FINAL REPORT

National Science Foundation Award ID 1013639

Program Officer Name: Dan Katz

EAGER: Impact on Scientific Discovery using Simulation-Based Engineering & Science

NCSA at University of Illinois at Urbana-Champaign

PI: Merle Giles

October 31, 2012
1) EXECUTIVE SUMMARY

The NSF community’s mission to support for all fields of fundamental science and engineering, stands to benefit from studies in science related to industrial simulation-based engineering & science (SBE&S). Leadership for this report came from the National Center for Supercomputing’s (NCSA) Private Sector Program (PSP) at the University of Illinois at Urbana-Champaign (UIUC), which gathered data from its partner company user base and from a comprehensive survey conducted by International Data Corporation’s (IDC) HPC Division. This report’s key focus is on specific ways in which advances in science are needed to accomplish gains, even breakthroughs, in simulation capability.

Both academic and industrial science communities will benefit from increased understanding of industrial SBE&S needs and activities, which are increasingly becoming multi-disciplinary. An increased understanding is timely since manufacturers in the U.S. FORTUNE100® admit that computational efforts are a) largely limited to current production efforts, b) are too-often limited to steady-state and single-phase modeling and c) fail to adequately simulate multiple components as they are actually assembled. ii In other words, the nation’s largest manufacturers readily admit that limitations exist in current-state high-performance computing (HPC) simulations, and that barriers to additional complexity are rooted in an inadequate understanding of the fundamental science.

2) KEY FINDINGS

All but one of the IDC survey-responding sites (96.3%) said technical computers were installed directly onsite that are used to support R&D, engineering, or science. Computational fluid dynamics (CFD) and finite element analysis (FEA) were tied as the most commonly occurring, most important methodologies. The foremost disciplines underpinning the top 3 applications were fluid mechanics/aerodynamics, mathematics, physics, bio/life sciences, meteorology/climate science, and chemistry. The top 3 programming models used were MPI, OpenMP, and shared-memory. iii

Fewer than 1 in 6 organizations in the IDC survey claimed their applications, as now written, would meet their requirements for the next five years. Nearly 4 in 10 sites said the underlying mathematical/algorithm needs to be improved. More than 3 in 10 responded that the underlying science needs to be improved, yet more than one-half (53%) of the respondents admit that it is not always easy to distinguish between needed improvements in the underlying domain science and in the related computational and computer science. iv 30% of the organizations named specific ways in which they now have to “dumb down” their problems in order to complete the runs in reasonable amounts of time, with the most frequently occurring strategies being the use of coarser meshes than desired and not fully exploiting the known science. v Even so, a surprising number of respondents also claimed that a significant limitation to achieving increased realism in their simulations is the limitation of the underlying science. vi

A large majority of the respondents (81.4% to 92%, depending on the application) believed that today’s known science could support a moderate or a large amount of additional realism in the applications. An important subset of the survey respondents (37%) said that taking the next step — advanc-
ing the known science — could add even more realism to their key applications. Depending on the application, from 91% to 100% of this subgroup believe that advancing the science could add "a moderate amount" or "a large amount" of realism to their most important codes.

The main technical limitations preventing the sites from dramatically increasing the problem size or other dimensionality of their simulations were models/algorithms not scaling enough, inadequate latency/bandwidth, inadequate processing power, and the need for advancement in the underlying science. Three-quarters of the responding sites said their people can handle today's technologies but need retraining to handle next-generation systems. Only one respondent said their workforce is ready to provide technical leadership for using next-generation HPC systems. 92% of the responding organizations said that additional investments in the science underlying their key applications could make a moderate impact, defined as 10–25% improvement, or a great impact, defined as more than a 25% improvement.\textsuperscript{vii}

Perhaps a prime example of this technical/science confusion was observed in NCSA's graduate team survey. Of the two dominant methodologies used among manufacturers (CFD and FEA), CFD is known to scale rather well using HPC, yet this team observed that a CFD simulation at one company scaled to 1000 cores while a CFD simulation at another struggled to use 32.\textsuperscript{viii} It was found that the underlying science differs in that the scalable domain is compressible fluids while the domain that struggles is incompressible fluids. What is not known, however, is whether this is a fundamental barrier in the algorithm or inadequate understanding of the fundamental science.

