Filing Information
October 2000
IDC #
Volume:
Tab:

Workstations and High-Performance Systems

Bulletin

September 2000 HPC User Forum: First Annual Meeting Notes

Analyst: Earl Joseph II, Ph.D., Christopher G. Willard, Ph.D. and Debra Goldfarb

IDC Opinion

The need for improved performance metrics is being addressed by a combination of users from industry, government and academia along with vendors of technical computers. The IDC2000 initiative was presented, debated and launched at the meeting. Users and vendors supplied their written feedback that was then presented and discussed during the meeting.

The initiative has two parts: 1) to develop a new, more meaningful industry standard benchmark; and 2) to develop a matrix of metrics for use to evaluate unique.

We thank Larry Davis and Steve Conway for their excellent summaries of the meeting that were used to help form this overview.

Executive Summary

The first annual meeting of the HPC User Forum was held in Richmond, VA on September 18th and 19th. Approximately 65 people participated in the meeting, with representatives from government, industry, university and all major HPC vendors. The full users forum meeting was preceded and followed by short meetings of the group's steering committee.

Opening Steering Committee Meeting

Earl Joseph convened the first steering committee meeting and discussed the structure of the user forum, including the newly-formed Performance Advisory Group. He suggested more regular meetings of this group, perhaps via conference calls. He also presented the list of user issues as judged by the members. The sizing of computers and mapping of applications to architectures was clearly the dominant issue, and was the principal discussion topic for the meeting. George Cotter of NSA made a plea for improved, more organized user collaboration. He presented a proposal that the User Forum organize as a formal structure with officers and a board of directors.

Performance Metrics: Issue and Solutions

The plenary meeting session ensued with discussion on the development of metrics and benchmarks to ascertain HPC system performance. Dr. Joseph introduced the topic and pointed out the shortcomings of LINPAK as an appropriate benchmark for HPC systems. He also introduced the two phases of the IDC2000 Users Forum benchmarking plan:

1) development of new simple benchmarks and metrics to characterize performance, and

2) the creation of a table of applications benchmarking results contributed by the member organizations that would be available to all members.

Larry Davis chaired the session on performance, which included presentations by the HPCMP, NSA, Ford, Boeing, Pittsburgh Supercomputer Center (PSC), SDSC, and NERSC.

- HPCMP -- Terry Blanchard gave a presentation, which provided an overview of their ongoing benchmarking activities including their use in current HPC hardware acquisitions.
- NSA -- Mike Merrill presented the NSA's principal benchmark (GUPS), which measures performance on random memory accesses. This led to a discussion on the general issue of memory vs. CPU performance and their relative importance in HPC systems, which continued later in the panel discussion.
- Industrial users -- Alex Akkerman of Ford and Barry Sharp of Boeing discussed the types of application benchmarks that they use. They stated that, if the HPC hardware vendors agreed to

release the results, they would be willing to contribute benchmark results to the User Forum activity to create a table of application benchmark results.

- PSC -- Jim Kasdorf discussed their use of benchmarks. He strongly supported continued updates to the NAS parallel benchmarks.

- SDSC -- Allan Snavely discussed the general problem of how to predict HPC system performance over a range of application problem sets from a sparse set of benchmark data points. He proposed a new effort to develop scalable synthetic benchmarks based on a careful study of how applications codes behave that would produce a “system signature” of performance in several performance dimensions. He discussed a simplified performance model for memory-bound applications that seems to produce a reasonable estimate of actual system performance in this case. He noted the complexity of the problem and stated that accurate predictions of HPC system performance can only come from an accurate understanding of both the applications software and the hardware.

- NERSC -- Bill Kramer and Bob Lucas discussed the NERSC effective systems performance test, which measures a system’s efficiency in scheduling and executing a typical system workload, and a new set of synthetic application benchmarks that will focus on memory and communications performance. This set of synthetic applications benchmarks form the heart of the IDC2000 initiative to define a more useful systems performance metric than LINPAK. NERSC’s eventual goal is to combine this set of benchmarks with their effective systems performance test to accurately characterize a system’s performance in a multi-user environment.

