FORSCHUNGSZENTRUM JÜLICH GmbH
Zentralinstitut für Angewandte Mathematik
D-52425 Jülich, Tel. (02461) 61-6402

Interner Bericht

Perspektiven für Supercomputer-Architekturen

Friedel Hoßfeld

FZJ-ZAM-IB-2001-13

September 2001

(letzte Änderung: 17.09.2001)

FZ Jülich – Zentralinstitut für Angewandte Mathematik
Interner Bericht – 17. September 2001

Perspektiven für Supercomputer-Architekturen

F. Hoßfeld

Forschungszentrum Jülich
John von Neumann Institut für Computing
Zentralinstitut für Angewandte Mathematik
Rheinisch-Westfälische Technische Hochschule
Lehrstuhl für Technische Informatik und Computerwissenschaften

1 Exponentielle Entwicklungen: Moore’sches Gesetz und TOP500

Zum ersten Mal in der Geschichte des Computers bietet sich die Gelegenheit, eine ausgewogene Pyramide der Rechnerleistung für das Wissenschaftliche Rechnen – vom Arbeitsplatzrechner über Workstation- und PC-Cluster und Parallelrechner bis zur Spitzenklasse des Supercomputer und deren Grid-Verbund über Breitbandnetze – aufzubauen und fortzuentwickeln /1/. In den ersten Jahrzehnten war die Entwicklung des Höchstleistungsrechnens von der numerischen Lösung partieller Differentialgleichungen dominiert, die über die Diskretisierung von Raum und Zeit und Linearisierung in die Lineare Algebra und ihre numerischen Konzepte und Algorithmen führt. Die Antwort der Rechnerarchitektur auf diese frühen Herausforderungen waren die Vektorrechner, die das Pipeline-Prinzip optimierten und effektive Werkzeuge für die Vektorisierung der Programme schufen /2/. Die Ausschöpfung ihres Leistungspotentials durch Programmiermethoden und Compilertechniken, Softwarewerkzeuge, Betriebssystemfunktionen und gemeinsamen Hauptspeicher im Übergang zum Parallelrechnen resultierte in einem effizienten Arsenal an Wissen und Erfahrung über die Stärken und Schwächen des Vektorrechnens. Die höchste Klasse der Vektorrechner – durchgängig als Multiprozessorsysteme ausgelegt – von Fujitsu, NEC, Hitachi und Cray sind immer noch „Arbeitspferde“ in vielen Produktionsumgebungen des Höchstleistungsrechnens in Wissenschaft und Wirtschaft. Obwohl im vergangenen Jahrzehnt die effektiv aus Vektorrechnern erzielbare Leistung (*sustained performance*) mit den massiv-parallelen Systemen mittlerer Größe bei vielen wichtigen Anwendungen gut mithalten konnte, hat die technologische Entwicklung bei den Mikroprozessoren und Speicherchips die Vektorrechner technisch und im Preis/Leistungsverhältnis sowie hinsichtlich der Spitzenposition in der Leistungspyramide zunehmend in den Hintergrund gedrängt. Das exponentielle Wachstum des Moore’schen Gesetzes für Mikroprozessoren und Speicherchips (Abb. 1) bestimmt auch bei den Supercomputern weitgehend die Architekturen und den Markt: „Commodities“ setzen proprietäre Prozessortechnologien unter Druck.

Jedoch muß festgestellt werden, daß aus der absehbaren Entwicklung erkennbar ist, daß Workstations und PCs, so leistungsfähig sie auch werden, das Potential massiv-paralleler Höchstleistungsrechner mit ihrem engen Verbund von starken Mikroprozessoren über Breitband-Verbindungsnetzwerke nicht erreichen können (Abb. 2). Die Schere zwischen der Leistungsfähigleit einzelner Workstations und PCs und den Höchstleistungsrechnern scheint vielmehr weiter aufzugehen. Die Performance der Supercomputer ist mit dem durchschnittlichen Faktor von fast 2 pro Jahr in den letzten Jahren sogar exponentiell stärker gewachsen, als das Moore'sche Gesetz bei den Mikroprozessoren aufzeigt (Abb. 3). Die Daten der TOP500-Listen /5/ weisen so einen geradezu dramatischen Wettbewerb aus. Massiv-parallele Rechner werden daher als die – derzeit einzige – Architektur angesehen, um dem Bedarf der großen Anwendungen (*Grand Challenges*) in Naturwissenschaft und Technik, in zunehmendem Maße aber auch in den expandierenden Anwendungsfeldern des Supercomputing in der Wirtschaft, wie Logistik und Börse, gewachsen zu sein.

2 Der ASCI-Impuls

Ein starker Impuls für die Weiterentwicklung gekoppelter SMP-Systeme geht vom ASCI-Programm (*Accelerated Strategic Computing Initiative*) des amerikanischen Department of Energy aus /4–5/. In diesem Programm werden Rechner der höchsten Leistungsklasse insbesondere für die Simulation von physikalischen und chemischen Prozessen zur Garantie der Sicherheit und Einsatzfähigkeit der Nuklearwaffen der USA entwickelt. Die Leistungsanforderungen zielen auf 10- bis 100-TeraFlops-Rechner für die Jahre 2001–2004 (TeraFlops: 10^{12} Gleitkomma-Operationen pro Sekunde). Diese Parallelrechner werden aus allgemein verfügbaren Komponenten bestehen; insbesondere sind die Rechenknoten sogenannte SMP-Systeme (s. Kap. 3). Die Entwicklung der Software für diese gekoppelten Systeme wird intensiv von den entsprechenden wissenschaftlichen Einrichtungen in USA zusammen mit ausgewiesenen Hochschulen betrieben. Diese Arbeiten werden das wissenschaftliche Hochleistungsrechnen primär in den USA, schließlich aber insgesamt befruchten. Allerdings sind wegen der Komplexität der Hardware-/Software-Konfigurationen, des durch die Entwicklung der Informationstechnik vorgegebenen raschen Wandels und auch der letztlich militärisch orientierten Zielsetzung des ASCI-Programms von dort keine allgemeinen und endgültigen Lösungen für die Software-Probleme bei gekoppelten SMP-Systemen zu erwarten.

Eine zweite in die Breite wirkende amerikanische Aktivität wurde von der National Science Foundation im Rahmen des PACI-Programms (*Partnerships for Advanced Computational Infrastructure*) angestoßen /6/. Mit neuer Strategie entstanden um das San Diego Supercomputer Center und das National Center for Supercomputing Applications in Urbana, Illinois, schlagkräftige Verbünde zur Förderung des Wissenschaftlichen Rechnens. Hier wird ein breites Spektrum an Know-how und konkreter Software aufgebaut, mit dem neue Rechnerarchitekturen weit jenseits der TeraFlops-Grenze betrieben und genutzt werden können. Grid Computing heißt das neue Paradigma /1/.

Des weiteren zeichnet sich ab, daß die Hersteller von Höchstleistungsrechnern künftig zusammen mit den Rechnern nur Basiskomponenten der für den Betrieb und die Programmierung notwendigen Systemsoftware (Middleware) liefern werden. Um die zur effizienten Nutzung der Systeme unzureichende Software um die notwendigen Komponenten zu erweitern, gibt es deshalb in den USA ab dem Jahr 2000 eine breite staatliche Förderung von Software-Entwicklungsprojekten bei amerikanischen Forschungszentren und Universitäten /7/. Dies ist konsequent angesichts der schon jetzt geübten Praxis, bei der die Hersteller mit strategisch

ausgerichteten Zentren auch international eine intensive wechselseitige Zusammenarbeit gepflegt haben, die insbesondere für die Anwender sehr fruchtbar war. Diese Zusammenarbeit ist gerade auch mit deutschen Partnern aus Forschung, Hochschulen und Industrie im Bereich des High Performance Computing etabliert und muß erhalten werden.

Trotz der Zielsetzung eines Metacomputing-Verbundes zwischen den Supersystemen der US National Laboratories Los Alamos, Livermore und Sandia entsprechend den ASCI-Pathforward-Plänen reichen die Anforderungen des ASCI-Programms weit über die verfügbare Technologie für Rechnerarchitekturen hinaus. Die Überlegungen für zukünftige Rechnerarchitekturen der PetaFlops-Klasse (PetaFlops: 10^{15} Gleitkomma-Operationen pro Sekunde) sind in vollem Gange /8/.

Wie im ASCI-Programm dargelegt, stellen ausgewogene Rechnersysteme im Bereich von 10 bis 100 TeraFlops unabdingbare Anforderungen an die Prozessorleistung, die Prozessorarchitektur, die Speicherorganisation, die Verbindungsnetzwerke und die Kommunikation sowie die I/O-Systeme. Daraus folgen für die verschiedenen Leistungsparameter diese Skalierungen:

1 TeraFlops Peak-Performance/
1 TeraByte Hauptspeicher/
50 TeraByte Plattenspeicher/
16 TeraByte/sec Cache-Bandbreite
3 TeraByte/sec Speicher-Bandbreite/
0.1 TeraByte/sec I/O-Bandbreite/
10 GigaByte/sec Disk-Bandbreite/
1 GigaByte/sec Archivspeicher-Bandbreite/
10 PetaByte Archivspeicher.

Weltweit warten die Supercomputer-Zentren auf die Umsetzung der ASCI-Maschinen in den Markt, um an dem Technologie- und Leistungssprung im Höchstleistungsrechnen teilhaben zu können, damit sie die Computersimulation in Forschung und Industrie weiterentwickeln zu können. Es ist nicht klar, wann außer der Hardware vom Markt die Vielfalt der Ergebnisse aus dem ASCI-Programm offen verfügbar sein werden. Deshalb ist es unverzichtbar, umgehend eigene Initiativen zur Entwicklung der für die neue Rechnerarchitektur erforderlichen Software-Funktionalität einschließlich Betriebssystemkomponenten, Programmiermodellen, Programmierwerkzeugen und Programmbibliotheken mit numerischen Methoden und nichtnumerischen Algorithmen zu starten /9, 10/.

