

## Onward to Petaflops Computing

David H. Bailey

December 6, 1996

With programs such as the U.S. High Performance Computing and Communications Program (HPCCP), the attention of scientists and engineers worldwide has been focused on the potential of very high performance scientific computing, namely systems that are hundreds or thousands of times more powerful than those typically available in desktop systems at any given point in time. Extending the frontiers of computing in this manner has resulted in remarkable advances, both in computing technology itself and also in the various scientific and engineering disciplines that utilize these systems.

Within the month or two, a sustained rate of 1 Tflop/s (also written 1 teraflops, or  $10^{12}$  floating-point operations per second) is likely to be achieved by the "ASCI Red" system at Sandia National Laboratory in New Mexico. With this objective in sight, it is reasonable to ask what lies ahead for high-end computing.

The next major milestone is a sustained rate of 1 Pflop/s (also written 1 petaflops, or  $10^{15}$  floating-point operations per second). It should be emphasized that we could just as well use the term "peta-ops", since it appears that large scientific systems will be required to perform intensive integer and logical computation in addition to floating-point operations, and completely non-floating-point applications are likely to be important as well. In addition to prodigiously high computational performance, such systems must of necessity feature very large main memories (somewhere between 10 Tbyte and 1 Pbyte, depending on application), as well as commensurate I/O bandwidth and huge mass storage facilities. The consensus of scientists who have performed initial studies on in this field is that "affordable" petaflops systems may be feasible by the year 2010. [PAWS]

To get some idea of the scale of a petaflops system, such a computer could dispatch in three seconds a computation that a typical state-of-the-art desktop system today would require a full year to perform. One Pbyte of memory is equivalent to the text content of approximately one billion books, or in other words about the combined libraries of one thousand universities. If one were to attempt to construct a petaflops system today, even if one employed low-cost personal computer components (ignoring for a moment the daunting difficulties of communication and software for such a system), it would cost some 50 billion dollars and would consume some 1,000 megawatts of electric power.

The need for such enormous computing capability is often questioned, but such doubts can be dismissed by a moment's reflection on the history of computing. It is well known that Thomas J. Watson, a founder of IBM, once ventured that there was a worldwide market of only about six computers. Even the legendary Seymour Cray, who recently

passed away following a tragic auto accident, designed his Cray-1 system on the premise that there were only about 100 potential customers. In 1980, after the Cray-1 had already achieved significant success, an internal IBM study concluded that there was only a limited market for supercomputers, and as a result IBM delayed its entry into the market.

In contrast, some private homes now have more than Watson's predicted six systems. Further, currently available personal computers have computational power and main memory comparable to or exceeding that of the Cray-1, and enthusiastic users are eager for more. Indeed, the increasing power of personal computers now poses a threat to the scientific workstation market, which is moving to advanced symmetric multiprocessor systems partly in response. But however this battle turns out, it is indisputable that scientists and engineers will demand ever more powerful desktop systems for their work.

There certainly is demand for more powerful systems at the high end of scientific computing, as scientists continue to press the outer limits of physical simulation. High-end systems traditionally have been the province of academic and government research laboratories. But in a significant recent development, parallel supercomputers are increasingly being used by persons in other arenas, including financial analysts in the Wall Street community and marketing analysts in the consumer banking and retailing industry.

In short, the demand for state-of-the-art computing power appears insatiable. Thus we may as well start planning now for petaflops systems. Some of the compelling applications anticipated for petaflops computers include the following [Petaflops]:

1. Nuclear weapons stewardship.
2. Cryptology and digital signal processing.
3. Satellite data analysis.
4. Climate and environmental modeling.
5. 3-D protein molecule reconstructions.
6. Real-time medical imaging.
7. Severe storm forecasting.
8. Design of advanced aircraft.
9. DNA sequence matching.
10. Molecular simulations for nanotechnology.
11. Large-scale economic modeling.
12. Intelligent planetary spacecraft.

To elaborate on just a single item, consider the application of 3-D protein molecule reconstructions, also known as the "protein folding problem". In designing a new drug agent, scientists need to examine many protein molecules, each with a specified nucleotide sequence. But at present it is not possible to determine, except by experiment, the actual three-dimensional structure of the resulting protein molecule. And without this knowledge, it is not possible to know whether the molecule will have the proper binding sites to

be an effective agent. Petaflops computers may be powerful enough to do the necessary computations to determine this 3-D structure in a reasonable amount of time. Needless to say, such a capability could be a powerful new tool for the pharmaceutical field.

Some of these anticipated petaflops computer applications will be scaled-up versions of present-day applications, with evolutionary enhancements. Others will consist of integrated simulations of multiple physical effects. Many of these applications will likely employ advanced visualization facilities, such as immersive or remote visualization environments, that are still under development today. But if the history of computing is any guide, a number of exotic new applications will be enabled by petaflops computing technology. These applications may have no clear antecedent in today's scientific computing, and in fact may be only dimly envisioned at the present time.

