

## Archived Material

### Historical Purposes Only

## **Testimony Hearings on the President's Information Technology Advisory Committee, Interim Report to the President**

### **Details**

October 6, 1998

Testimony before the U.S. House of Representatives Committee on Science, Subcommittee on Basic Research

Joe Thompson

Distinguished Professor of Aerospace Engineering Founding director, NSF ERC for Computational Field Simulation, Mississippi State University

### **Introduction**

Chairman Pickering and distinguished members of the Subcommittee on Basic Research, my name is Joe Thompson. I am a Distinguished Professor of Aerospace Engineering at Mississippi State University. I was the founding director of the NSF Engineering Research center (ERC) for Computational Field Simulation at Mississippi State, and I currently am leading a national ten-university team providing Programming Environment & Training (PET) support for the DoD Major Shared Resource center (MSRC) at CEWES in Vicksburg MS. I am also Special Assistant in High Performance Computing to the Vice President for Research at MSU, and am leading our effort as a charter member of the Internet2 consortium.

Mississippi has now moved into third position among all the states in terms of total unclassified high performance computing facilities within the state, with some 41% of DoD's unclassified HPC power: two of the four DoD Major Shared Resource centers (MSRCs) are located in Mississippi. Mississippi State University ranks 20th among universities in HPC, and fourth in the Southeast. The NSF ERC at Mississippi State is still the only one of the NSF ERCs having its focus directly on high performance computing.

Although I am a member of the President's Information Technology Advisory Committee (PITAC), I am speaking today of the PITAC report and high performance computing in general from the perspective of a computational scientist - one of only three computational scientists on the PITAC. Ken Kennedy has given the committee a good discussion of the PITAC report.

I do strongly believe in the PITAC report, and in fact I think that we must make it even more forceful as we move from the interim to the final report later this year.

I want to make two main points today, both of which are directly related to the PITAC report:

- That we have neglected to fund software research commensurate with and concomitant to funding of hardware acquisitions.
- That we are reaping the fruits of the last decade's research in information technology while neglecting research in this decade.

On the first point, today we have high-end hardware that is only 20% effective for computational science because the necessary software to fully utilize the hardware has not been developed. This is being taken into account, in fact, in strategic planning for hardware acquisitions: purchasing hardware rated, in terms of peak performance, several times as powerful as needed in order to realize the capability that is actually required in real computational science applications. But, in fact, we know more about how to design hardware today than we do about how to design software. PITAC described software as the new physical infrastructure of the information age, and as among the most complex of human-engineered structures. That poses a daunting challenge.

The PITAC report notes that we have consistently neglected to adequately address the software advances that are required to fully utilize advances in hardware. More powerful machines do not proportionally increase capability in computational simulation applications until software suitable to the hardware architecture is developed. In this regard, the report called specifically for software research to be made a substantive component of every major information technology initiative. But hardware has a strong vendor lobby, while support for software is diffuse. There have even been repeated outside efforts to divert funding set by agencies for essential software development into unplanned hardware procurement.

I want to be very careful here not to appear to be arguing against funding for hardware. In fact, I agree strongly with the call in the PITAC report for major funding and effort in high-end systems. But I am saying, in concert with the report, that carefully considered technical decisions are necessary that consider hardware and software together in funding programs. High-end systems means high-end hardware and the necessary systems and applications software. And that consideration must necessarily have an effect on the direction of high-end architecture. We must come to think of a computing system - hardware and software appropriate to each other - rather than a machine: that mechanical analogy may be unfortunate.

It is false economy to consider hardware acquisition apart from software requirements. The need to amortize the high cost of new chip development, and especially of new fabrication facilities, over high volume sales units favors high-end systems based on large numbers of commodity processors operating in parallel. While there is definitely logic in such a scalable approach, it is only the commodity processors that are off the shelf - the required software is left out of the cost equation. This can result in under-utilization of expensive hardware for lack of appropriate software. It is essential that the cost of both hardware and software development be considered in high-end acquisitions, and this consideration may change the relative positions of hardware alternatives.

This is not to imply that the primary metric for the success of computer systems should be the approach to full utilization. Not at all, for utilization is a function both of problem complexity and mix, and of system design

and operation. And criticism of systems on the basis of utilization alone is not warranted. Further, there is the question of a balance between the necessity to continue to run mission-critical software in the face of architectural advances and the time required for migration into new approaches. But significant under-utilization does raise questions that must be addressed.

The important point here is one of balance: software development coupled with hardware development, maintenance of legacy code coupled with new code development. It is simply a mistake to consider only one element of this balance. Raw hardware power will not suffice to address the challenges of computation science. But neither can we allow new directions in hardware architecture to be bound by existing software. This is a fundamental reason why research in software design is vital.