Mit den Leistungssprüngen werden aber die Supercomputer-Zentren und ihre Anwender zunächst mit ernsten Problemen der Stabilität und Zuverlässigkeit der Systeme konfrontiert werden. Hinzu treten aufgrund der nunmehr möglichen Behandlung bislang unerreichbarer Problemgrößen in den Anwendungen ganz neuartige Anforderungen an die Verifizierbarkeit und Validierbarkeit der mit diesen Supercomputern erzielbaren Ergebnisse. Hier stellt sich auch die Frage der Rechnerarithmetik mit automatischer Ergebnisverifikation von neuem, wie sie seit langem angemahnt wird /11, 12/. Auch sollten zunehmend für Computerexperimente dieser neuen Größenordnungen die Methoden des *Experimental Design*, die in anderen experimentellen Disziplinen zur Optimierung der gewinnbaren Information und der Experimentkosten genutzt werden, besser erschlossen werden /13/. Diese Anstöße führen aber unabdingbar auch zu neuen Anforderungen an die Aus- und Weiterbildung einschließlich der Modifikation oder Neukonzeption von Hochschul-Curricula.

3 Strukturen und Komponenten der SMP-Architekturen

Die in den vergangenen Jahren im wissenschaftlichen Hoch- und Höchstleistungsrechnen vorherrschende Rechnerklasse ist die der Parallelrechner mit verteiltem Speicher. Diese Rechnerklasse, auch *MPP-Systeme* (*massively parallel processors*) genannt, wird meist nach dem Programmiermodell des *Message Passing* genutzt. Eine wachsende Zahl von Anwendungen konnte erfolgreich für MPP-Systeme parallelisiert werden, vorwiegend aus dem wissenschaftlichen Bereich. Der Einsatz von MPP-Systemen in der industriellen Praxis war hingegen längere Zeit eingeschränkt. Dies war einerseits in der vergleichsweise schwierigen Portierung von Anwendungen nach dem Message-Passing-Modell begründet, andererseits folgten die Software-Häuser dem Trend zu MPP-Systemen nur zögernd, so daß für die Industrie wichtige Standard-Software auf sich warten ließ.

Demgegenüber haben Parallelrechner mit gemeinsamem Speicher ständig wachsende Bedeutung erlangt. Auf dem Markt werden sowohl Intel- als auch RISC-basierte Systeme angeboten. Es gibt verschiedene Techniken für den Bau solcher Systeme, die es erlauben, daß derzeit bis zu 256 Prozessoren logisch auf einen gemeinsamen Speicher zugreifen können. Diese Klasse von Rechnern wird als *SMP-Rechner* (*symmetric multiprocessor systems, shared memory multiprocessors)* bezeichnet /10/.

Parallelrechner mit gemeinsamem Speicher sind zu Standardsystemen für Simulationsrechnungen im mittleren Leistungsbereich geworden. Solche Rechner werden bereits heute von vielen Computer-Herstellern angeboten, z. B. von Compaq, Hitachi, HP, IBM, NEC, SGI, Sun, auf der Basis proprietärer, d.h. aus eigener Entwicklung produzierter Prozessoren (wie Alpha-, Power-, MIPS-, Sparc-Chip) oder auch von vielen anderen Firmen auf der Basis von Intel-Prozessoren. Die SMP-Rechner haben im Gegensatz zu den speziell für Wissenschaft und Technik entwickelten homogenen MPP-Systemen ein günstigeres Preis/-Leistungsverhältnis und werden auch für kommerzielle Anwendungen breit genutzt. Es ist zu erwarten, daß diese Rechnerklasse nahezu alle Anwendungsbereiche des Computing, insbesondere auch die kommerziellen Anwendungsfelder, beherrschen wird. Damit werden die Parallelrechner die dominante Rechnerstruktur sein.

Im wissenschaftlichen Computing werden SMP-Rechner heute für numerische Simulationsaufgaben mit mittleren Leistungsanforderungen sowohl in der industriellen als auch in der akademischen Forschung eingesetzt. Aufgrund des einfacheren Programmiermodells des *Shared Memory* gibt es für sie eine verhältnismäßig große Basis an parallelisierter Software, die ihre Wurzeln zum Teil in Entwicklungen für Cray-Mehrprozessor-Vektorrechner hat. Unter anderem die chemische Industrie nutzt zur Modellierung von Abläufen und zur Berechnung von Stoffeigenschaften in hohem Maße SMP-Rechner. Ein Einsatzschwerpunkt von SMP-Rechnern liegt im kommerziellen Bereich als Datenbank-Server. Es ist zu erwarten, daß die hohen Leistungsanforderungen des kommerziellen Marktes, der zudem durch großes Umsatzvolumen und starken Wettbewerb gekennzeichnet ist, dazu führen werden, daß zunehmend größere und leistungsfähigere SMP-Rechner verfügbar sein werden.

Im Gleichklang mit dieser Entwicklung werden spezielle Parallelrechner für den vergleichsweise kleinen technisch-wissenschaftlichen Markt vom Preis-/Leistungsverhältnis her nicht mehr lohnend produziert werden können und an Bedeutung im Markt verlieren. Einzelne SMP-Rechner werden allerdings trotz der erwarteten Leistungssteigerungen auch in Zukunft nicht in der Lage sein, die ständig wachsende Nachfrage nach sehr hoher Rechenleistung zu decken. Daher greifen die Hersteller für diesen Markt auf die Technik des

Clustering zurück und bieten als Systeme für die höchste Leistungsklasse gekoppelte SMP-Systeme an. Dies geschieht z. B. im amerikanischen ASCI-Programm /4/. Auf Grund der Vielzahl von technischen Lösungen zur Kopplung von SMP-Rechnern ist es schwierig, die Architekturen in Klassen einzuteilen. Hier soll vorerst nur die folgende Unterscheidung getroffen werden: Systeme aus SMP-Rechnern, die eng über ein dediziertes Netz gekoppelt sind und sich dem Benutzer als einheitliches System darstellen, bezeichnet man als *Parallelrechner mit SMP-Knoten.* Systeme aus SMP-Rechnern, die lose, z. B. über ein nicht-dediziertes LAN gekoppelt sind, heißen *vernetzte SMP-Rechner* /10/.

Parallelrechner mit SMP-Knoten für die höchste Leistungsklasse sind bereits als Produkte erhältlich und werden von den im technisch-wissenschaftlichen Markt aktiven Herstellern verstärkt angeboten. Vernetzte SMP-Rechner sollten insbesondere im industriellen Umfeld eine hohe Bedeutung haben, wo häufig mehrere über ein Intranet miteinander verbundene SMP-Rechner vorhanden sind. Hier sollte es beim Auftreten hoher Anforderungen an Rechenzeit zukünftig vermehrt möglich sein, diese Rechner temporär für einzelne Anwendungen zu einem größeren System zu verbinden. Gekoppelte SMP-Systeme führen eine zusätzliche Hierarchiestufe in die Rechnerstruktur ein und stellen daher für die Programmierung eine neue Herausforderung dar. Als Programmiermodell für diese Systeme könnte man ein hybrides Modell wählen: Innerhalb jedes einzelnen SMP-Rechners wendet man das Modell des gemeinsamen Speichers an und zwischen den SMP-Rechnern das Modell des Message Passing. Für beide Programmiermodelle gibt es weitgehend akzeptierte Sprach-erweiterungen oder Kommunikationsbibliotheken wie OpenMP, MPI oder HPF. Diese können aber nicht immer koexistieren, und es ist vielleicht sinnvoller, eines der Modelle auf das gesamte gekoppelte SMP-System zu übertragen. Neben dem adäquaten Programmier-modell fehlen weitgehend die Werkzeuge zur Programmentwicklung, wie z. B. zur Leistungs-analyse und -optimierung. Das Daten-Management, insbesondere die parallele Ein-/Ausgabe, stellt einen weiteren Problemkreis dar. Die Kopplung von SMP-Rechnern führt hier wie auch bei der Ressourcenverwaltung und beim Nutzungsmodell zu einer Steigerung der ohnehin schon hohen Komplexität.

Die Entwicklung paralleler mathematischer Software konzentriert sich bislang entweder auf Rechner mit gemeinsamem Speicher oder auf MPP-Systeme mit verteiltem Speicher. Software-Bausteine für die Verteilung der Daten und den Zugriff auf diese haben eine besondere Bedeutung. Es gibt zwar Basisbibliotheken, welche Speicherhierarchien, insbesondere Caches, berücksichtigen, es ist aber definitiv zu klären, ob diese Konzepte auf gekoppelte SMP-Systeme durchgängig zu übertragen sind und wie ihre möglichst effiziente Implementation aussehen kann. Schließlich ist offen, welche Strategien bei der Entwicklung bzw. Portierung von kompletten Anwendungen verfolgt werden sollen, z. B. bezüglich der Datenverteilung oder Lastbalance. Es ist leider zu erwarten, daß die Herstellerfirmen aus eigenen Kräften nur begrenzte Anstrengungen unternehmen können, die Programmentwicklungs- und Programmausführungsumgebungen für das Höchstleistungsrechnen auf Parallelrechnern mit SMP-Knoten auf den erforderlichen Stand zu bringen. Dies gilt in noch höherem Maße für die mathematische Basissoftware und für Anwendungen, deren Entwicklung bereits bei der Rechnergeneration der MPP-Systeme von den Rechnerherstellern weitgehend an Softwarehäuser und die Wissenschaft abgegeben wurde. Dabei schließt das Hochleistungsrechnen neben den historisch gewachsenen und weiter wichtigen technisch-wissenschaftlichen Anwendungsgebieten vermehrt auch kommerzielle Software ein, die für den Standort Deutschland schließlich von hoher Bedeutung ist. Das Thema der vernetzten SMP-Rechner hat viele Berührungspunkte mit dem Metacomputing und Grid Computing, das in Deutschland im BMBF-geförderten Projekt zur Entwicklung von UNICORE /14/ und im EU-geförderten Grid-Projekt EUROGRID /15/, in den in den vergangenen Jahren über den DFN

geförderten Gigabit-Testbed-Projekten /16,17/ oder auch in dem durch das Land NRW geförderten, inzwischen gleichfalls abgeschlossenen Projekt HPCM /18/ untersucht wird. In diese Richtung zielen auch die amerikanischen Projekte zum Grid Computing /1/, bei denen die breite Verfügbarkeit und leichte Nutzung einer nationalen Computing-Infrastruktur angestrebt wird.