Likewise, structural materials simulations scale relatively poorly. As an example, Simulia shared at ISC’09 that its Abaqus code is used 90% of the time on 4 computational cores or less, and 80% of the time on 8 cores or less.\textsuperscript{ix} What is not known is whether these barriers are due to the fundamental science or algorithms. Licensing costs have been cited as economic barriers that exist among a widely installed base of users at small and medium manufacturers, which further confuses the issue.

A face-to-face focus group with large-organization industrial HPC/scientists revealed that domain science in industry is often “Edisonian”, claiming that significant cultural gaps exist to believe that HPC is adequate to do real science. Engineers routinely run minimalistic models because they want to run on a laptop, so this drives gross errors. Removing barriers to scale will improve model fidelity, but these are cultural issues. Lack of scientifically accurate physics seems to contribute to this perception among manufacturers, yet specifics are hard to get. For instance, the science in some structural and fluid physics is adequate, but to achieve realism 10-100X compute power is needed in the same time frame. Output is lacking detail, so digital simulations must still be validated with physical testing.

Likewise, time-to-solution is important in life sciences, so capacity computing provides critical capability. Ensembles must be run to generate hypotheses; hence, the iterative process is real science, but the human/computer interface needs to be better. The “on-ramp” to build better models is simply too long. At the same time, focus group participants shared that realism in biology is unresolved, so better science is needed. Yet biology is really a digital science now, so there is an increasing need to better understand the atomistic level so as to get more accurate chemistry to devices.
Specific computational challenges exist in achieving strong scaling, with Navier-Stokes equations cited as key examples of mathematics that are 200 years old. Lack of strong scaling results in traditional engineers doing more of the same and not changing their expectations. Innovative user communities do have ideas, but don’t have access to sufficient computing power for development and demonstration. Large memory computers are helpful, as they offer a platform to try out ideas at scale.

It is important to note that for manufacturers (machine-based firms and those in life-sciences), the most important key to simulation is that it can help make better design decisions. Proper physics, therefore, need to be packaged for the designer, not just the professional engineer, presenting a challenge to the HPC community to package its capabilities for non-experts. Furthermore, there is great danger in using a single point solution to evaluate a design without regard to the variability associated with both the design and application. Industrial simulation experts complain that large-scale HPC systems are often built for science (both architecturally and politically), not uncertainty quantification. An example: attempts to run a commercial physics code in an HPC environment, when the application is optimized for use on a single computer, routinely results in the code thrashing back and forth between disk and memory, running for weeks. With sufficient memory it might run in minutes.

The UIUC business school graduate team surveyed three companies: a) one ranked in the FOR-TUNE50®, b) one in the FORTUNE250®, c) one was a medium-sized supply chain company. They concluded that:

a) All three companies are fundamentally limited by the current understanding of their applied science and engineering domains. Companies A and B, both of which produce large and highly-complicated products, are constrained by models that do not yet accurately capture the interaction of multiple domains and components in real-time. The two firms contend this is the most significant barrier to higher-level computational and simulation modeling for large OEM’s such as themselves, and that academic communities and relevant government agencies must understand and address this modern reality. Without significant advances in basic research, Companies A and B doubt they can significantly improve their computational and simulation modeling efforts, fully leverage petascale HPC systems, and increase their rates of innovation.

b) The computational and simulation modeling efforts of all three companies are limited by access to fundamental tools. With respect to software, each company reports being either somewhat or very limited by current ISV (independent software vendor) commercial licensing costs that exceed associated hardware expenditures by factors of 5-30x. The supply chain Company C is, in the words of its top technical manager, “desperate” for more affordable access to ISV codes.

c) Insufficient access to high-performance hardware resources is a common frustration, though the problem does not seem to be one of cost. Limited access to ISV codes and/or self-imposed policies means top engineers at each company can rarely use more than half, let alone the full power of, their most advanced technical computers. As a result, engineers can neither per-
form a desired number of simulations nor optimize the algorithms or scaling capabilities of proprietary codes. Companies B and C emphasized that increasing simulation counts within existing fixed periods is critical to improving the quality and reliability of the science and engineering behind their products.

