



Der Kern-Job-Dämon verteilt Daten auf die Speicher

Mit 32 Knotenrechnern arbeitet Suprenum auf der Industriemesse in Hannover

Mitten in der Halle 18 der Industriemesse Hannover, der Halle des Computer Integrated Manufacturing, steht er lebhaftig schräg gegenüber von IBM: Deutschlands erster selbstentwickelter Super-Rechner, der „Suprenum“ – knapp zwei Meter lang, ebenso hoch und 780 Kilo schwer (F.A.Z. vom 21. Februar). Drei blaßgraue Schränke strahlen im Glanz der Scheinwerfer auf dem Podest wie ein Monument gemeinsamen Schöpferwillens. Schade, daß Suprenum nicht schon auf der Cebit gezeigt werden konnte, aber er läuft erst seit einer Woche.

Suprenum ist ein Parallelrechnersystem: In zwei der drei Schränke arbeiten je 16 Rechner mit ihren eigenen Speichern – unansehnlich „Vektorknoten“ genannt. Dazu gibt es je Schrank Hilfseinschübe für Diagnose, eine Steuereinheit für ein oder zwei Ein-GigaByte-Plattenlaufwerke, die als Verlängerung der lokalen Speicher dienen können, und natürlich ein Einschub für die nötige Kommunikation. Im dritten Schrank befindet sich die eigentliche Verbindung zur Außenwelt und zu den externen Massenspeichern, ein Unix-Steuerrechner von Krupp-Atlas.

Die Leistung eines Parallelrechners kommt aus der Struktur, seiner „Architektur“. Bei klassischen technisch-wissenschaftlichen Hochleistungsrechnern werden auch mehrere Rechenwerke zugleich bemüht. Die greifen direkt auf riesige gemeinsame Speicherfelder zu. Für diese „speichergekoppelten“ Rechner braucht man nur die Rechenarbeit richtig aufzuteilen, nicht so sehr die Daten, denn bei Systemen mit gemeinsamem Speicher kann ein Rechner auch im Feld des anderen tätig werden. Es gibt automatische Compiler (Sprachübersetzer), die Rechengänge logisch aufteilen und parallelisieren. Aber ohne zuviel Zugriffs-

verzögerung, also untätiges Warten, bis der andere Rechner aus dem gemeinsamen Speicher bedient wurde, geht das nur bis zu vielleicht acht angeschlossenen Rechnern. Die muß man deswegen so höllisch schnell machen, daß in der Regel Wasserkühlung nötig wird. Mehr Leistung gibt's auf diesem Weg nicht.

Anders bei speicherunabhängigen Parallelrechnern wie Suprenum: Hier hat jeder Rechner seinen eigenen Speicher. Es können also theoretisch beliebig viele Einheiten parallel arbeiten, mit normalen Geschwindigkeiten und ruhig luftgekühlt. Der neue Rechner bringt bloße 10,5 Kilowatt Heizleistung – so kühl wie dieses Jahr war es zu einer Industriemesse noch nie, und die von Suprenum versprochene 15 Grad wärmere Abluft hätte man sich in Hannover wohl allgemein gewünscht.

Der Haken an speicherunabhängigen parallelen Rechnern ist, daß die Daten erst in diese Einzelspeicher geladen werden müssen (dazu bedient sich Suprenum eines „Kern-Job-Dämons“) und daß Berechnungen an den „Grenzen“ Datenaustausch mit den Nachbarrechnern nötig werden lassen. Das berechnete Feld darf ja auch bei einem Parallelrechner keine Bruchstellen zeigen, die Aufteilung muß unsichtbar bleiben. Die Programmierung des Suprenum (in besonderem Fortran) ist diffiziler als bei Multiprozessorrechnern mit gemeinsamem Speicher.

Schwierig wird es in einem Parallelrechner immer, wenn Daten umgelagert werden müssen – das Rechnen ist immer noch das Schnellste. Bei großen Datenfeldern müssen Informationen aus den 16 lokalen Speichern auf der Magnetplatte zwischengespeichert werden, ein möglicher „Flaschenhals“, obwohl die Rechner im Cluster mit zwei 64 Bit breiten und 160

Megabyte pro Sekunde schnellen „Bussen“ verbunden sind. Berechnungen mit Größen aus mehreren „Feldern“ müssen erst die Daten mit Nachbarrechnern austauschen. Am Anfang und zum Schluß muß alles herein- und hinausgespeichert werden. Und schließlich muß all das noch glatt laufen, weil ungleiche Belastungen durch ausgefallene Ressourcen das System arg ins Hinken brächten. Der Durchsatz eines so komplexen Systems liegt nicht an den einzelnen Rechenwerken allein.

Die Erwartungen sind hoch: Jeder der in Hannover integrierten 32 Einzelrechner soll nominal 20 Megaflops leisten – ein Megaflop sind eine Million „floating point operations“ oder Gleitkommarechnungen pro Sekunde. Dafür ist er großzügig bestückt mit einem zentralen Mikroprozessor (Motorola MC68020 mit 20 MHz Takt und MC68882 skalarer Gleitkommaarithmetik), einem Gleitkomma-Vektorrechner mit Gate-Array-Chips von Weitek (WTL 2264 und 2265), bis zu 8 Megabyte eigenem Speicher und einem Kommunikationsrechner zu den anderen Einzelrechnern und zur lokalen Platte. Dort können auf Wunsch periodisch Zwischenzustände abgespeichert werden (Checkpoints), damit bei Fehlern nicht wieder ganz von vorne begonnen werden muß.

Was Wunder, daß einem stolz der mattsilberne Glanz der Recheneinschübe entgegenstrahlt, wenn vorne einer der großen Schränke geöffnet wird. Hinten aber, nur von Eingeweihten beachtet, sieht man ganz in Gold die schnellen Daten-Autobahnen, die aus diesen Einzelrechnern erst den parallelen „Superrechner“ machen. Ganz neue Software ist nötig, um darauf zu fahren, um wirksam zu parallelisieren. Ob Suprenum sein Betriebssystem deshalb „Peace“ nennt? FRITZ JÖRN