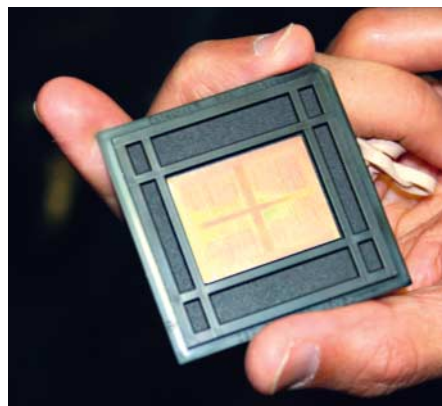
Der beeindruckendste neue Rechner ist wohl der Mare Nostrum 4 am Supercomputerzentrum in Barcelona. Allein schon seine Umgebung ist einzigartig: die alte Kapelle „Torre Girona“, die auch schon seine Vorgänger beherbergte. Und dann das Konzept, das unter Federführung von IBM Zug um Zug in Zusammenarbeit mit Lenovo, Intel und Fujitsu umgesetzt werden soll. Am Barcelona SC sollen nämlich gleich vier verschiedene Architekturen aufgebaut werden. Die erste und größte Ausbaustufe wurde jetzt in rekordverdächtiger Bauzeit fertiggestellt. Lenovo begann im Mai mit dem Aufbau der 48 Racks mit den ThinkSystem-SD530-Servern, bestückt mit Intels nagelneuen Skylake-SP-Prozessoren (Platinum 8160). Und schon im Juni erreichte die Ausbaustufe mit insgesamt 3400 Knoten 11 PFlops und konnte damit als zweit-schnellster Europäer Platz 13 in der Top500-Liste erreichen. Etwas kleiner ist der fürs nächste Jahr geplante Cluster mit Power9 und Nvidia V100, der auf 1,5 PFlops veranschlagt wird. Dann sollen noch zwei Systeme mit Intel Xeon Phi und Fujitsu Post-K (auf ARM64-Basis) hinzukommen, jeweils etwa mit 0,5 PFlops. Alle Rechner sollen über das PRACE-Projekt europaweit genutzt werden.

Apropos IBM und Power9, man liege gut im Zeitplan, so IBMs HPC-Chef Dave Turek zu c't. Der Summit-Supercomputer fürs Oak Ridge National Lab soll gegen Ende des Jahres zu einem Viertel aufge-

baut sein, mit seinen Nvidia-V100-Karten dann vielleicht 40 bis 50 PFlops liefern. Der Vollausbau auf 4600 Knoten, ein jeder mit 40 TFlops, soll dann im nächsten Jahr erfolgen. Der mit einer theoretischen Spitzenleistung von 100 PFlops etwas kleinere Sierra fürs Lawrence Livermore National Lab soll ebenfalls 2018 in Betrieb gehen. Als erster Prozessor unterstützt Power9 auch PCIe 4.0 – als einzige Karte gibt es derzeit allerdings nur 200-GBit-InfiniBand von Mellanox.

ThinkSystem statt Thinkpad

ThinkSystem, so heißt ab jetzt die Serverlinie bei Lenovo. Sie wurde für die neue Xeon-Generation erstmals komplett von Lenovo entwickelt, ist also kein Erbstück mehr von IBM. Für die PC-Linie sieht Konzernchef Yang Yuanqing indes keine Zukunft mehr, Lenovo will jedoch Platz eins bei Datenzentren, Supercomputern und im mobilen Internet anstreben. Dabei mithelfen soll als Serverchef übrigens ein



IBMs HPC-Chef Dave Turek konnte keinen Power9-Chip zeigen, doch Power9-Architekt Shakti Kapoor hatte gleiche mehrere in seiner Tasche, sogar einen geöffneten (hier in Shaktis Hand).

guter alter Bekannter der Szene: Kirk Skaugen war 24 Jahre bei Intel und hatte dort die Serversparte und im letzten Jahr die Client Computing Group geleitet.

HPE und DelleMC werden das Terrain jedoch nicht kampflos räumen, hinzu kommen mit Inspur und Huawei weitere stark expandierende Konkurrenten aus China. Alle haben auch neue Multi-GPU-Serverssysteme im Angebot, für PCIe schon länger, nun auch für den SMX-Sockel mit Nvidia NVlink. HPE und Dell bieten Serverboards für vier Tesla P/V-100 an. Inspur hat ein System für acht, ähnlich Nvidias DGX-1, und dürfte vermutlich deutlich preiswerter sein. Nvidia zeigt derzeit überall schon Tesla V100 (Volta) als PCIe- oder SMX-Karte herum, in Betrieb sah man aber noch keine.