d) Just as important, it appears firms describe how they organize computational and simulation modeling efforts in different and unanticipated ways. Based on input from PSP, the graduate team expected the three companies to indicate most computational and simulation modeling is done within R&D departments. Instead, two of the companies described most of these efforts as occurring as part of product “design.” Here, “R&D” may have been confused for “basic research only,” exclusive of any design processes. In addition, each of the senior technical managers we spoke with had difficulty describing the exact ROI computational and simulation modeling delivers, or how they otherwise communicate its value to senior executives.

3) ACTIVITIES OF THE PROJECT

a) Funding through NSF’s EAGER program began July 1, 2010
b) Project completion: September 30, 2012
c) 2010: A graduate student team from UIUC’s business school engaged three (3) market-leading manufacturers possessing distinct competencies and wide respect for their science and engineering leadership. Each firm completed a probing and comprehensive survey designed to meet two objectives:
   i) Provide a highly detailed, stand-alone “snapshot” of the company’s computational and simulation efforts.
   ii) Position NCSA to make relevant comparisons across the manufacturing landscape, if and when more firms provide such feedback.

d) 2011/2012: on behalf of NCSA, IDC surveyed 30 organizations from industry (76.6%), government (16.7%), and academia (6.7%). Specific goals of this survey were to:
   i) Document which HPC applications manufacturers are using today and for what purposes,
   ii) Determine the scalability of these applications and their current limits,
   iii) Identify barriers preventing greater scalability, including core science issues and barriers,
   iv) Determine what methods manufacturers are using, if any, to get past these scalability limits,
   v) Recommend areas where additional investments and research could improve the scalability of applications.

e) IDC Survey focus group
   i) Participants: BP, Caterpillar, DuPont, Ford, NASA, National Cancer Institute
   ii) Questions and topics included:
      (1) Identify the factors that currently limit the sustained performance and realism of applications codes in scientific and related industrial domains.
      (2) Propose what could be done to improve the performance/realism of the codes.
      (3) How much do you think the sustained performance/realism of your key codes could be improved with sufficient effort?
(4) What steps do you think should be taken to improve the sustained performance and realism of key codes in your field?

(5) Are there steps that research organizations such as the NSF (National Science Foundation) and NCSA (National Center for Supercomputing Applications) could take that would broaden your utilization of supercomputing resources available at NSF and National Lab sites?

(6) What would you like to see improved the most in support of your technical applications?

(7) What are the main technical limitations to dramatically increasing the problem size, or dimensionality of the simulation (pertaining to the underlying science and computational methods used)?

(8) For applications where the underlying science needs to be improved, are there scientific improvements that can be accomplished in one to two years?

(9) For your key applications, how mature is the underlying science?

f) NCSA Private Sector Program Partner EAGER workshops and meetings:

i) May 2011 workshop topics included:
   (1) Defining Virtual Realism, Boeing
   (2) Multiscale and Multiphysics, UIUC
   (3) Virtual Duty Cycle Analysis, John Deere
   (4) End-to-End Product Design Exploration, John Zink
   (5) Industrial HPC Resolution Challenges, Rolls-Royce
   (6) Variability and Uncertainly Quantification, Caterpillar
   (7) Computational Engineering Challenges, Rolls-Royce

ii) May 2012 workshop topics included:
   (1) High-Speed Impact, GE, Boeing, Caterpillar, John Deere, Rolls-Royce
   (2) Speed – Why is iForge (NCSA’s industry-dedicated cluster) so Fast?, NCSA’s PSP
   (3) Why is Multiphysics so Difficult?, Illinois Roctar, Simulia, UIUC, NCSA
   (4) Workflow and Data, Nokia-Siemens Networks, Adaptive Computing, GE, Cray, Caterpillar
   (5) Is the Solver the Holy Grail?, NCSA, Lawrence Livermore, Sandia, Argonne
   (6) Modern Software Code Implementation, UIUC, Numerical Algorithms Group, NCSA
   (7) NDEMC – An American Success Story, John Deere, Council on Competitiveness, NCSA

iii) June 2012 International Industrial Supercomputing Workshop at NCSA
   (1) HPC centers from ten countries represented: USA, South Korea, South Africa, Sweden, Netherlands, Germany, Spain, Italy, Japan, UK
   (2) Prime topic: industry-induced innovation for HPC, science and engineering barriers