Zur Erzielung hoher und höchster Performance werden SMP-Rechner über ein dediziertes, leistungsfähiges Netzwerk fest zu einem Hochleistungssystem verbunden (*Parallelrechner mit SMP-Knoten*) oder ständig oder nur temporär über Ethernet oder ein allgemeines Local Area Network oder Wide Area Network zu einem Rechnerverbund zusammengeschaltet (*vernetzte SMP-Rechner*). Für beide Architekturtypen verwendet man die gemeinsame Bezeichnung *gekoppelte SMP-Systeme*. Diese Architekturen sind noch besonders hinsichtlich der Software in der Entwicklung, und die neue Hierarchiestufe in der Rechnerstruktur führt bei der Anwendungsentwicklung zu einer erhöhten Komplexität, deren Beherrschung jedoch sowohl für den breiteren Einsatz vernetzter SMP-Rechner als auch für die effiziente Nutzung künftiger Höchstleistungsarchitekturen unumgänglich ist. Forschung und Entwicklung sowie erzielbare Verbesserungen von Software-Komponenten im obersten Segment des High Performance Computing werden jedoch wegen der grundsätzlich möglichen Skalierbarkeit aus den SMP-Bausteinen unmittelbar Einfluß auf die SMP-Komponenten und damit auf die Systeme im breiten kommerziellen Markt haben.

Die Technologie stellt mit den heutigen Mikroprozessoren sehr leistungsfähige Basisbausteine sowohl für Workstation-Arbeitsplätze als auch für Höchstleistungsrechner zur Verfügung, die über eine beeindruckende Spitzenleistung verfügen. Mit diesen Zahlen wird vielen Nicht-Fachleuten – auch den Medien und den Entscheidungsträgern in Wirtschaft und Industrie – das Gefühl vermittelt, daß die Performance-Optimierung immer unwichtiger wird und zumindest die nächste Chip-Generation jeden heute beobachtbaren Performance-Flaschenhals beseitigen wird. Seit dem Einsatz von MPP-Maschinen stehen dieser Erwartungshaltung jedoch andere Erfahrungen entgegen. Die erzielte Leistung für reale Anwendungsprogramme liegt in sehr vielen Fällen im geringen Prozentbereich der Spitzenleistung (5 bis 20 Prozent), und nur für sehr wenige auf die Architektur angepaßte hochoptimierte Programme stellt sich in etwa der Geschwindigkeitsvorteil ein, den der Zuwachs der Spitzenleistung verspricht. Der Grund für diese nachteilige Entwicklung liegt zu einem wesentlichen Teil in der Komplexität der Architektur moderner Mikroprozessoren und der Speicherorganisation. Sowohl die verfügbare als auch in naher Zukunft zu erwartende Systemsoftware der Hersteller wird die parallelen Funktionseinheiten und hochkomplexen Speicherhierarchien nur selten in adäquatem Maß nutzen und damit das volle Leistungspotential ausschöpfen können.

Hier liegt eine Kernaufgabe für die europäischen Forschungs- und Entwicklungsarbeiten. Während bei der Entwicklung und Produktion von Prozessorchips – zumindest im Bereich von universell einsetzbaren Rechnern – Europa keine Rolle mehr spielt, erscheint es sehr erfolgversprechend, den Bereich der Software – und hier insbesondere die Methodenentwicklung für die Programmierung und die Performance-Optimierung – zu fördern und damit einen hohen Wertschöpfungsfaktor zu erzielen. Da in der nächsten Generation der Hochleistungsrechner – entgegen der bisherigen Praxis bei Vektor- und massiv-parallelen Systemen – allgemein verfügbare Basis-Komponenten verwandt werden, stehen die im Bereich des Hochleistungsrechnens entwickelten Methoden und Werkzeuge unmittelbar den industriellen Anwendern zur Verfügung. Dies ist eine wesentliche Änderung der bisherigen Situation, und es kann bei einer Investition in dieses Gebiet ein kräftiger Innovationsschub auf dem Gebiet der Modellierung und Simulation auch in der deutschen Wirtschaft erwartet werden.

Parallelrechner sind Computer mit mehreren Prozessoren, wobei die Prozessoren gemeinsam, d. h. koordiniert, an einer oder mehreren Aufgaben arbeiten. Es gibt unterschiedliche Typen von Parallelrechnern, wobei eine wichtige Klassifizierung die hinsichtlich ihrer Speicherarchitektur ist. Die Speicherarchitektur eines Parallelrechners hat Auswirkungen auf die Programmiermodelle, die auf diesem Rechner sinnvoll und effizient anwendbar sind. Wichtige Klassen sind nachfolgend beschrieben:

- In der Klasse der UMA-Rechner (*Uniform Memory Access*) haben alle Prozessoren gleichberechtigten und im Prinzip gleichschnellen Zugang zu allen Speicheradressen des gemeinsamen Adreßraums. Beispiele: CRAY T90, IBM 390, SUN E10000.

- In NUMA-Rechnern (*Non Uniform Memory Access*) existiert ebenfalls ein gemeinsamer Adreßraum für alle Prozessoren, die Zugriffszeit auf verschiedene Speicherstellen kann jedoch auf einem Prozessor unterschiedlich groß sein: Der Zugriff auf eine „nahe“ Speicherstelle kann schneller erfolgen als der Zugriff auf eine „entfernte“ Speicherstelle.

- Rechner der Klasse ccNUMA (*cache coherent Non Uniform Memory Access*) sind NUMA-Rechnern ähnlich, haben damit wiederum einen gemeinsamen Adreßraum für alle Prozessoren, können jedoch den Inhalt entfernter Speicherstellen lokal puffern, so daß nach einem ersten Zugriff weitere Zugriffe auf die gleiche Speicherstelle eventuell schneller erfolgen können. Dies geschieht kohärent zu den Originaldaten und kohärent zu eventuell vorhandenen weiteren Kopien. Dort wirkt sich eine Modifikation der Originaldaten auch auf die Kopie bzw. eine Modifikation einer Kopie ebenfalls auf die Originaldaten aus. Beispiel: SGI Origin2000, HP V-Class.

- In der Klasse der NORMA-Rechner (*No Remote Memory Access*) haben die Prozessoren keinen gemeinsamen Adreßraum, auf den alle Prozessoren direkt zugreifen können. Jeder Prozessor besitzt seinen lokalen Speicher und Adreßraum, auf den andere Prozessoren nicht direkt zugreifen können. Beispiele: CRAY T3E, IBM RS/6000-SP, PC-Cluster.

Als *SMP-Rechner* werden nunmehr solche Parallelrechner bezeichnet, bei denen alle Prozessoren unter der Verwaltung eines Betriebssystems stehen, auf einen gemeinsamen Adreßraum zugreifen können und aus Betriebssystemsicht (und aus Sicht des Anwendungsprogrammierers) als gleichberechtigt angesehen werden. Es ist bei SMP-Rechnern also z. B. möglich, einen vom Betriebssystem auf einem Prozessor suspendierten Prozeß auf einem anderen Prozessor fortzuführen. Dabei ist es aufgrund des gemeinsamen Adreßraumes nicht nötig, den vom Prozeß benutzten Speicher zu kopieren oder zu verschieben.

In einem Parallelrechner mit NORMA-Architektur bezeichnet man als *Knoten* eine Basiseinheit bestehend aus einem oder mehreren Prozessoren einschließlich Speicher und eventuell weiteren Komponenten. Diese Basiseinheiten werden über ein Verbindungsnetzwerk miteinander verbunden, über das sie miteinander kommunizieren können. *Parallelrechner mit SMP-Knoten* sind solche Parallelrechner, in denen die Knoten selbst SMP-Rechner sind, die über ein dediziertes Netz miteinander verbunden sind, um gemeinsam Aufgaben zu bearbeiten. *Vernetzte SMP-Systeme* sind dagegen solche Parallelrechner, in denen SMP-Rechner über ein nicht-dediziertes Verbindungsnetzwerk (System Area Network, Intranet, Internet) miteinander verbunden sind. *PC- bzw. Workstation-Cluster*, im weiteren einheitlich als PC-Cluster bezeichnet, sind Spezialfälle von Parallelrechnern, die aus handelsüblichen

Komponenten (COTS: *Commodity Off The Shelf*), meist aus dem PC-Bereich, aufgebaut werden können. Auch hier sind als Knoten des PC-Clusters SMP-Komponenten (Mehrprozessor-Boards) möglich. PC-Cluster erlauben durch die Verwendung preisgünstiger PC-Technologie den Aufbau kostengünstiger Parallelrechner, jedoch müssen unter Umständen Zugeständnisse an Effizienz, Zuverlässigkeit u.ä. gemacht werden.

Unabhängig von der Art der hier betrachteten Rechner (Parallelrechner mit SMP-Knoten, vernetzte Parallelrechner, PC-Cluster) werden die einzelnen Knoten eines solchen Rechners über ein Verbindungsnetzwerk miteinander verbunden, über das der Datenaustausch zwischen den Knoten erfolgt. Die Leistungsfähigkeit eines solchen Netzwerkes hat wesentlichen Einfluß auf die Leistung der Programme, die auf diesem Rechner ablaufen.

In einem statischen Netzwerk sind die Knoten fest mit anderen Knoten (Nachbarn) verbunden. Beispiele für Topologien statischer Netze sind Stern, Ring, Gitter, Hypercube, Fat Tree u. a. Dynamische Netze sind eventuell mehrstufig aus Schaltelementen aufgebaut. Die Eingänge und Ausgänge des Schaltnetzwerks sind jeweils mit den Knoten verbunden (Permutationsnetzwerk, Omega-Netzwerk), und beim Verschicken einer Nachricht von einem Senderknoten A zu einem Empfängerknoten B werden die Schalter so gesetzt, daß ein Weg im Netzwerk vom Eingang A zum Ausgang B dynamisch geschaltet wird. Beispiele dynamischer Netze sind Kreuzschienenverteiler, Omega-Netze oder Benes-Netze. Zu einem Verbindungsnetzwerk läßt sich eine Reihe von Metriken /19/ angeben, die prinzipielle Aussagen über die Leistungsfähigkeit des Netzwerkes erlauben.