Dr. Joseph opened the session to general discussion and noted that the first milestone for the IDC2000 initiative would be the announcement of the selection of the first set of pseudo-applications to be used as benchmarks by SC2000. Eric Strohmeier reviewed some of the details of the IDC 2000 initiative.

The discussion began with feedback on the IDC2000 performance metrics plan from the three sets of users represented at the meeting: government, university and industry, and HPC hardware vendors. Each group generally endorsed the plan. Government users included a request to the vendors to allow release of benchmark results on their systems. The hardware vendors expressed a willingness to dedicate resources to benchmarking efforts, but asked that government acquisition officials be mindful of the effort required to perform benchmarking activities. The performance panel then discussed the relative importance of CPU vs. memory performance. There was general agreement that a balance between the two was necessary for systems serving broad applications areas.

Vendor Presentations

Several of the hardware vendor representatives provided short presentations on current products and future directions. Vendors represented included SRC, Fujitsu, Compaq, IBM, SGI, Sun, and Cray. Of notable interest was the recent success of Compaq in capturing several very large acquisitions (ASCI 30 TF system and PSC NSF system). In addition, Etnus, which makes the TotalView parallel debugger, gave a short presentation on its products and plans.

IDC Research Update

IDC provided a brief update on a number of different HPC research studies including automotive/aerospace research, clustering trends, Linux market directions, vendor market shares and HPC forecasts.

ASCI Program Update

Jose Munoz of the DOE presented an update on the ASCI program. The ASCI White system at Lawrence Livermore National Laboratory is currently being constructed, and will have a 12 TF IBM SP P3 system. The ASCI Q system, planned for 2002, will be a 30 TF Compaq system with Alpha processors. Mr. Munoz presented some utilization results from the ASCI Blue Mountain system (SGI Origin architecture) that showed typical utilization on the system ranging from 60 to 80 percent. He stated that most of ASCI's current simulations run on 2,000 to 4,000 processors, and that the organization is working to determine which kinds of platforms best run various applications. He discussed the need to maintain overall system balance over several system parameters and presented diagrams showing capabilities in these dimensions for ASCI's current and future systems. He also discussed scientific visualization activities.

The presentation created considerable discussion concerning ASCI's acquisition policies, which seem to stress maximum CPU performance in peak GFs. Mr. Munoz pointed out that although this may be true, ASCI has made major investments in software applications to be able to use these theoretically very powerful systems efficiently.

Singapore Institute for High Performance Computing Introduction

David Kahaner of the Asian Technology Information Program then introduced Dr. B.T. Cheok of the Singapore Institute for High Performance Computing. This organization has a mandate to use leading edge high performance computing resources to enhance Singapore's global competitiveness. The institute sponsors research collaborations as well as programs aimed at developing human capital. It also operates HPC systems for Singapore industry and academia. Dr. Cheok extended an open invitation for User Forum members to visit them in Singapore.

Distributed Computing Update

Paul Muzio and Steve Karwoski presented reviews of their respective distributed centers. Both discussed their respective center's capabilities and activities.

IDC Forecast Update

Chris Willard of IDC then presented an overview of the IDC market forecast for the technical computing market, accounted for $5B in 1999. The IDC report segmented the market into the following four categories: capability, enterprise, divisional, and departmental. In response to a question, high performance computing was described as including systems in the capability and enterprise categories, with the dividing line between enterprise and divisional systems being $1M.

Dr. Willard also presented IDC's detailed taxonomy of HPC systems categorized by memory and processor distributions. IDC forecasted fairly strong growth in the HPC marketplace over the next several years, with a higher growth rate for total technical computing. Dr. Willard also presented a market forecast for the use of Linux in HPC systems. It shows substantial Linux growth through 2004, but predicted that Unix will remain the primary operating system for technical computers throughout the period.

Closing Steering Committee Meeting

The final steering committee meeting focused on the organization of the HPC User Forum. NSA again presented a proposal to formalize the organization immediately with a group of officers and a board of directors. There was general agreement that a tighter organizational structure was needed, but also general skepticism that it needed to be as formal as NSA's recommendation. The compromise position, which seemed to gain a consensus, was that we should immediately form a small group of interested participants that would guide the forum. It is probably essential that this group (whether it is called a steering committee or a board of directors) have representatives from each major user organization, but also be kept to 10-15 members so that progress can be made. It was also generally felt that IDC should provide the executive director for the forum. IDC will review the surveys that each user completed stating their willingness to be involved at several different levels and then work to form the small group that will manage the user forum activities.