4) ADDITIONAL SURVEY and WORKSHOP FINDINGS, August 31, 2010 Manufacturing Supply Chain Summit, Chicago IL

a) FORTUNE50® companies

   i) SBE&S is needed because of the pace of production
   ii) Must use virtual learning to replace slow, physical, expensive experimentation
   iii) Physical prototypes are not accurate
iv) The OEM (Original Equipment Manufacturer) models components because the supplier doesn’t

v) Need predictive MS&A to speed up innovation

b) SMEs
   i) Rapidly replacing metallic products with composites, but they do not have the tools to validate the final design
   ii) They do not know, nor do they have access to, the material science inputs for validation
   iii) How does the company introduce economies of scale without such knowledge?
   iv) General use physics code does not work for modeling mold flows; injection molding simulation software fails to represent the realism of the chemistry and fluid-structure-interaction; mold flow is a dual-domain technology
   v) Need constitutive models; better programming is needed, but today’s research is inadequate to fully describe the fluid-structure-interaction
   vi) With composite materials, fiber alignment must be captured and materials properties understood; in other words, translation of orientation into properties has not been done

c) ISVs (Independent Software Vendors)
   i) High cost in employing domain experts for application support
   ii) Suffer from lack of MS&A investment in physics-based libraries and tools, especially materials

5) ADDITIONAL SURVEY and WORKSHOP FINDINGS – May 2011, NCSA Private Sector Program Annual Meeting, Urbana IL

a. Caterpillar’s Keven Hofstetter shares the challenge of having a good understanding of variability in both the product design and the environment and incorporating that into simulations to achieve high quality designs. There is great danger in using a single point solution to evaluate a design without regard to the variability associated with both the design and application. The “Variability and Uncertainty Quantification” graph below illustrates evaluation of a structural component. The designer may accept a particular design because the average component stress is well below the average material strength. However if variability in component loading and material properties is considered, it may be found that there is a high level of failure, hence an under-designed product. Or, one may make a design decision based on the minimum material strength being a multiple of the maximum component stress, which could lead to an over-designed product. Ultimately, a simulation must account for variability within both the material strength and expected component loading and then minimize the overlap to achieve a product that meets customer requirements without excessive cost.
b) Scientific discovery from advanced MS&A is driven earlier in the timeline to help avoid the increased cost to fix problems as the design moves toward production. (Product Development Timeline)

c) Rolls-Royce’s Shyam Neerarambam: “Simulation and modeling have essentially become a way of life. What are the challenges that simulation and modeling have to conform to if they deliver on the promise of virtual realism where we start relying more and more on virtual realism? Although this case study is specific to aircraft engine design, the challenges associated with
the design in industry are common and can translate across multiple industries. As we start introducing more realism, how does that fit into the design process?

d) Challenges with MS&A software:

i) Caterpillar’s John ‘Bo’ Salomon: “I would say that the actual size of our models is not the constraint. If I build one that 30, 40, 50 million cells, it’s hard for me to imagine we need to get too much bigger than that. But currently with just the physics I do right now which is multi-phase oil slingin’ around in an axle, a model like that might take a year to run and maybe just a few revolutions of the drive train. Our customers ask how long would it take us to run out for 2 or 3 minutes so things can heat up a little bit? The answer to that’s forever. So that’s the challenge for us is that, how can we increase the level of detail in our models that are needed and run ‘em out long enough that give us valuable insight.”