The Nation cannot afford to neglect its high-end computer needs. While we should leverage commodity systems to the fullest extent, we cannot allow other nations to exceed our capability for computational science applied in scientific discovery, and in engineering analysis and design, for any reason. Currently the market for high-end machines is some three times the market for beef jerky. Dependency on commodity processors for high-end systems forces a concomitant major research investment in software design and development that has not been fully acknowledged.

Major effort is required in both systems software and applications software. Systems software makes the hardware work: such things as compilers, schedulers, file managers, debuggers, security, communications, etc. Applications software performs computational simulations of physical problems. But, behind each of these is sorely needed fundamental advances in software design and engineering, so that both systems software and applications software can be developed to be reusable, robust, and secure. Software design must be brought to a higher level, similar to that which now enables complex chip design. We simply cannot continue to tolerate the labor-intensive and error-prone software development that is now the case. The PITAC report notes that neither are we adequately improving the efficiency of software construction nor are we training enough professionals to supply the needed software.

One way to get a grasp of the software problem is to consider the contrast between replacing an entire fleet of trucks with replacing a software system. Replacement of an entire fleet of GM trucks with Ford trucks would pose no great problem either for drivers (users) or mechanics (maintenance). The same could even be said about the Air Force replacing all its Boeing aircraft with aircraft from Lockheed. But today replacing a major software system is a traumatic experience for businesses, universities, or government. While chip design (hardware) is fundamentally done with similar design principles and tools in different companies, such effective design tools for software do not yet exist. Yet we are basing the essential infrastructure of commerce and government on this esoteric medium. Contrast the effect of being denied access to all Ford mechanics and engineers with loss of access to the supporters of a major software system.

And, as the report notes, in addition to the impact of insufficient funding for software research, there is the additional effect of the very strong attraction of young researchers into short-term commercial development of Internet systems. Stock valuation, rather than revenue, has become the currency of success in our presently technology driven economy, and this has taken focus on the short term to the extreme. There is little impetus for young computer scientists to apply their efforts to the type of long-term research that, in the past decade, laid the foundation for the present rapid succession short-term commercial successes.

While we should not seek to control commercial directions, Federal initiatives must address the long term if our leadership position in information technology is to be sustained. And the PITAC report expressed serious concern over the fragility of the software infrastructure. It is easy to overlook the fact that today businesses are betting the company on an essential free infrastructure - the Internet - that they did not design and over which they have little control. That is not the usual business approach.

The Internet was, of course, not designed for the purpose it now serves for business, and we are in considerable danger of the consequences of that fact as commerce becomes more dependent on the Internet. The wild successes of the present decade make me think of what William Faulkner, from my part of the country, said: A mule will work for a man faithfully for ten years just for the pleasure of kicking him once. Our ten years may soon be up, and it only takes one kick from a mule to do major damage.

That is why the research in network technology supported by the Next Generation Internet initiative is so important to the country. This is essential research to address fundamental problems with the present Internet, such as security, quality of service, scaling, and management. Such effort, analogous to the Federally-supported effort of the past two decades that made the Internet available to businesses in the present decade, is essential to future commercial success, but such long-term effort is not going to attract commercial investment in today's orientation toward short-term launch of start-ups and rapid release of a succession of Internet applications.

The need for research in software design also impacts indirectly on the current workforce demands in information technology in two ways: First, without effective software design tools, software development is esoteric and very labor-intensive, increasing the requirements for the workforce. Second, the lack of such design tools makes software design, in the team-oriented approach that is necessary for large systems, a tedious operation that is not attractive to many students who might well be interested in the individual logical creativity that is the heart of software innovation. Without effective the software design tools that will result only from coordinated long-term research, we may be in danger of having information technology support and maintenance become the dog-work of the next decade.

Funding of long-term research is essential to address both of these points: to advance sorely-needed software development and to provide the impetus that is now lacking for young researchers to pursue graduate study and to make these advances to sustain our leadership position in information technology. Here we need to increase the funding of fellowships and graduate assistantships in areas directly related to information technology in order to attract more students into graduate study. Today students, and especially women, are much more likely to go into the life sciences than into the information technology areas: computer science, computer engineering, computational science and engineering. But recent data shows that the PhD output in the life sciences has now passed the demand, while demand in information technology continues to greatly exceed the output of the graduate programs.

The PITAC report also addresses the need for more centralized leadership of the cross-cutting Federal program in information technology, and - not wanting to call for creation of a new Federal agency - recommended NSF as the most appropriate agency to provide this leadership. It might be noted, however, that the present pervasive importance of information technology to both our economy and security makes this period not

unlike that when NASA was created in response to the launch of the space age.