Meeting Notes – Full Version

OPENING STEERING COMMITTEE MEETING: 9/18/00

Introduction (Earl Joseph, IDC)

Reviewed July meeting notes and the efforts made to address issues raised then:

- Logistical issues addressed included: more elbow room in meetings, more control over presentations and break times, and name tags.
- Operation of Performance Advisory Group (PAG) and SIGs (*currently in progress).*. Only 5 people responded to email request for opinions. Dr. Joseph recommended that the group move to schedule conference calls once a month for PAG.

Reviewed the agenda and logistics for the full User Forum meeting. Explained new CD-ROM membership card. Proposed PAG Operating Procedure:

1.PAG reviews a proposal and creates a recommendation.

2. Steering Committee reviews and approves/changes/cancels.

3. All members review & comment.

4. IDC implements.

Discussion:

- Suggest the PAG start with a conference call to provide structure, then maybe a chat room would be meaningful.
- Conference call participation. There are currently 18 members in the PAG, so we would expect 7-8 participants in a each conference call.

Prioritizing the Issues For the User Forum To Address:

A discussion was held on prioritizing issues for the user forum, and on the overall organization of the forum. The following is a record of this discussion.

Earl Joseph: The priorities have been consistently rated in the last 3 meetings. The highest priority is performance measurement, then mapping applications to architectures. On some issues, rankings were bimodal: e.g., policy recommendations were ranked most important by about 1/3 of members, not very important by 2/3. These situations call for SIGs to address the issue.

George Cotter: I continue to be troubled by the lack of serious organization on the user side to pursue the issues that are causing this community so much trouble in so many different areas. I was reviewing material I pulled together about this User Forum a year ago. According to my vision at that time, the balance between IDC and the users themselves is out of whack. It's hard for me to see this User Forum being able to tackle the really tough issues (without more user involvement). For example, performance measurement: a lot of study and work needs to go into this. We need a serious user-oriented, user-populated group to come up with an agenda and a set of actions, and how to deal with them, how to organize in collaboration with IDC. The IDC viewpoint has consistently been that the user community won't organize itself and put the requisite effort into this without substantial support from IDC. I think the balance is wrong (and I would like to see the users take a more active role).

Dr. Joseph: We wouldn't put it in those negative terms. We'd say users are too busy and IDC is providing a service. The response from this group (Steering Committee) will tell us which way is best to follow.

Mr. Cotter: What's happening in high end computing that makes improved user collaboration essential? Market economics. System architectural directions. Little long-term R&D. Fragmented linkages to academia. Policy mechanisms are often erratic. (However, there is) high payoff when collaboration occurs.

The Steering Committee should only exist long enough to get things started. By organization, I mean having a Board of Directors, with officers, committee chairs, directors-at-large, an executive director (IDC). Also officers (president, VP, secretary, treasurer). Committees: membership, programs, policy, etc.

Dr. Joseph: Right now, every user member who joins is eligible to be on the Steering Committee. We need to deal with this and decide if Steering Committee membership should be more limited, and how it should be organized. We also look at the Steering Committee as a Board of Directors, in the sense that users set the issue agenda, determine what needs to be done, reviews the progress and changes the direction as needed. IDC provides a service in organization, in researching the issues, in making proposals, and in seeing that a solution is implemented the helps the industry.

Comment: Maybe IDC was thinking that the type of organization Mr. Cotter outlined would emerge over time as the forum size grows. Maybe now is the right time for some additional structure. Maybe certain objectives need to be stated, for example that architectural trends are very bothersome to some members of this group. We don't know which architecture will provide the best cost/performance in the next few years and are concerned some may go nowhere.

Mr. Cotter: The choice is between a tight structure run by users themselves, and a looser group run by IDC. We're drifting into the latter model where not much is being done by users between meetings.

Comment: I support a strong user community, but the flip side is to know where we're going. Sitting here today, it's difficult to make an argument for one architecture or another.

Mr. Cotter: I don't think the user community can come out with a position on where architectures should go. A well-organized user community would evolve a position on this.