ii) ANSYS Barb Hutchings: “I’m from ANSYS and ANSYS is a commercial software company. We build and sell engineering simulation software used for structural analysis, CFD, electromagnetic and so on. How does ANSYS more effectively respond to our key customers when they ask us, what can ANSYS and high performance computing do to give me a strategic advantage? What’s the next level that you can take me to? I’m not sure the biggest bottleneck is actually on the software development side. We do a ton of high performance computing tuning with every release. And some of it is incremental, some of it’s kind of algorithmic - breakthrough rather than GPU, or like we’re doing now, hybrid parallelism to accommodate the many core architectures. The bigger challenge for us is demonstrating that in a way that’s commercially relevant.”
iii) Rolls-Royce’s Norm Egbert, lead North American engineer: We’ve moved out of a lot of the fundamental research and have that done in universities through long standing arrangements that provide for significant funding and basically make the universities an extension of our company. We have 40,000 employees; a little bit of a quarter are engineers, between 10-11,000. The vast majority of those engineers do not work in the virtual reality business. A little over 60% work in the aerospace industry, and most of the engineers are deployed in new product development and concept identification, technology insertion in existing products, support of those products, providing the services and a range of activities that support the aftermarket. Two percent of our engineering work force approximately, is involved in the methods and the total computational analytical support that underpins the virtual product modeling. It’s two percent because we actually are business case driven and there’s an opportunity cost.”

iv) The sequence of graphs below describes steps required to achieve virtual realism, as presented by Norm Egbert at May 2011 PSP annual meeting.

**Why is virtual realism difficult in our industry?**

_Because our reality is really difficult!_
We routinely use advanced simulations to model reality today …

- Aeromechanics
  - Stall
  - Flutter
  - Forced response
- Fluid Mechanics
  - Unsteady turbomachinery
  - Combustion
- Extreme events
  - Fan blade off
  - Bird strike
  - Ice ingestion

Current Challenges … why don’t we employ more advanced simulations today

- Computational solutions
  - Complexity of physics
  - Technology readiness levels of methods
  - Data reduction and display - intuitive insight
- Scale of the business
  - Six gas turbine engines certified in 2009 alone
  - Marine ships and propulsion product range
  - Nuclear power plant (totally different physics disciplines)
- Cost effectiveness – empirical can still be most effective
- Regulator and customer input – demands for physical demonstration
- Range of research advancements required for high fidelity modeling
Current Challenges … range of physics problems within different CFD solutions

Fluid Mechanics (CFD)

Unsteady Turbomachinery  Combustion  Cavitation  Nucleate Boiling

Mathematical models are different for each physical phenomena

Full Virtual Product Modeling … what is required?

- Reduce development and certification testing
  - ~100s of operating conditions
  - Satisfy all safety requirements
  - Demonstrate functionality, durability and reliability

- Multi-disciplinary solutions
  - Fluids
  - Mechanicals
  - Materials
  - Controls

- Estimated HPC Requirement per event/operational scenario
  - Double size of methods organization
  - Cores ~Million
  - Run time ~1 Year

3X
Virtual Product Modeling ... what is required from the academic community?

Collaboration ... to develop accurate/faster/cheaper simulation technology

Virtual Product Modeling ... what is required from entire community?

Collaboration ... to retain focus, provide purpose and deliver real value
v) Norm Egbert sums it up the value proposition as follows:

**In these challenging times, why should we pursue virtual realism?**

- Improved understanding
- Better & safer designs
- Reduce costs, risks, time

Accelerate industry and national competitiveness

We have the opportunity to fundamentally change engineering

- Real time simulation driving decision making
- New ways to design (simultaneous solutions)
- Move resource from current tasks to high value innovation
- Deliver new product functionality

vi) U of I’s Bill Gropp, computer science professor: “One of the more important things that is emerging is that there’s a realization that the modern architectures require much more attention to how one divides computation of the cores and the processors. Even a regular grid code put on a node that has no other user processes running actually is victimized by the other stuff that is going on, leading to lower performance. This is true on anything. It becomes increasingly easier to observe it on large parallel programs, because what happens is that imbalance accumulates until your parallel program coordinates or exchanges data, at which point you blame MPI for everything. But that’s not where the problem came from. And fixing it is going to require some thinking about how we arrange our codes.”