One illustration that more coordination of this cross-cutting initiative is necessary is apparent lack of real coordination between the very large DoD HPCMP and DoE ASCI programs, both of which are concerned with the utilization of computational science on high-end machines for computational simulation of physical phenomena and processes. It should be realized that similar enabling computational science technology is required for computational simulation of warfighting systems in DoD HPCMP and of the nuclear stockpile in DoE ASCI. Coordination between such programs is thus essential if duplication is to be avoided, gaps are not to be left, and the best available technology is to be effectively brought to bear. Cross-cutting elements in agency budget requests are still, of course, reviewed independently by Congressional committees - an inherent problem in cross-cutting initiatives.

It is important to stress that the PITAC committee also notes that it will be necessary for NSF to make some adjustments in regard to fulfilling a National leadership role in information technology. Otherwise, I would think the creation of a new agency for information technology would, in fact, have to be considered. Never in history, perhaps, have we been so deeply and broadly dependent on a body of technology. I believe that the importance of information technology to the Nation, and the complexities involved in mounting a coordinated initiative will require oversight by an advisory body such as the PITAC in any case.

The fundamental research in software needed will require strategic coordination in order to ensure that the range of needed developments is addressed. Also, much of the research needed is inherently multidisciplinary, and that is unfortunately somewhat counter to the cultures of both NSF and universities. Cross-disciplinary research requires an orientation more toward centers than the individual head-down investigators that is more the norm at NSF and in the view, perhaps, of the Science Board. While this emphasis on individual investigators is properly important to NSF, a different approach will be necessary in fulfilling this leadership role in long-term research in systems and applications software. Multidisciplinary collaborative research in the center mode should be legitimized and incentivized from the Science Board level to the program level.

The PITAC report calls strongly for such multidisciplinary centers. And it is important to note that centers focusing on enabling technology are called for, as well as centers of visionary opportunity. The Federal HPCC program, now concluded, had opportunities as its focus, properly forging ahead to catch the country's imagination as to the possibilities opened by high performance computing and communications. Now it is time to address the obstacles, to develop the cross-cutting enabling technology that the opportunities all require.

Two recent examples of collaborative multidisciplinary effort between universities and Federal labs in high-end computing software effort are the Programming Environment & Training (PET) component of the DoD High Performance Computing Modernization Program (HPCMP) - in which I lead a team of ten universities - and the university alliance component of the DoE ASCI program. Speaking from direct experience in the DoD PET effort, this collaboration between university researchers and DoD high-end computer users is a definite culture challenge to both, but with good results.

An approach that might be taken by NSF in this multidisciplinary initiative in software research is to require individual investigators to associate the proposed effort with a relevant research center. The Associates

Program of the National Research Council provides a successful model for this, in that proposers must first engage the interest of an appropriate Federal laboratory and must propose to conduct the work collaboratively in residence at the lab. While neither the restriction to Federal labs nor to residency is appropriate in the present case, the analogy holds, so that an individual investigator at a university could be expected to propose to work in networked collaboration with a relevant university research center or Federal lab.

Computational simulation of physical phenomena and processes has great potential to improve engineering analysis and design in industry, as well as scientific investigation in general. The computer has become a new and very powerful device for scientific discovery. Impact is already being felt in the aerospace and automotive industries, and is now moving into the shipbuilding industry. At the ERC we have direct experience with the importance of computational science to these industries - in particular with Lockheed-Martin, Boeing, Ford, General Motors, and the aircraft engine companies through NASA Lewis Research center. And we have also worked with smaller companies making industrial fans and aircraft components to increase their capability in computational simulation.

In the latest listing of the Top 500 supercomputer sites in the world, 162 are in industry: about one third. And two-thirds of those industrial sites are in the US. The top site in the chip, pharmaceutical, geophysical, and electronic industry is in the US. But the top site in the aerospace, telecommunications, automotive, and finance industry is outside the US. An indication of just how widespread HPC is in industry is that Waste Management Systems, Rubbermaid, and Walgreens are among the Top 500 sites.

This greatly enhanced design capability, which can both shorten the design cycle and broaden the range of considerations, will ultimately impact throughout industry from toys to biological systems. Another example is the work of the DoD CEWES MSRC in Vicksburg MS on using computational science on high-end computers to simulate recent terrorist bombing events, in order to determine effective preventive measures.

In contrast to many business and financial problems, where a massive number of essentially independent alternatives or events must be evaluated or scheduled, computational science at the high end poses an applications software problem, in addition to the systems software problem, when reliance is placed on large parallel systems of commodity processors: Field solutions, such as fluid mechanics, heat and mass transfer, electromagnetic and plasma dynamics, are massively parallel in that they inherently involve simultaneous solutions at a multitude of locations throughout the field. But, because these solutions are simultaneous, much communication is required between neighboring locations. An analogy would be for everyone in this room to be intending to go to dinner tonight at the same restaurant - but having to decide on the restaurant by speaking only with those in adjacent seats. Computational science thus requires not only fast processors, but also fast communication networks connecting the processors, fast access to data, locality of data to processors, and enabling software. We don't get the rest of this with the commodity processors.