Die Vernetzung von Rechnern wird schon seit längerer Zeit in LANs (*Local Area Network*) und WANs (*Wide Area Network*) betrieben, und Standardkomponenten sind dafür erhältlich, die prinzipiell auch für ein Verbindungsnetzwerk in einem Parallelrechner mit SMP-Knoten verwendet werden können. Neben diesen auf die Bedürfnisse eines LANs oder WANs ausgerichteten Komponenten gibt es auch dedizierte Netze, die speziell für Parallelrechner entwickelt wurden (SAN: *System Area Network*). Aus dem Bereich der Local Area Networks sind einige Netzwerke bekannt, die weitverbreitet im Einsatz sind. Diese Technologien werden auch zum Aufbau von vernetzten SMP-Systemen und PC-Cluster genutzt. Zum Einsatz kommen hier vorwiegend Ethernet in verschiedenen Leistungsvarianten, FDDI, ATM und HiPPI. Der Nachteil der oben genannten Standardtechniken ist häufig, daß sie für den robusten Einsatz in Büro- und Fabrikumgebungen konzipiert wurden und vielfach nur über tiefe Protokoll-Stacks im Betriebssystem mit entsprechenden Latenzzeiten bei der Datenkommunikation nutzbar sind. Sogenannte *System Area Networks* (SAN) wurden aus diesem Grunde entwickelt, denn sie können für Übertragungen mit geringer Latenzzeit und hoher Bandbreite ausgelegt werden, und für sie existieren Programmierschnittstellen, um direkt – nach einer entsprechenden Initialisierung – aus einem Benutzerprogramm heraus zu kommunizieren, ohne den Umweg durch Betriebssystemschichten zu nehmen. Verbreitet sind hier Myrinet und SCI (*Scalable Coherent Interface*).

Parallelrechner mit SMP-Knoten verfügen über ein dediziertes, in der Regel sehr leistungsfähiges Verbindungsnetzwerk und werden im allgemeinen für hohe und höchste Anforderungen entwickelt und genutzt, wobei sowohl Anforderungen hinsichtlich Leistung (z. B. Rechenleistung, I/O-Leistung etc.) als auch Zuverlässigkeit, Software-Ausstattung, Platzbedarf, Wärmeentwicklung etc. befriedigt werden müssen. Entsprechend diesen Anforderungen ist das professionelle Einsatzgebiet von Parallelrechnern mit SMP-Knoten überwiegend die Lösung mittlerer bis größter Probleme in weiten Anwendungsbereichen wie Crash-Tests, Strömungs- und Verbrennungsprozessen, Wettervorhersage und Klimamodellierung, Elementarteilchenphysik und Festkörperforschung oder Data Mining.

PC- oder Workstation-Cluster /20/ haben in letzter Zeit ein reges Interesse in Forschung und Industrie gefunden, da sie mit relativ geringem finanziellem Aufwand für die Hardware eine beachtliche Peak-Prozessorleistung aufweisen können. Im universitären Umfeld werden meist Cluster selbständig aus Einzelkomponenten (Mainboards, Interface-Karten, Verkabelung, Switches) aufgebaut und verwaltet, jedoch sind auch Komplettlösungen von Herstellern lieferbar (z. B. Siemens hpcLine, Cray SuperCluster). Der wesentliche Gedanke bei diesen Cluster-Lösungen, die auch nach dem ursprünglich in den USA entworfenen System Beowulf-Cluster genannt werden, ist der mutmaßliche Preisvorteil bei der Verwendung von COTS (*Commodity Off The Shelf*), d. h. Komponenten, wie sie überall „von der Stange" gekauft werden können. Sie beschränken sich praktisch auf Intel- und Alpha-Prozessoren; dedizierte Cluster für parallele Anwendungen benutzen zum überwiegenden Teil diese Prozessoren als Basisbausteine.

Parallelrechner, die aus SMP-Knoten gebildet werden, haben gegenüber monolithischen Parallelrechnern den Vorteil, daß sie aus leistungsfähigen, aber dennoch relativ preiswerten Standardkomponenten modular aufgebaut werden können und entsprechend den Leistungsanforderungen in relativ weiten Grenzen konfigurierbar sind. Eine Voraussetzung für effektive Leistung ist jedoch die Ausgewogenheit des Gesamtsystems, insbesondere auch die Verwendung eines leistungsfähigen Verbindungsnetzwerkes. Man muß unterscheiden zwischen Systemen, die moderate Leistungs- und Verfügbarkeitsanforderungen bedienen können, und Hochleistungssystemen, die sowohl in Bezug auf Leistung als auch auf Zuverlässigkeit professionellen Ansprüchen an den Produktionsbetrieb dieser Rechner genügen müssen. Bei PC- oder Workstation-Cluster muß man gleichwohl auf ein sorgfältiges Design der Komponenten achten, wie z. B. mehrere, voneinander unabhängige PCI-Busse mit hoher Bandbreite, genügend Bandbreite zwischen Prozessoren und Hauptspeicher etc. Die PC-Technologie ist jedoch noch ungeeignet, sehr hohe Leistungsanforderungen zu erbringen. Hier sind niedrige Bandbreiten, Platzbedarf, Wärmeentwicklung und deren Abfuhr, geringe Leistungsdichte und mindere Zuverlässigkeit des Gesamtsystems sowie umständliche Handhabbarkeit und damit vergleichsweise hoher personeller Aufwand derzeit den Einsatz beschränkende Faktoren. Das Defizit im Bereich der Handhabbarkeit ist jedoch erkannt und in der Weiterentwicklung aufgegriffen worden, und es gibt insbesondere im expandierenden Einsatzbereich des Betriebssystems Linux Bestrebungen, hier Abhilfe zu schaffen.

Die Speicherbandbreite von Prozessoren zum Hauptspeicher ist ein wichtiger, wenn nicht der wichtigste Faktor für die Leistungsfähigkeit eines Systems. Hier ist insbesondere bei SMP-Systemen ein sorgfältiges Design erforderlich, um nennenswerte Leistungssprünge durch das Hinzufügen weiterer Prozessoren erreichen zu können. In der PC-Architektur für Intel-Prozessoren waren bislang enge Grenzen durch den gemeinsamen Bus gesetzt, über den die Prozessoren auf den gemeinsamen Hauptspeicher zugreifen mußten. Wichtige weitere Aspekte auch schon bei moderat parallelen Systemen sind zum einen die Höhe der Stromaufnahme und zum zweiten die Wärmeentwicklung, die primär durch die Logikbausteine erzeugt wird, und deren geregelte Abfuhr. Teilweise produzieren heutige Mikroprozessoren alleine schon mehr als 50 Watt Wärmeleistung pro Chip; die nächste Generation von Prozessoren wird weit mehr als 100 Watt erzeugen. Weitere Bausteine, Stromversorgung etc. können beträchtliche zusätzliche Wärme erzeugen. Schon der Aufbau von Parallelrechnersystemen moderater Größe, z. B. mit 32 Prozessoren, benötigt eine entsprechende Planung hinsichtlich der Stromzufuhr und Wärmeabfuhr. Ein weiterer Aspekt für Systeme, die hohen und höchsten Leistungsanforderungen genügen müssen, ist die Stellfläche bzw. das benötigte Volumen. Für den Power4-Prozessor von IBM wird es durch verschiedene Techniken (2 Prozessoren auf dem Chip, 4 Chips in einem Multi-Chip-Module, 4 Multi-Chip-Module auf einem Board) zu

einer hohen Rechenleistung von 140 GigaFlops pro Board kommen. Würde man die bisherige PC-Technologie (500 MHz Einzelprozessorsystem, 250 Watt Netzteil) zur Erzielung solcher Parallelrechnerleistungen nutzen wollen, benötigte man 280 solcher PCs und käme auf ein Volumen von ca. 11 cbm (0,04 cbm pro PC) und einer Leistungsaufnahme von 70 kW mit entsprechend hoher Wärmeentwicklung /21/.

Die Anforderungen der neuen Generation der SMP-Höchstleistungsrechner an die Klimatisierung der Maschinenhallen sind enorm. Da es sich bei den SMP-Systemen um luftgekühlte Anlagen handelt, muß die Abwärme – im Gegensatz zu den Vektorrechnern vom Typ Cray und den MPP-Rechnern wie Cray T3E – ausschließlich über die Luft abgeleitet werden. Ein weiterer kritischer Punkt der heutigen Mikroprozessor- und Speichertechnologie als Grundlage für Höchstleistungsrechner jenseits der TeraFlops-Grenze ist in der effektiven Stellfläche zu sehen, die leicht einige tausend Quadratmeter umfassen kann. In der Regel sind für solche Systeme neue Maschinenhallen oder komplette neue Rechenzentren zu erstellen, wie dies in den amerikanischen Zentren geschehen ist; auch für die deutschen Supercomputer-Zentren in Jülich, München und Stuttgart sind solche Maßnahmen in der Planung.

4. Von TeraFlops via PetaFlops zum Ende von Moore's Law

Heute beträgt die Steigerung der Leistungsfähigkeit der Spitzensysteme, wie sie die TOP500-Liste ausweist, fast den Faktor 2; dabei wächst die Zahl der Prozessoren durchschnittlich um den Faktor 1.30, die Prozessorleistung um den Faktor 1.40 (gegenüber 1.58 des Moore'schen Gesetzes). In der TOP500-Liste vom Juni 2001 /3/ war die Hitachi-Maschine des Leibniz-Rechenzentrums in München bereits auf Platz 12 zurückgefallen und hat somit den Platz unter den ersten 10 innerhalb eines Jahres wieder verloren (Abb. 4). Die ersten Plätze werden von den ASCI-Systemen eingenommen; unter den Herstellern dominiert IBM mit den SMP-Cluster-Systemen vom Typ SP-Power3, der neue Prozessor Power4 wird die Dominanz verstärken.

Dieser Entwicklung kommt um so mehr Bedeutung zu, als die Firma Compaq vor kurzem angekündigt hat, daß sie den Alpha-Chip über das Modell SV7 hinaus nicht weiterentwickeln wird, was einen drastischen Endpunkt einer großen Prozessorlinie und damit de facto auch das Ende der von Compaq für die 30-TeraFlops-ASCI-Maschine in Los Alamos entworfene SMP-Architektur bedeutet. Hinzu kommt, daß der „Merger" von Silicon Graphics mit Cray Research und die vor einem Jahr vollzogene Trennung der beiden Firmen für die Weiterentwicklung der erfolgreichen T3E-Linie in eine SMP-Architektur keinen Raum ließ und beide Firmen im Grunde drei Jahre Zeit für innovative Entwicklungen verloren haben. Durch die Abschottung des amerikanischen Marktes gegen japanische Superrechner wurde deren Weiterentwicklung trotz des relativen Erfolges auf dem außeramerikanischen Markt zweifelsfrei nachhaltig gehemmt /22/. Die jüngst erfolgte Öffnung der USA für NEC-Systeme im Zuge der Stabilisierung von Cray Inc. muß erst noch ihre Wirkung zeigen. SGI verfolgt eine Doppelstrategie mit ihrer genuinen, aber leistungsmäßig zurückfallenden MIPS-Linie und der neuen Intel-Prozessor-basierten SMP-Linie. Sun Microsystems will ambitioniert ins Höchstleistungsfeld vorstoßen; die neue Leistungsklasse ihrer Chips muß sich aber erst noch bewähren. Auch jetzt sind also „darwinistische" Selektionsprozesse weiter im Gange /23/.