vii) Blue Ribbon Panel Bob Meisner, DOE: “There’s another half of this education part which is the programming side, the computer science side. It’s also the awareness the engineering side. So it’s not your area of expertise but it is the engineering curriculum owners. I’ve asked this question of several deans of mechanical engineering of top schools in the last month. I said how many of you have a mandatory finite element class using software in your undergraduate curriculum? And none of them had it. And these are top 25 engineering schools. Let me suggest that the problem as you’ve kind of outlined it so far is actually harder than you imagined. Because I think all the interesting science and engineering problems are interdisciplinary and I don’t think the universities have quite set up their departments to work in an interdisciplinary manner. We need three different disciplines to
work together with the computer scientists to actually work on a problem that is of interest to anybody. And I would say at the University of Illinois you had one of our line centers for ten years. You did rocket science there and you were pretty successful at bringing together teams and running things at large scale, working problems of interest to us, turbulence problems. But more importantly problems of interest to the community there. And I think they were pretty successful.”

viii) P&G’s Tom Lange, Director of Modeling and Simulation: -Ok, let’s move on to the physics computation side. We have petascale on the horizon. We have terascale already available to many of us even within our own firms. The embarrassingly parallel kind of problems, the use of these big computers to do design of experiments on relatively simple non scaling problems is already upon us. And the other thing is for the most part we get spatial scaling. As the geometries get more complicated the elements counts go up, the atom counts go up in the molecular dynamics we’re able to scale these problems and use ever more cores. But I have to say that in both structural and some of the fluid mechanics problems that are finite element in their calculations we are running into the barriers where I don’t think we’re scaling element problems beyond a few hundred cores. I know that in my experiences with the NNSA labs with the Sierra Suites they are clearly using larger counts there because that code was built for that. But for the most part this spatial scaling is sort of on track but explicit problems are not. So could you speak to the more general subject about what are the challenges you face in bringing the science and engineering codes to use the 10,000 to hundreds of thousands of cores? And how far away are we and what can we do to fix that? Are we investing in it or what’s goin’ on with it?

ix) Answer to Tom Lange from Bob Meisner, DOE: “There are methods that people have been using that are the right choice for solving problems of modest size on modest numbers of cores. They are currently using the right algorithms for that. But those algorithms have a limit at which point they stop scaling. So they’re not the wrong algorithms. They just the right ones for certain domains. The time domain remains a challenge. There are people who are working at that. There are some ideas. It will always be harder. There will be some barriers.”

6) ADDITIONAL SURVEY and WORKSHOP FINDINGS – March 2012, HPCC Conference, Newport RI

a) State of the art in virtual product development (VPD), Doug Post, DoD CREATE:
   i) Science-based computation is increasing in adoption
   ii) Physics-based testing is rigid and not responsive to new requirements; design flaws are discovered late, leading to rework
   iii) Systems integration happens late in the process (too late)
   iv) However, computation is highly scalable and responsive to new requirements; design flaws can be corrected early in the process
   v) VPD can replace rule-of-thumb extrapolations of existing designs with physics-based designs; physics can be injected into the design process early and throughout the design process
vi) Problem: commercial codes do NOT solve challenging problems, nor are they validated well enough at DOD for performance prediction

vii) Major technical challenge is software application tool development and deployment

viii) Major business challenge is getting management buy-in. They prefer the devil they know rather than the angel they don't. Most managers are risk averse.

ix) Question - what about solvers? Doug believes it can be done with libraries in a simpler way.

x) Key takeaway: A small delta in HPC capability can generate 3-digit ROI.

b) Addison Snell, Intersect 360:
   i) HPC architecture shifts put increasing burden of scaling on software developers.
   ii) Each core still is being used as a separate processor with fixed memory
   iii) This leads to a fundamental question as to which doesn't scale - the application or the model?
   iv) 95% of US manufacturers are companies with <100 employees