Comment: We could be a driving force with vendors.

Mr. Cotter: We can't work those issues in large meetings like this. That needs to be done through SIGs.

Comment: We should present our requirements. We're not architects.

Comment: I think the disconnect is in mapping applications to architectures. There's a middle ground where users and vendors need to collaborate.

Comment: This would also benefit vendors. Vendors don't know if their architectures will succeed. They don't know what they have to do to succeed with users.

Dr. Joseph: If users want to form more committees, etc., we're here to support that and will gladly help in every way. But if users don't have time for this level of participation, IDC can do more to drive the process. The process is that the user steering committee is like a Board of Directors in setting the issue agenda, setting and changing the direction of the forum. IDC is here to help drive the process, conduct the research and make positive changes for the overall industry.

Debra Goldfarb, IDC: Let's do this as a written vote (structural options for User Forum). We'll do up a written ballot and pass it around. (This was done and then discussed at the end of the forum meeting).

FULL USER FORUM MEETING

Introduction (IDC)

Debra Goldfarb: [Welcoming remarks. The heart of the day will be a discussion of the performance issue.]

Earl Joseph: [Explained packet and walked through agenda.]

"Performance Metrics: Issue and Solutions"

Dr. Joseph [PowerPoint]: review of prior User Forum discussions on this issue.

In 2000, the range in prices as a ratio of peak performance was 8x. Ergo, peak is nearly meaningless. We want to promote a more meaningful industry standard benchmark for procurements, and also create a matrix with many performance metrics and applications results.

Advice from one user (warning): We need to get it not just technically correct, but think about who will be using the numbers.

If the benchmark is too complex, no one will like it. The PAG should resist complexity at every turn.

IDC2000 Initiative. (1) Quickly put in place a new standard metric; (2) then create a matrix of applications and kernels for use in procurements.

LINPACK correlates strongly with peak, so it's also nearly meaningless except as a measurement of CPU speed.

Session on Performance Measurement

Larry Davis, from the DOD Modernization Program moderated a panel session on performance measures and strategies currently in use by Forum members.

Mr. Davis: This performance measurement topic is extremely important to us. We've structured this session to show some of range of benchmarks being used out there by users.

Terry Blanchard, NAVO: "Benchmarking for HPC System Acquisitions."

HPCMP Goals: Provide computer services to the R&D community within the DOD. We are looking for production systems that can provide cycle services to a broad band of users.

There are four major initiatives within the program, four major centers, 17 distributed centers. Networking and specialized software support is also provided.

Next procurement: we refresh our technology every year. This will be the first program-wide refreshment, as opposed to each center doing its own. There are ten computational technology areas, and about 5000 users. Each center provides support for some, not all of the application areas. CFD is supported at all four major centers.

Acquisition criteria – wide range of criteria. Performance, cost, support, confidence, upgrade, capabilities.

Every year we survey our users to find out what their requirements are. We categorize these requirements by HPC system class: distributed memory, shared memory, parallel vector. Acquisitions made for future HPC systems. Decision-makers must consider both technology push and requirements pull.

Benchmarking team: formed to develop benchmarks useful over long period of time, as well as for next refreshment round. Representatives from all four Major Shared Resource Centers. We plan to share information we gather with the HPC community.

Benchmark objectives: based on operational requirements.

Benchmark structure: three types of benchmarks: (1) applications codes (complete or code essence, but not kernels) representative of all our applications areas, executed with varying CPU counts; (2) synthetic benchmarks: HW, algorithms, OS, file system, etc.; (3) mixes: based on utilization by applications across the program…two-way. Goals: to understand workloads on a system, and how the system will scale with upgrades.

Benchmark rules (draft): make runs with no coding changes; run job mixes with production queuing systems; rules intended to encourage vendors to provide operational systems, not benchmark systems.

Schedule: mid-October rules (draft version); November 13 official release; January 22 results due; refresh and update.

Comment: You haven't mentioned specific benchmarks.

Reply: We know some we'll use. We'll decide on others based on information recently gathered from users about their requirements.

Comment: Will you share this information?

Reply: Yes, and we hope others in this community will share their benchmarks with us.

Comment: How do you decide which platforms to benchmark on?