Es läßt sich aber voraussehen, daß im Jahre 2005 nicht nur das erste 100-TeraFlops-System installiert sein wird, sondern daß auch keine Maschinen unter 1 TeraFlops mehr einen Platz in der TOP500-Liste haben werden. Bis zum Jahr 2010 dürfte der erste Rechner mit einer Spitzenleistung von einem PetaFlops verfügbar sein. Die einschlägigen Forschungs- und

Entwicklungsaktivitäten in den USA nehmen an Breite und Tiefe zu /8/, sie scheinen aber derzeit noch auf die USA beschränkt zu sein. Um diese Ziele zu erreichen, sind aber erhebliche technologische Anstrengungen nötig, denn mit der gegenwärtigen und absehbaren SMP-Prozessortechnologie und ihren Anforderungen an Stromversorgung, Klimatisierung und Stellfläche würden die noch sinnvollen Grenzen gesprengt. Es zeichnen sich aber – derzeit sicher noch spekulative – neue Ansätze für zukünftige Prozessoren ab /24, 25/, und die sogenannte PiM-Technologie (PiM: *processor in memory*), wie sie wohl vorrangig von IBM verfolgt wird, ist gerade für Höchstleistungssysteme vielversprechend. Für den Einsatz in der Gentechnik und Bioinformatik entwickelt IBM parallel zur offiziellen Produktlinie den neuartigen Rechner „Blue Gene", der mit 64.000 PIM-Chips mit je 32 Prozessoren und einer Leistungsfähigkeit von 32 GigaFlops mit insgesamt einer Million Prozessoren die Größenordnung von 10^7 Operationen pro Maschinenzyklus und damit effektiv ein PetaFlops leisten soll (Abb. 5 und 6). Die publizierte Produktlinie von IBM weist für die Jahre 2003-2004 Systeme mit 30 TeraFlops, für 2005-2006 mit 100 TeraFlops und für 2008-2009 mit 1 PetaFlops aus; der Parallelitätsgrad, d.h. die einzusetzenden Prozessorzahlen, betragen 10.000, 20.000 und 40.000 (Abb. 7).

Im Jahre 2002 soll der von NEC mit 640 CMOS-Vektorprozessorknoten entworfene spektakuläre „Earth Simulator" (Abb. 5, 8 und 9) mit einer Leistungsfähigkeit von 40 TeraFlops in Betrieb gehen /26/. Cray Inc. arbeitet mit Unterstützung der US-Regierung an dem neuen Supercomputer SV2, der aus Architekturbausteinen von 50- bzw. 100-GigaFlops-Knoten aus vier CPUs (12.8 GigaFlops) bestehen soll (Abb. 5 und 10). Bis zu 4096 CPUs lassen sich über ein schnelles Verbindungsnetzwerk (100 GigaByte/sec Bandbreite zu den Knoten) zu 50 TeraFlops hochskalieren. Der für das Jahr 2002 vorgesehene Termin für die Markteinführung dürfte aber zu kurzfristig angesetzt sein.

Diese Entwicklungen sind absehbar und erscheinen daher relativ gesichert, wenngleich Planungssicherheit über die kurze Frist hinaus im Supercomputing stets ein Abenteuer war. Die Frage bleibt aber, was für die Technologie und damit schließlich für die Höchstleistungsrechnerarchitektur jenseits des Zeitpunktes erwartet werden kann, an dem die Gültigkeit des optimistischen Moore'schen Gesetzes (Abb. 11) aufgrund der quantenmechanischen Effekte zusammenbrechen wird (Abb. 12); die Prognosen zielen auf die Zeit zwischen 2015 und 2020. Die vagen Hoffnungen für die „Zeit danach" ruhen auf der das Moore'sche Gesetz schließlich zerstörenden Quantenmechanik selbst (Abb. 13): Das Potential des Quantencomputers /27, 28/ stellt Theorie und Technologie vor neue Herausforderungen, die verstärkt und ohne weiteren Verzug auch von der Forschung und Entwicklung in Deutschland angenommen werden sollten.

Literatur

/1/ I. Foster and C. Kesselman (eds.), *The Grid: Blueprint for a New Computing Infrastructure*, Morgan Kaufmann Publishers, San Francisco, 1999

/2/ K. Hwang, *Advanced Computer Architecture: Parallelism, Scalability, Pro-grammability*, McGraw-Hill, 1993

/3/ H.-W. Meuer, J. J. Dongarra, E. Strohmaier, and H. D. Simon (eds.), *TOP500 Supercomputer Sites*, 17th Edition, June 21, 2001, RUM 62/2001,University of Mannheim, and LBNL-48241, University of Tennessee; Proceedings 16-th International Supercomputer Conference 2001, Heidelberg, 20–23 June 2001; www.supercomp.de

/4/ ASCI: www.llnl.gov/asci

/5/ ASCI-Pathforward: www.llnl.gov/asci-pathforward

/6/ L. Smarr, *Toward the 21st Century*, Comm. ACM 40(1997), No. 11, 28-32. – Ph. L. Smith, *The NSF Partnerships and the Tradition of U.S. Science and Engineering*, Comm. ACM 40(1997), No. 11, 35-37. – D. A. Reed et al., *Distributed Data and Immersive Collab*oration, Comm. ACM 40(1997), No. 11, 39-48. – R. Stevens et al., *From the I-Way to the National Technology Grid*, Comm. ACM 40(1997), No. 11, 51-60. – K. Kennedy et al., *A Nationwide Parallel Computing Environment*, Comm. ACM 40(1997), No. 11, 63-72. – G. J. McRae, *How Application Domains Define Requirements for the Grid*, Comm. ACM 40(1997), No. 11, 75-83. – J. P. Ostriker and M. L. Norman, *Cosmology of the Early Universe Viewed Through the New Infrastructure*, Comm. ACM 40(1997), No. 11, 85-94

/7/ Report from the President's Information Technology Advisory Committee (PITAC Report to the President, February 24, 1999), Information Technology Research: Investing in our Future; www.ccic.gov/ac/report

/8/ J. J. Dongarra and D. W. Walker, The Quest for Petascale Computing, Computing in Science & Engineering Vol. 3(2001), No. 3, 32-39

/9/ F. Hoßfeld, *Verbund der Supercomputer-Zentren in Deutschland – Ansichten, Einsichten, Aussichten* –, in: H.-W. Meuer (Hrsg.), *Supercomputer 1998, Anwendungen, Architekturen, Trends*, K. G. Baur, München, p. 160-171

/10/ F. Hoßfeld et al., Gekoppelte SMP-Systeme im wissenschaftlich-technischen Hochleistungsrechnen – Status und Entwicklungsbedarf – (GoSMP), Analyse im Auftrag des BMBF, Förderkennzeichen 01 IR 903, Dezember 1999

/11/ U. Kulisch, Wissenschaftliches Rechnen mit Ergebnisverifikation, Mathematical Research Volume 58, Akademie-Verlag, Berlin, 1989

/12/ U. W. Kulisch, Advanced Arithmetic for the Digital Computer – Design of Arithmetic Units, Report, Universität Karlsruhe, Institut für Angewandte Mathematik, April 2001

/13/ P. J. Roache, Verification and Validation in Computational Science and Engineering, Hermosa Publishers, Albuquerque, 1998

/14/ UNICORE: http://www.fz-juelich.de/zam/RD/coop/unicoreplus/index.html

/15/ EUROGRID: http://www.fz-juelich.de/zam/RD/coop/eurogrid.html

/16/ Gigabit-Testbed West: http://www.fz-juelich.de/zam/RD/coop/gigabit/gigabit.html

/17/ Gigabit-Testbed Süd: http://gtb.rrze.uni-erlangen.de/

/18/ HPCM: http://www.fz-juelich.de/zam/RD/coop/hpcm.html

/19/ H. Richter, Verbindungsnetzwerke für parallele und verteilte Systeme, Spektrum, Akademischer Verlag, Heidelberg, 1997

/20/ NOW: http://now.cs.berkeley.edu

/21/ T. Mudge, Power: A First Class Architectural Design Constraint, IEEE Computer 34(2001), No. 4, 52-58

/22/ Y. Oyanagi, Developmewnt of Supercomputers in Japan: Hardware and Software, Parallel Computing 25(1999), 1545-1567

/23/ E. Strohmaier, J. J. Dongarra, H. W. Meuer, and H. D. Simon, The Marketplace of High-Performance Computing, Parallel Computing 25(1999), 1517-1544

/24/ S. Vajapeyam and M. Valero, Early 21st Century Processors, IEEE Computer Vol. 3(2001), No. 3, 47-50

/25/ G. S. Sohi and A. Roth, Speculative Multithreaded Processors, IEEE Computer Vol. 3(2001), No. 3, 66-73

/26/ K. Tani, Status of the Earth Simulator System, Proceedings 16-th International Supercomputer Conference 2001, Heidelberg, 20-23 June 2001; www.supercomputer.de

/27/ P. Zoller, The Post Moore's Law Era: Quantum Computing, Proceedings 16-th International Supercomputer Conference 2001, Heidelberg, 20-23 June 2001; www.supercomputer.de

/28/ M. A. Nielsen and I. L. Chuang, Quantum Computation and Quantum Information, Cambridge University Press, Cambridge, UK, 2000