And even here, where applications software is being developed to address the various physical systems, there is cross-cutting infrastructure beyond that of systems software. Before any physical system can be simulated computationally, it is necessary to represent the geometry of the system computationally and to build a grid (or mesh) filling the physical region on which to solve the governing partial differential equations, i.e. on which to do the computational simulation. This geometry/grid problem is a common underlying feature of

computational science applied from airplanes to oceans to biological systems, and happens to be my own area of concentration.

But the commonality of the geometry/grid problem has resulted in lack of ownership among the funding agencies. Small separate projects have been funded, but the coordinated effort of the magnitude to finally address the problem has yet to be mounted, with the result that geometry/grid remains a major pacing item in computational simulation for real applications. Numerous statements from the aerospace industry, in particular, attest to this fact. This particular problem area is not one of computational speed, but rather that the process requires days of person-time with the software systems presently available. So this is an area where the Nation needs strong coordinated and cross-cutting effort in support of computational science, and an example of an area that could be best addressed by a virtual center bringing together needed expertise resident in separate universities.

As the PITAC report notes, it is essential that university researchers, as well as Federal labs, have access to high-end computer facilities. Currently, major Federal funding programs are providing new and advanced high-end systems for DoD through the HPC Modernization Program and for DoE through the ASCI Program. But only NSF, through its two PACI centers is providing high-end facilities that are readily accessible to university researchers, although both DoD and DoE do involve specific university partners. The pyramid concept of regional high-end facilities supported by centralized facilities at the highest end that was put forward in the Branscomb report and is now applied to a degree in the NSF PACI centers and even in the DoD mix of Distributed and Major Shared Resource centers deserves, I believe, broader attention and application. There is a need for distributed facilities in university research centers, as well as for centralized highest-end facilities at Federal laboratories and the two NSF PACI centers, with networked distributed usage. On that Top 500 list, the three top universities are all in Japan.

## **Conclusion**

Finally, I had the opportunity to testify before the Science Committee about this time last year on the need for high bandwidth connectivity to universities regardless of geographical location. That need is noted in the NGI Implementation Plan, released in February, and also in the PITAC report. Information technology is having a leveling effect on universities, allowing researchers to collaborate regardless of affiliation, and thus making it ever more important that geographical location not be a factor. This point is also noted in the recent science policy report from the House Committee on Science. And access to universities regardless of location is important not only to enable needed research collaborations but also so that graduates are produced with experience in high-bandwidth networks to meet the increasing workforce demand.

This difficulty of access has not, however, lessened the desire and preparation of universities in non-urban areas, such as in the EPSCoR states, for participation in Internet2. Seventeen of the nineteen EPSCoR states (one actually a territory) are represented among the universities in the Internet2 consortium. And 25 universities from EPSCoR states are in the consortium, with funds committed to establishing the local networks necessary for connection. With the latest round of NSF vBNS awards, 23 universities in 15 of the 19 EPSCoR states are included in the 128 universities with vBNS connection awards from NSF.

High bandwidth connectivity of universities is absolutely critical to the National effort in network technology.

Both Canada and Japan have recently announced such efforts, with Japan stating the direct intention of surpassing the U.S. in the information technology that now constitutes a very significant portion of our economic growth.

With information technology accounting now for some 30% of our economy, PITAC's call for fundamental long-term research has to be considered differently from such a call for research in fluid mechanics or nuclear physics. Information technology now constitutes fundamental infrastructure on which science, engineering, commerce, education, and even entertainment are being built. Never has a particular area of research been so critical to the Nation in such a fundamental and pervasive way.

And there is the matter of response time: we are now faced with almost immediate response time in our fundamental infrastructure - for better or worse. The negative and concerning factor here is that the time available for damage control is shrinking drastically. And this affects communications, the power grid, financial transactions, and most aspects of commerce and security. We are being placed by information technology into the position of not being able to rely on marshalling our own response to crisis; rather, we are becoming dependent on the reliability and security of the infrastructure system and its own capacity for response to failure. As response time decreases through the advance of hardware and software, our stability is thus a direct function of the software.

So the serious concerns raised in the PITAC report: that fundamental research in software, focus on long-term research, connectivity of universities, and more coordinated direction are - along with advances in high-end hardware systems - absolutely vital to the Nation's future cannot be stated too forcefully. This is a far greater challenge than any we have yet faced.

The recent science policy report from the House Committee on Science contains a very perceptive and far-reaching statement: "We must all possess the tools necessary to remain in control of our lives." That is software.

I want to thank the Chairman and all the members of the Subcommittee for this opportunity to speak with you today, and I especially want to thank you for your interest and support in this matter of great importance to the country.