Reply: We release so all vendors can respond if they choose to.

Q: How many vendors do you expect to respond?

Reply: 4-6, probably.

Mike Merrill, NSA: GUPS

GUPS = Giga Updates Per Second. You create a really large table, then you count how many times per second you can update it. You're updating 64-bit words. GUPS measures how well memory, bandwidth and processors are balanced. Motivation: We need a significant amount of random memory access (30% in a lot of our 50 application codes). Our code ~NPB CG, but also includes reads and writes.

GUPS per $10M Spent, 1986-2000 (graph). Peaked in 1996 at 3.0 with Cray T90, then sharp decline to 0.3 once T90 was no longer available.

Alex Ackerman, Ford.

Benchmark approach: Usually done on annual basis. Attribute based (e.g., safety, durability, CFD. We look at different one of these each time). Internal models (applications): published results (papers) may not be applicable, especially. when it comes to scaling.

All vendors are invited to participate. We establish guidelines as to how we want the systems configured.

Comment: Do you have problems getting vendors to configure systems large enough?

Reply: Not really. We tend to stick to existing rather than future systems.

Throughput performance (mix of multiple jobs for a given period of time). Accuracy is evaluated by end users. Performance analysis (usually based on throughput) – we only care about time-to-solution, not flops. Decisions are based on a combination of performance and price/performance.

Comment: Do you factor in 'pain' of moving from an existing to a new vendor?

Reply: Yes, these are serious issues. We still have people running on C90s and we have a hard time getting them to move off of those. We have machines from almost every vendor today.

Comment: How long does it take to create the benchmarks, and how often do you do it?

Reply: Comes out to about once a year = about once every 4 years for each of our 4 attributes.

We still do much of our safety work on T90s, nothing else matches that but we can get close on other machines at lower cost.

Comment: Future goals?

Reply: We'd like to eliminate all this and move to a standard HPC benchmark. We'd be happy if software vendors would supply us with small kernels for benchmarking and also new versions of applications.

It would be good if vendors would allow us to share performance results with the community, but not possible today because of two-way NDAs.

Vectors are still the only machines where we can solve our large capability problems. These systems have stood still for about 8 years, so we're solving the same problems in the same amount of time.

Powers of 3 (price/performance): clusters, cluster systems, SMPs, vector. You pay a price for performance. So far our management has been okay with that, and we think that will continue in the future.

Comment: If the tariff came down would you buy Japanese?

Reply: We might need to.

Comment: We always look at price/performance but it really comes down to the cost-effectiveness of the system. E.g., Cray T90 cost-effectiveness is pretty equal to the others despite the big price/performance difference.

Comment: The fact you are still staying with Cray T90 shows price is not as crucial as business results.

Reply: We see it as different classes of problems. Some make most sense to run on PCs. Others can't be solved on PCs in reasonable amount of time.

Barry Sharp, Boeing

We're interested in 3 areas: Structural codes, CFD, propulsion. Very similar to Ford approach: looking at machines and how they perform.

Comment: Can you contribute to the [User Forum] applications matrix?

Reply: I don't see why not.

We're primarily project-driven. We run a single piece of code needed for a project, then make our decision based on that.

Comment: Maybe in this group we've been attaching too much importance to price vs. performance.

Two years ago, FAA wanted Boeing to analyze an issue with one of our planes (flap skew). We needed to find out quickly where this code would run well. Found to our surprise it ran well on a machine we normally wouldn't consider. Another example: TWA accident requiring enormous amount of study on center fuel cell explosion. Did near-real time analysis on Cray T90. This in itself required us to procure hardware to support this effort. Third example: model landing gear on Boeing 767 rather than physically test, to save lots of time and money by modeling. Saved Boeing $5-6M.

Comment: How does Boeing make these quick decisions on hardware procurement?

Reply: We first look at what our existing machines can do. We base decisions on time/cost of simulation vs. experimentation, and also consider cost of not doing it.

Jim Kasdorf, Pittsburgh Supercomputing Center

Our interest: procurement benchmarking, real applications benchmarking. (generally not ISV codes), system performance for large jobs (not capacity). PSC's Teraflop Computing Scale (TCS) Design. For a broad range of applications, memory-intensive, communications-intensive, I/O-intensive. We don't have to solve all the problems; there are other machines in NSF program. Needed real performance numbers on real codes.