Abbildungen

Abb. 1

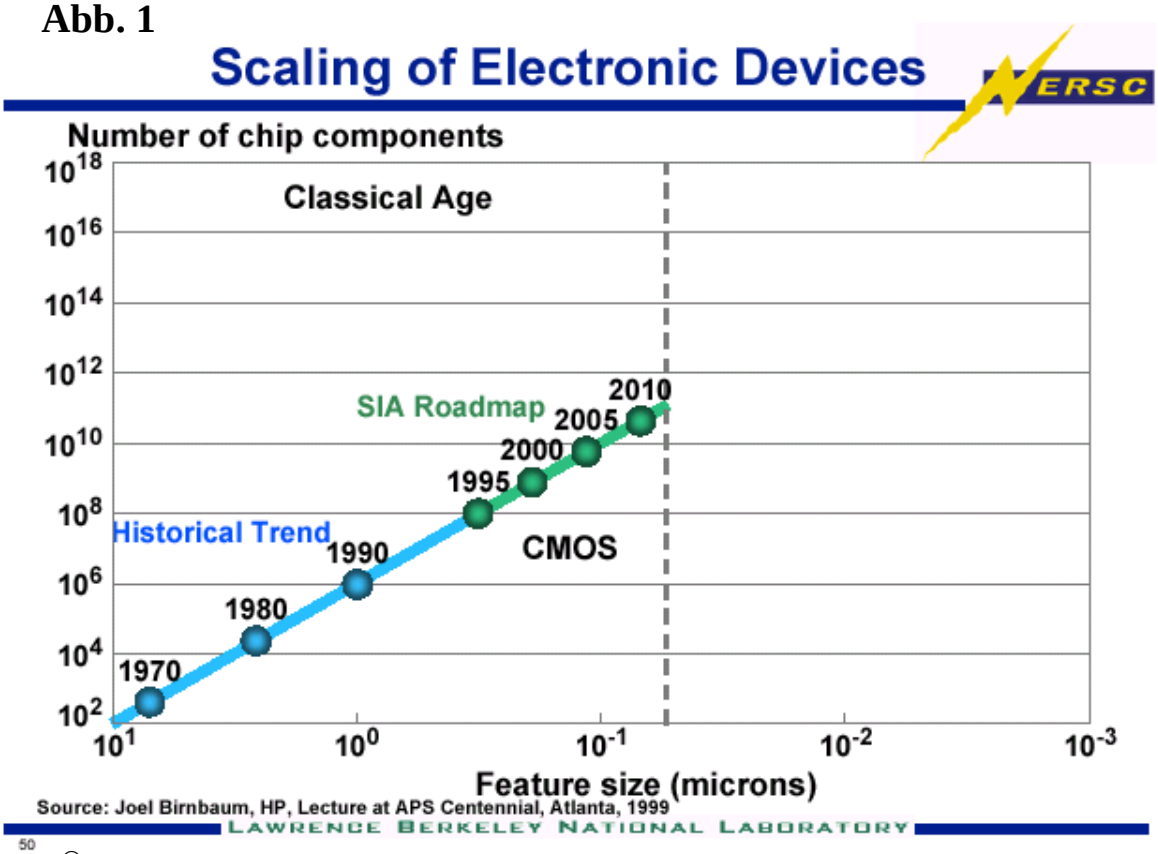
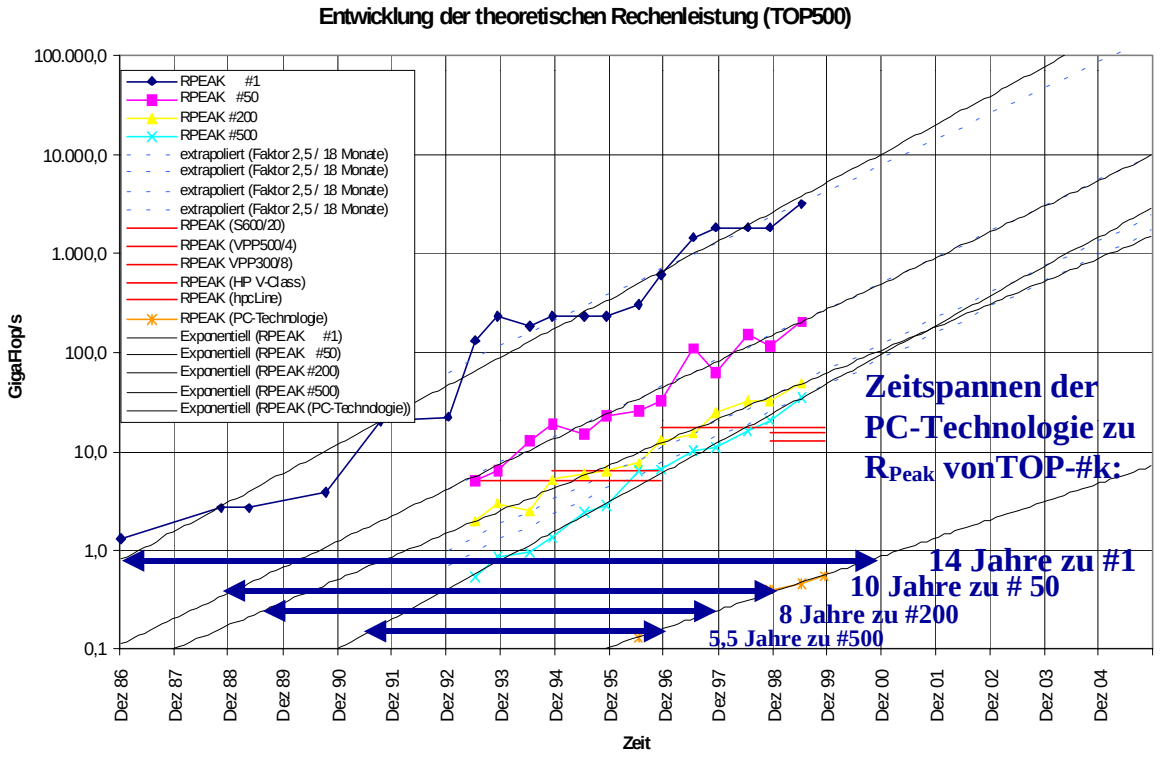


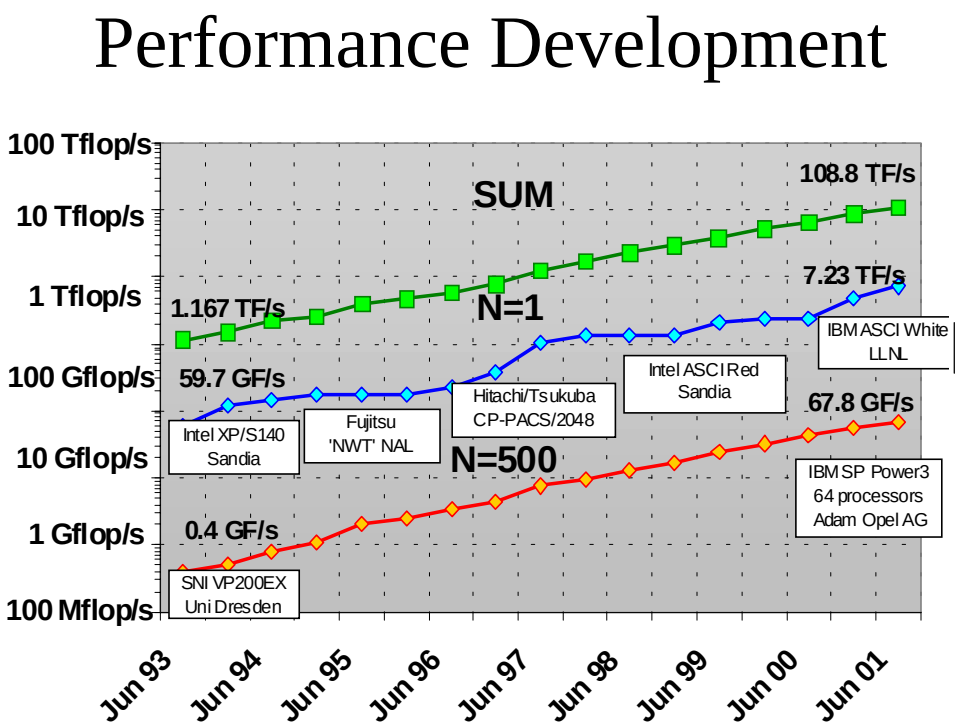
Abb. 2



Zeitliche R_{Peak}-Leistungsabstände zwischen HPC und PC

© Ch. Bischof, RWTH (Ergänzung: F. Hoßfeld, NIC-ZAM)

Abb. 3



© E. Strohmaier 2001

Abb. 4

TOP10

out of TOP500 List - 21 June 2001

Rank	Manufacturer	Computer	R_{max} [TF/s]	Installation Site	Country	Year	Area of Installation	# Proc
1	IBM	ASCI White SP Power3	7.23	Lawrence Livermore National Laboratory	USA	2000	Research	8192
2	IBM	SP Power3 375 MHz	2.53	NERSC/LBNL	USA	2001	Research	2528
3	Intel	ASCI Red	2.38	Sandia National Laboratory	USA	1999	Research	9632
4	IBM	ASCI Blue Pacific SST, IBM SP 604E	2.14	Lawrence Livermore National Laboratory	USA	1999	Research	5808
5	Hitachi	SR8000/MPP	1.71	University of Tokyo	Japan	2001	Academic	1152
6	SGI	ASCI Blue Mountain	1.61	Los Alamos National Laboratory	USA	1998	Research	6144
7	IBM	SP Power3 375Mhz	1.42	Naval Oceanographic Office (NAVOCEANO)	USA	2000	Research	1336
8	NEC	SX-5/128M8 (3.2ns)	1.19	Osaka University	Japan	2001	Academic	128
9	IBM	SP Power3 375Mhz	1.18	National Centers for Environmental Predicition	USA	2000	Research	1104
10	IBM	SP Power3 375Mhz	1.18	National Centers for Environmental Predicition	USA	2001	Research	1104

