

Nicht mehr die Nummer eins

■ So gern man an tradierten Gewissheiten festhalten möchte, KI reißt einfach alles um und erschafft die Welt neu. Die Rede ist nicht von KI-Artefakten wie Bildern mit sechs-fingrigen Händen – Polydaktylie war schon immer Teil der Evolution. Nein; KI schickt sich an, Intels Vorherrschaft zu beenden. Doch wie konnte es so weit kommen?

Den Anfang machten wohl die vielen Entwicklungsschwierigkeiten bei den HPC-GPUs Intel Phi, für alle sichtbar am Vorzeige-Exascalor Aurora, der einfach nicht fertig werden wollte. Ihnen folgten etliche Verzögerungen beim Nachfolger Ponte Vecchio. Nun kommt der riesige, aus fast 50 Chiplets zusammengesetzte Beschleuniger bei KI-Anwendern nicht einmal gut an. Bereits im März 2023 verschrottete Intel dessen geplanten Nachfolger und verrannte sich in der Idee einer XPU, eines CPU-GPU-Zwitters namens Falcon Shores, um dann festzustellen, dass die Welt noch nicht reif ist für eine XPU und der Markt „viel dynamischer als gedacht“, sprich Nvidia ihnen davongaloppiert ist. Bis Falcon Shores 2025 nun als diskrete GPU erscheint, haben Nvidia mit dem Hopper-Nachfolger Blackwell und AMD mit der Instinct MI300X, der einzigen H100-Konkurrenz in Sichtweite, das Feld längst bestellt.

Auch bei den Server-CPU's strauchelt Intel inzwischen. Was KI-Anwender wollen, sind hoher Durchsatz, geringe Latenz, Architekturen für parallelisierbare Aufgaben – und noch mehr Durchsatz. Das liefert derzeit nur AMD: 128 PCIe-Lanes und 64 CXL-Lanes mit einer CPU treten da gegen 80 PCIe-Lanes bei Intel an. Nun zieht die SPEC über 2600 offizielle Ergebnisse des Benchmarks SPECint 2017 für Systeme mit Xeon-SP-CPU's Gen4 Sapphire Rapids wegen illegaler Tricksereien ein. Und auch der PC-Markt schrumpft und schrumpft und schrumpft.

Als der Chip-Riese noch die unangefochtene Nummer eins war und die Regeln diktieren konnte, hat er genau das getan: Kaum spürte der Zwerg AMD mit den Opteron-CPU's leichten Aufwind unter den Flügeln, drohte Intel selbst großen Systemhäusern mit Nichtbeliefern, sollten sie es wagen, Opteron-Server ins Programm aufzunehmen. Mit solchem Gebaren macht man sich keine Freunde – und hat dann auch keine in der Not. Und die ist groß, denn erstmals könnte AMDs Serversparte 2024 an der viel größeren von Intel vorbeiziehen. Beide steuern nämlich auf einen Quartalsumsatz von 3,5 Milliarden US-Dollar zu, nur der eine von unten, der andere von oben. Viel Mitleid dürfte der schwankende Riese angesichts seiner langjährigen Knebelpraktiken aber nicht ernten.

Bleibt noch die Auftragsfertigung, die Intel nun ausbaut, unter anderem für Nvidias H100-GPUs. Große Gewinnspannen gibt es hier aber nicht, und erst recht keinen Glamour, wie ihn die drei wertvollsten Unternehmen Apple, Microsoft und Nvidia versprühen. Der einzige Trost: Die Regierungen in Europa und den USA locken mit spendablen Beihilfen für neue Chipfabriken. Erfahrungen mit Chipknappheiten während der Coronapandemie, Störungen der Schifffahrt – erst im Suezkanal, nun im Roten Meer –, Handelsstreitigkeiten zwischen den USA und China, geopolitische Spannungen rund um Taiwan, der Heimat des weltgrößten Auftragsfertigers TSMC, und nun erwartbare Lieferengpässe vor allem bei KI-Chips, wie sie Nvidia bereits vorhersagt: Das alles lässt die Regierungen großzügig werden gegenüber jedem, der bereit ist, eine Chipfabrik auf ihrem Territorium zu bauen.

Von den 20 Milliarden Euro der deutschen Regierung hat sich Intel auch gleich die Hälfte für den neuen Standort in Magdeburg gesichert, doch Globalfoundries, TSMC, Infineon und andere sagen da auch nicht Nein. Auch auf dem Feld der Chipfertiger ist Intel nicht allein und ganz sicher nicht die Nummer eins.

Susanne Nolte

SUSANNE NOLTE