We think NAS Parallel Benchmarks are very good.

Comment: In comparing IBM and Compaq, how much did you take into consideration system software for exploiting the capabilities?

Reply: Quite a lot. Neither vendor's system is a walk in the park.

Comment: This system replaced what?

Reply: Nothing. It's a new capability for the research community.

Alan Snavely, San Diego Supercomputer Center: "Performance Modeling & Prediction"

How do you gain insight into how algorithms map to architectures? Need to understand every performance characteristic of the hardware and software; understand how algorithms are implemented in real applications (need to know how an algorithm stresses a machine); develop techniques for mapping characterized applications to systems…this is the hard part!

First, the standard stuff: Complete a table of "market-techture" specs = vendor-supplied… Difficult, because vendors don't want to provide specs where their systems don't look good. Better: a comprehensive table of demonstrated specs (observable performance is almost always less than what vendors claim).Maintain table of traditional benchmarks and applications run on systems. All of the above are limited, crude comparisons. We're looking to extend standard techniques with more powerful methods and tools, similar to GUPS.

- Scalable benchmarks that give a picture of memory performance across a wide range of problem sizes, memory access patterns, etc.
- Pre-execution and run-time tools for profiling algorithms and applications.
- Performance algebra for mapping quantitative models of applications to characterized hardware systems.
- Simulators and meta-simulators.

MAPS – SDSC's Scalable Benchmark

MAPS for machine comparison

- Alpha and MTA…shows MTA much better for GUPS-type problems
- Performance prediction
- Complex but tractable by fundamental methods
- Software attributes influence mapping

Work to be done:

- Come up with scalable benchmarks
- Develop precise methods for mapping functions to computers
- Need profiling tools
- Need software tools that give us a guess about how the machine will do

Comment: Seems like modeling processors is getting harder and harder. How little work do we have to do to get useful predictions? One thing is write a simulator at the architectural level, as Cray does. Problem is that these simulators are very slow, so tough to do real applications in this mode. Goal should be to model just enough (simplify and abstract enough) to be reasonably predictive.

Comment: What are the 4 to 5 key attributes of applications?

Reply: Memory access = number 1; peak floating point rate next. If I can profile applications regarding arithmetic operations and memory access, that gives me a good start.

Comment: Your current benchmarks are ok for a uniprocessor.

Reply: Absolutely. We need to address system-level performance.

Bill Kramer, NERSC

"The NERSC Effective System Performance Test". "The real value of an information system is properly measured by ANSWERS-per-month, not bits-per-microsecond." "Fallacy – MIPS (or MFLOPS) is an accurate measure for comparing performance among computers.

NERSC deals with broad range of applications/disciplines Impact of increased effectiveness. If you can increase efficiency from 55% to 90% over system lifetime (18 months), you get Moore's Law improvement at no additional investment

We look at theoretical peak vs. efficiency/utilization to measure performance. How much scientific work can be done for a given quantum of effort?

Concept of the test = simulate/measure a day in the life of an MPP. Ability to utilize a large computer is almost as important as the speed of the computer. Large capability mainframes rarely had idle cycles

Robert Lucas, NERSC

The NERSC 'NAP' Test for application domain performance, useful for a broad community of users. NAP is a NERSC project:

Create a small set of synthetic application benchmarks. Can be used on variety of HPC systems. Tests different system attributes . Useful for 10+ years.

Performance: how much scientific work can be done for a given quantum of cpu time. Test all aspects of a system – CPU, Memory, Communication, I/O simultaneously. Criteria: synthetic application benchmarks; data structures similar to real applications; formally defined descriptions; reference implementations; scalable problem sizes; performance attributes independent of problem sizes; run times independent of problem sizes on 'balanced' systems

IDC2000 Benchmark Milestones:

- Phase 1: select 3 applications or kernels to address each performance area
- Phase 2: improve portability, provide flexible problem size definition, run initial experiments and analyze results

- Phase 3: define benchmark description, release ‘officially; collect performance results
- Phase 4: analyze these performance results to see if objectives have been met; prepare parallel reference implementation.