Abb. 5

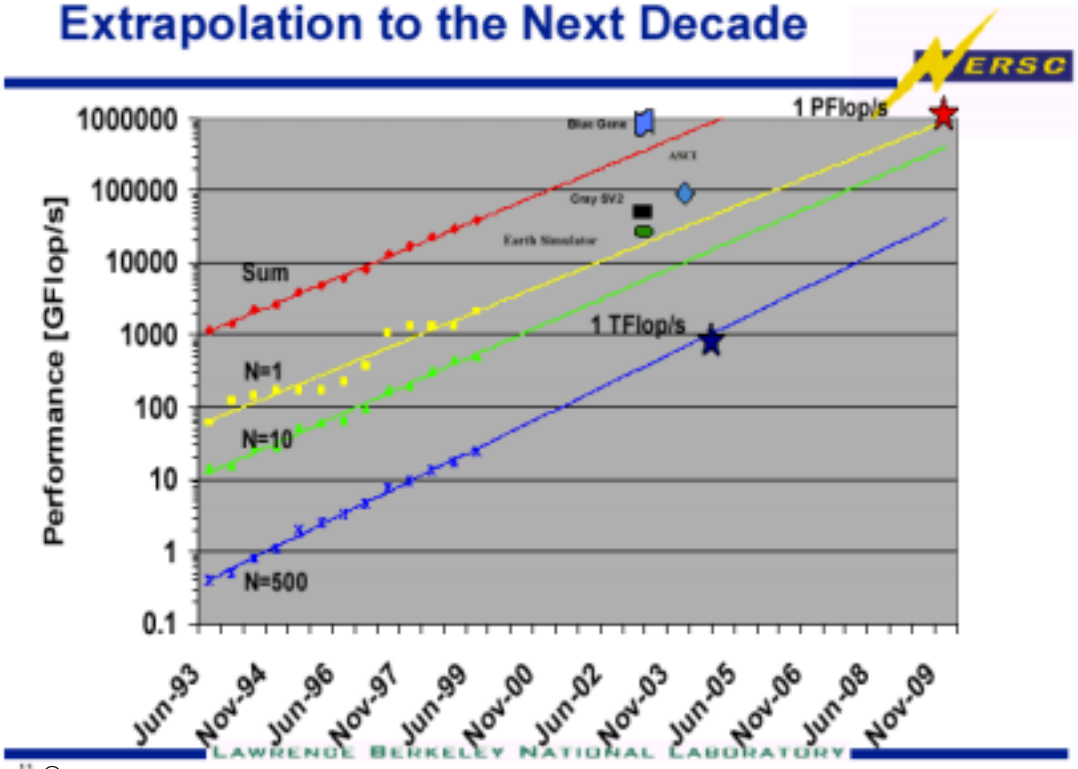


Abb. 6

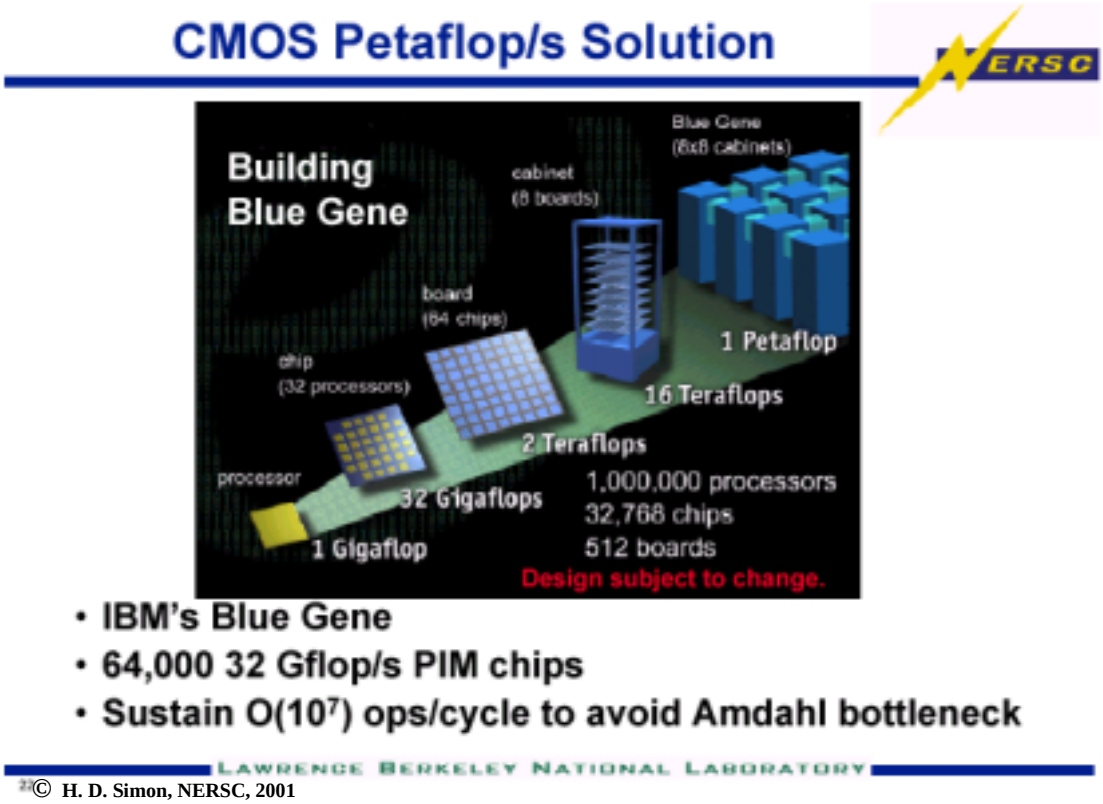


Abb. 7

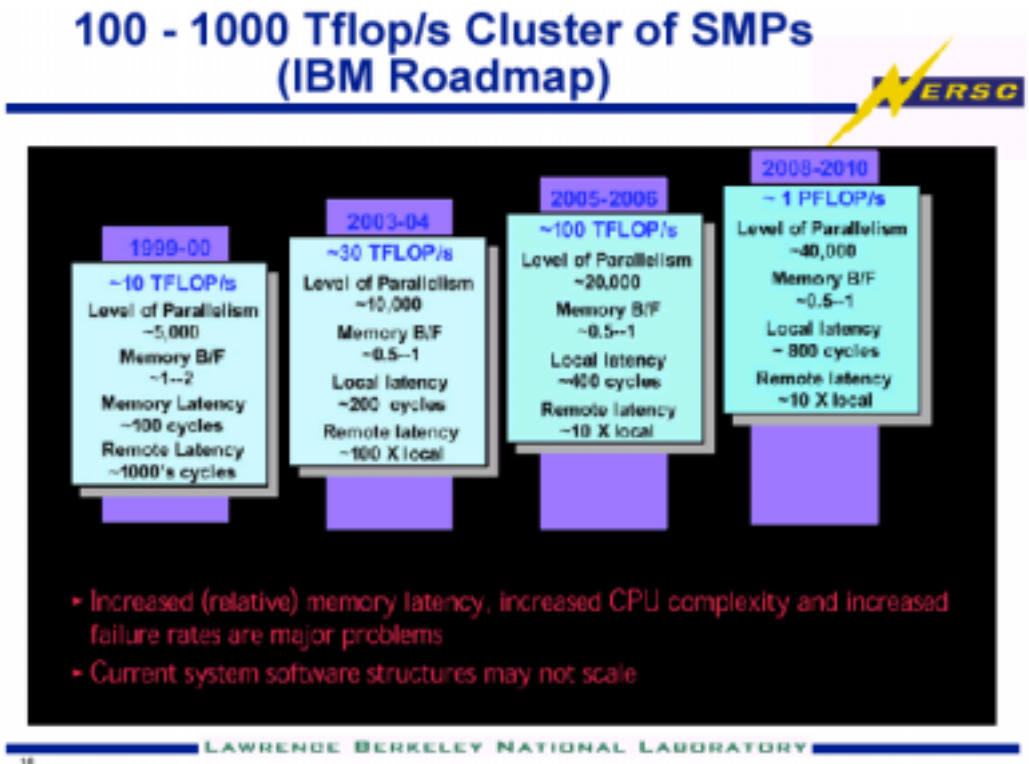


Abb. 8

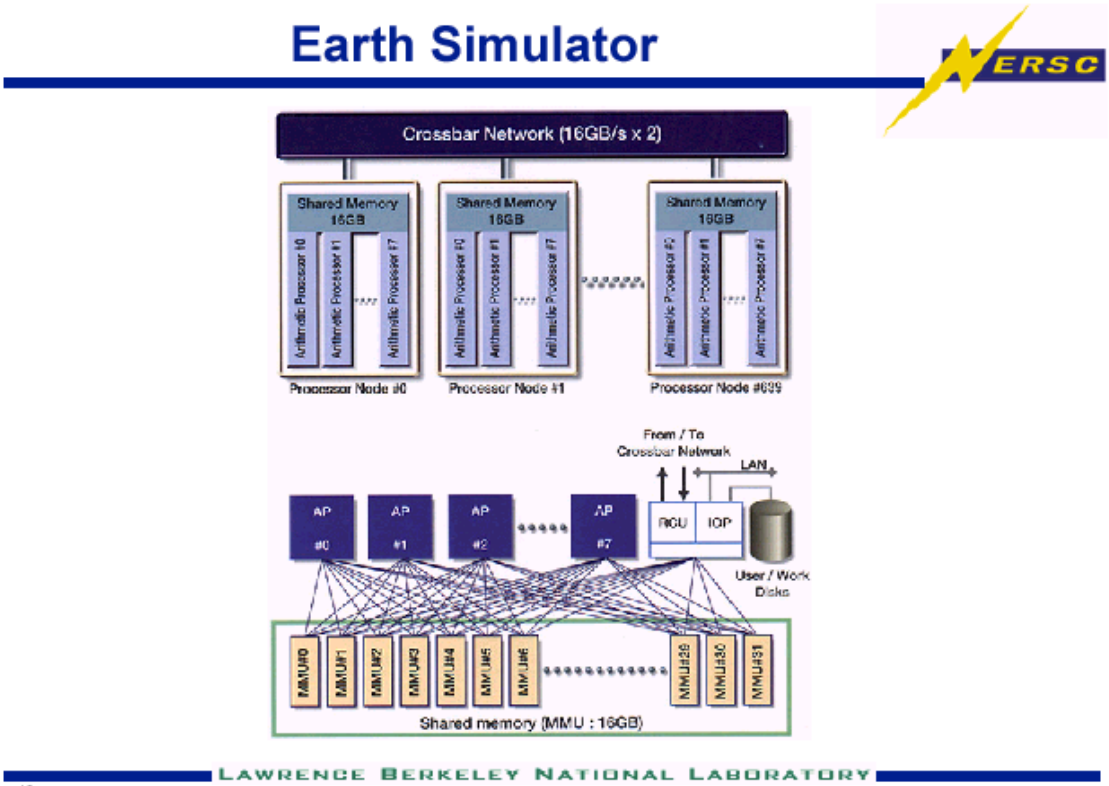


Abb. 9

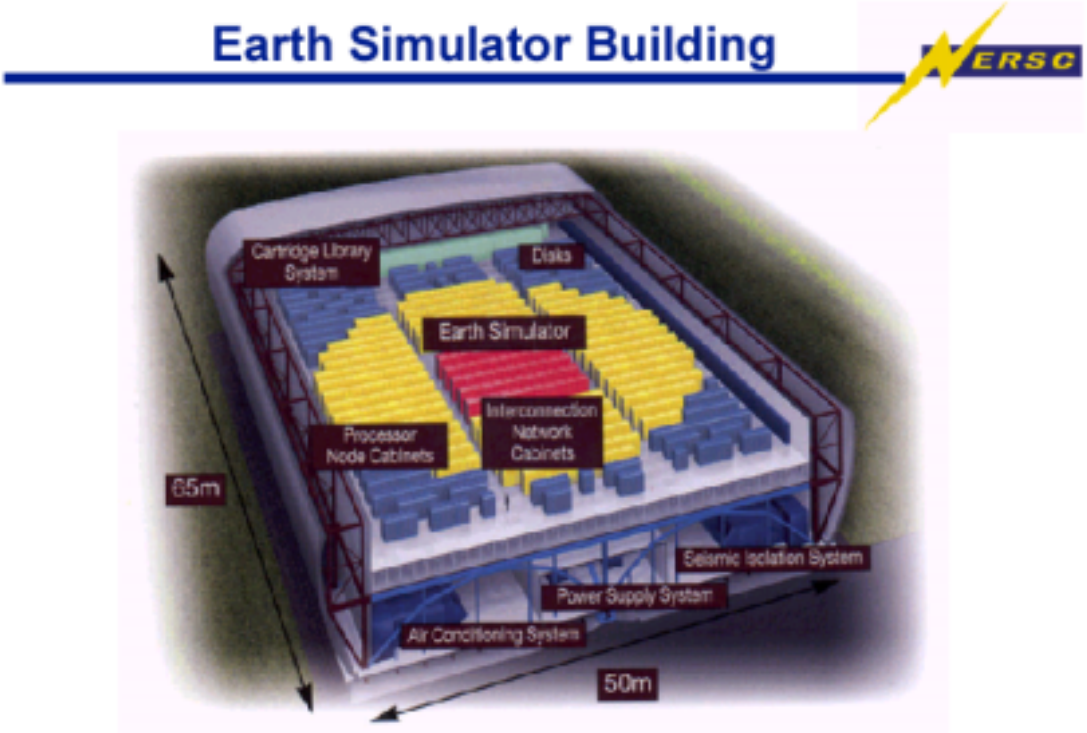


Abb. 10