In spite of such potential, it is not at all certain that the evolutionary advance of scientific computing systems, as produced by private industry, will achieve usable petaflops systems by 2010. One of the reasons for this conclusion is the recent turmoil in the scientific computing marketplace, which has led computer vendors to cut long-term research in favor of near-term development, and to focus on the more lucrative low- and mid-level systems instead of high-end systems. This phenomenon has been described as the "truncated pyramid" of the current computing marketplace. Thus it is likely that government agencies will need to provide a substantial part of the required research and development funding to make these systems a reality.

Beyond purely economic considerations, there are a number of difficult technical problems that need to be solved in the next few years if we are to achieve the goal of petaflops computers by the year 2010. Indeed, the anticipated difficulties of developing the underlying hardware technology, determining an optimal system architecture, producing effective system software, devising efficient algorithms, and ultimately of programming petaflops systems present challenges unprecedented in the history of computing.

A key issue for these systems is latency management. When citing the breathtaking increases in memory device density during recent years, a consequence of Moore's law, we often forget to note that the access time of these memory devices has not improved very much during this time, nor is there any reason to expect dramatic improvements in the foreseeable future. Thus the gap between processor speed and memory speed is expected to worsen in the future. Significant advances in processor technology may accelerate the feasibility of petaflops-level performance, but they will only further exacerbate the challenge of latency.

Latency can be dealt with by exploiting concurrency, such as in pipelined or multi-threaded architectures. This fact, together with the need to achieve 1 Pflop/s aggregate sustained performance, will mean that enormous system concurrency will be required. For instance, even with optimistic projections of future processor power, it is likely that

petaflops systems will incorporate at least 100,000 processors and possibly one million. Concurrency of this scale is well beyond anything attempted heretofore in high performance computing. Indeed, the coupled challenges of managing concurrency and latency will drive much of the research that needs to be done. Some specific research questions that need to be answered during the next few years are the following:

#### Hardware

1. Can we produce a usable petaflops system using commercial, off-the-shelf (COTS) hardware components?
2. Will an exotic hardware technology approach, such as superconducting RSFQ logic [Polonsky] or optical interconnect technology [Jans], achieve the 1 Pflop/s milestone sooner or cheaper?
3. Will a multiple-instruction, multiple-data (MIMD) distributed memory architecture be satisfactory, or will some novel system architecture be required?
4. What hardware facilities are needed to manage latency and multiple layers of memory hierarchy?
5. How can mass storage and I/O be handled on such a system?

#### Software

1. What operating system design can reliably manage 100,000 to 1,000,000 processors?
2. Are radically new programming languages needed, or can existing languages be extended?
3. What specific new language constructs will be required?
4. What is the best way to support I/O, debugging, graphics and virtual reality?
5. What software facilities are needed to manage latency and the memory hierarchy?

#### Algorithms

1. Do there exist latency tolerant variants of known algorithms?
2. How will the operation count, memory requirement, data locality and other characteristics of various algorithms scale on these future systems?
3. Will variations of classical algorithms suffice for key applications, or must we find completely new algorithms?

#### Applications

1. Can anticipated petaflops applications be structured to exhibit the required 100,000+ concurrent threads?
2. What is the best way to implement various applications on proposed system designs?
3. What will be the memory and I/O requirements of future applications?
4. What completely new applications will be enabled by petaflops systems?

These research questions raise provocative issues about the future of all computing, not just high-end scientific computing. For example, high levels of parallelism are inevitable

for all classes of computing, even home systems. Thus it is likely that answers to these questions may have impact far beyond the realm of large-scale scientific computing.

There is already a growing research community working on these and related problems of petaflops computing. For example, recently the National Science Foundation awarded a number of research grants to explore system architectures for petaflops computers. These projects presented on their proposed designs at the recent Frontiers '96 conference, held in Annapolis, Maryland at the end of October. More studies are planned. We look forward to the findings of these investigations.

Onward to petaflops computing!

### **References:**

[Jans] J. Jans, "Planar Packaging of Free-Space Optical Interconnections", *Proceedings of the IEEE*, vol. 82, no. 11 (1994).

[PAWS] Report of the Petaflops Architecture Workshop, to be published, 1996.

[Petaflops] Proceedings of the Petaflops Summer Study Workshop (Aug. 14-23, 1995, Bodega Bay, CA), [http://www.anl.gov/summer\\_study](http://www.anl.gov/summer_study).

[PetaSoft] Report of the PetaSoft Workshop, to be published, 1996.

[Polonsky] S. V. Polonsky, A. F. Kirichenko, V. K. Semenov and K. K. Likharev, "Rapid Single Flux Quantum Random Access Memory", *IEEE Transactions on Applied Superconductivity*, vol. 5, no. 2 (June 1995), pg. 3000-3005.

[Sterling] Thomas Sterling, Paul Messina and Paul Smith, *Enabling Technology for Petaflops Computers*, MIT, Cambridge, 1995.