AMD war nur mit einem arg kleinen Stand vertreten und zeigte die neue Radeon Instinct mit Vega-Architektur. Zwei Epyc-Systeme standen arbeitslos auf dem Tisch zwei Tage verschlossen herum, die öffnete AMD erst am letzten ISC-Tag nach dem Stapellauf in Austin ganz unspektakulär – kein Sekt, kein nichts, keine Party ... Bei den Board-Herstellern Supermicro, Gigabyte, Tyan sah man ansonsten noch das ein oder andere Epyc-Board, nicht aber bei HPE oder Dell. „Wir haben noch keine Prozessoren von AMD gekriegt“, so verlaute es entschuldigend von HPE. Nicht einmal das geplante Storage-System Cloudline CL3150 konnte man bewundern. Und der Ex-HPLer Ed Turkel, jetzt HPC-Chef von Dell, deutete an, dass man ein PowerEdge-System mit Epyc wohl erst weit später im Jahr anbieten werde.

„Wir evaluierten auch AMDs neue Serverprozessoren, schließlich schauen wir auf eine jahrelange gute Zusammenarbeit zurück“, so auch Crays Chief Strategy Officer Barry Bolding zu c't. Cray hat nicht nur den schnellsten nichtchinesischen Rechner in der Top500-Liste – Piz Daint am CSCS –, sondern zum Auftakt der ISC gerade wieder einen dicken Auftrag aus Neuseeland vom nationalen Institut für Wasser- und Atmosphärenforschung an Land gezogen, zwei XC50 und ein CS400-Cluster im Wert von 18 Millionen Dollar. Beide Institute beschäftigen sich auch intensiv mit Klimaforschung, etwas, was in den USA ja nicht mehr so auf der Tagesordnung steht.

Ab Oktober will NEC neues Leben in die „Beschleunigerszene“ bringen und erste Muster vom Vektorrechner Aurora auf einer PCIe-Karte ausliefern. Er ist kein



Erfolgsverwöhnt: Das Siegerteam der Student Cluster Competition von der Tsingha Universität bekam den Preis von Prof. Thomas Sterling (rechts) überreicht.

wirklicher Beschleuniger, sondern läuft stand-alone und benutzt das gastgebende System nur für I/O. Wie groß der High-Bandwidth-Speicher on board sein soll, hat NEC immer noch nicht verraten („groß ...“). Auch die Zahl der Kerne ist noch etwas unscharf: „Höher als 4 und niedriger 16“, hatte Chefingenieur Shintaro Momose schon auf der CeBIT angedeutet, also vielleicht etwa 10. Mit über 150 GByte/s pro Kern kommt Aurora dann mindestens auf 1,5 TByte/s – zum Vergleich, Nvidia Tesla V100 hat 900 GByte/s.

Student Cluster Challenge

Das Team aus Boston musste kurzfristig absagen, so kämpften dann nur elf Studenten-Teams aus USA, China, Deutschland, Spanien, Schottland, Singapur und Südafrika um die Benchmark-Krone, wer die besten Ergebnisse im Rahmen des vorgegebenen Power-Limits von 3 kW schafft. Aber nicht nur Benchmarkkünste waren gefordert, zehn Prozent der Punkte wurden auch im Rahmen einer sehr anspruchsvollen Befragung vergeben.

Benchmarks wie Linpack, Tensorflow, HPCC, HPCG und so weiter standen auf dem Programm, und es war schon im Vorfeld klar, ohne mindestens eine Handvoll Nvidia Tesla P100-Karten hat man keine Chance. Des ungeachtet stürzte sich frohgemut, wenn auch chancenlos, das Team aus Barcelona mit einem Cavium-ThunderX-Cluster ins Rennen, wofür es auch den Publikumspreis erhielt. Ebenfalls ohne GPUs ging das All-Girl-Team von NERSC von verschiedenen US-Universitäten ins Rennen – damit kam es gegen die harte Konkurrenz nicht an. Das erfolgsverwöhnte Siegerteam der Tsingha Universität in China hatte zuvor schon die Asia Student Supercomputer Challenge (ASC) und die auf der SC16 gewonnen. Es wurde jetzt dicht verfolgt von dem Team aus dem Centre for High Performance Computing (CHPC) aus Kapstadt. Vom Kap kamen ebenfalls schon mehrmals die Sieger auf der ISC. Den besten Linpack schafften mit 37,07 TFlops aber die Studenten der Friedrich-Alexander-Universität aus Erlangen-Nürnberg, gut ausgestattet mit 12 Tesla P100-Karten.

(as@ct.de) **ct**

Anzeige