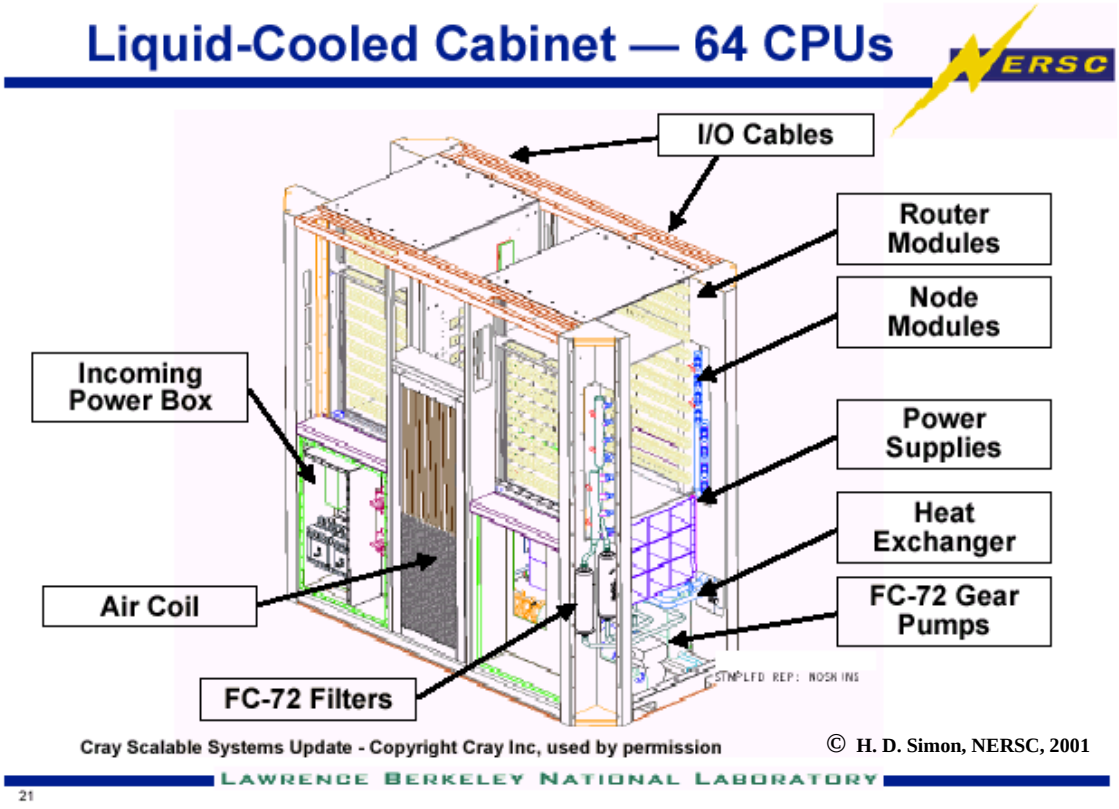
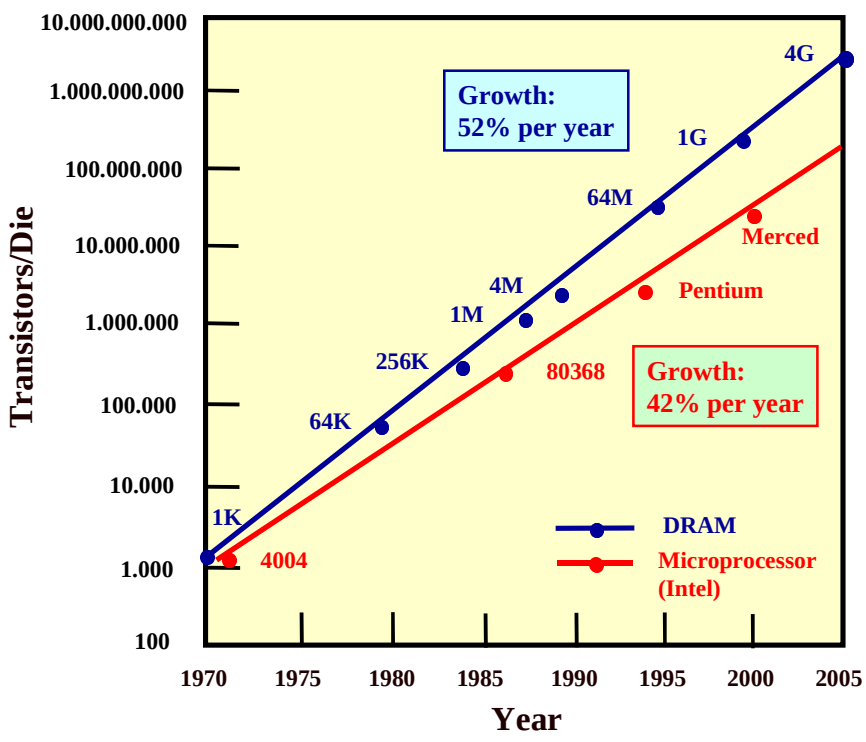


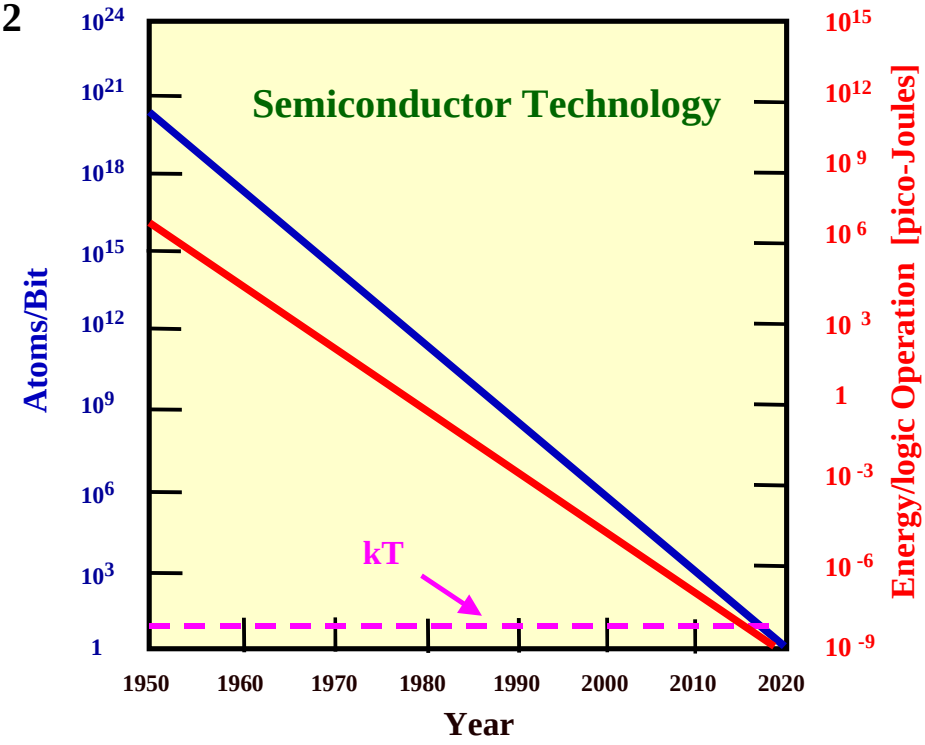
Abb. 11



Moore's Law in Semiconductor Technology

© F. Hossfeld 2000

Abb. 12



Information Density & Energy Dissipation

(adapted from C. P. Williams et al., 1998)

Abb. 13