c) Tom Lange’s three steps, P&G:
   i) Pathology: explaining through models and simulations why things that already exist occur. (Bottle broke when fell of shelf)
   ii) Virtual Trial-and-Error: predicts through models and simulations what will happen with things that are describable but don't yet exist. (Would this new bottle design survive a fall off the shelf?)
   iii) Analysis-Led Discovery: forecasts using models and simulations where `desired or undesired' outcomes COULD exist ... (What material properties would have to be possible for bio-plastic to replace the full range of bottles currently in market?).
   iv) For the above 3... the organization tends to use 1) to build confidence in the models because they directly are validated, 2) as a next step once the M&S is ‘deemed validated’, and 3) as an extension of 1 and 2... ONCE you have sufficient analysis automation and computing capacity. It is usually well suited to ‘embarrassingly parallel' approaches.
   v) At P&G, “I have seen all 3 approaches used in different businesses at different levels of competitive and business stress. The more stressed the business the more likely you do 3, the more successful the competitive position the more likely you do 1 and 2.”

7) ADDITIONAL SURVEY and WORKSHOP FINDINGS – April 2012, IDC focus group, Richmond VA

a) Question 1: In your field of expertise, what are the main factors limiting the sustained performance and realism of applications codes today?
   i) Domain research in industry is often Edisonian. A cultural challenge exists to believe HPC is adequate to do real science.
   ii) Issues are with scale; most codes are commercial, off-the-shelf (COTS), showing negative scaling returns at 24-32 cores.
   iii) Companies face a poor payback on each software dollar spent with ISVs.
iv) Some physics domains have grown their computational requirements by 100X in five years.
v) New physics regimes are increasingly complex; just five years ago the computational modeling for these complex projects would have taken 400 years to do the modeling.
vi) Innovative user communities have ideas but don’t have access to sufficient computing power.
vi) Large memory computers are helpful, as it offers a platform to try out ideas at scale.
vi) Need more people to implement better science; specifically those at the intersection of science and HPC.
ix) Performance issues trump realism today. Models with adequate resolution need to run 10-100X faster.
x) In manufacturing, the most important key to simulation is that it can help make a better design decision.
xi) Lack of scientifically accurate physics seems to be a problem, but specifics are hard to get.
xii) Moving data is a real issue.
ixiii) There is an increasing demand for uncertainty quantification (UQ), so the data challenge will grow dramatically.
ixiv) Industry codes need to increase strong scaling. Lack of strong scaling means that engineers are doing more of the same stuff and not changing their sights.
ixv) Proper physics need to be packaged for the designer, not just the engineer, presenting a challenge to the HPC community to package its capabilities for non-experts.
ixvi) Finding experts in MS&A is a real issue; universities provide people with interest, but not with experience. Too few analysts sit between the scientists and the designers. Scientists often don’t seem to care about specific use cases; hence, the high demand for translators.
ixvii) Realism in biology is unresolved – we need better science.
ixviii) HPC systems are often built for science (both architecturally and politically), not UQ.
ixix) Consensus is that universities are not teaching computational awareness. PhDs in computer science don’t know how to map an application to an HPC architecture.
xx) Too many jobs are running on laptops. Attempts to run the same code on HPC results in the code thrashing back and forth between disk and memory at it runs for weeks. With sufficient memory it might run in minutes.
xxi) Lots of open source solutions are written in Python or Perl, which are not scalable.
xxii) Let’s get access to real computing power in high school.
xxiii) Time-to-solution is important in life sciences, so capacity computing provides critical capability. Must run ensembles to generate hypotheses; hence, the iterative process is real science, but the human/computer interface needs to be better. The “on-ramp” to build better models is simply too long.

b) Question 2: Referring again to your field of expertise, how important is it to improve the sustained performance and realism of key applications codes? What difference could improvements make?
i) Time-to-product is more important than time-to-solution.
ii) High need to visualize time series data.
iii) The science in some structural and fluid physics is adequate, but to achieve realism 10X compute power is needed in the same time frame. Output is lacking detail, so digital simulations must still be validated with physical testing.

iv) Lots of computational science is similar.

v) Repeated references to poor programming, not bad algorithms. But how does one find the programmers?

vi) Published benchmarks often cheat by making the problem 10X larger, not 10X more resolve.

vii) Bottlenecks: material science and imaging. Some images in life sciences reach 70k X 70k. Visualization is a communication tool.

viii) Biology is really a digital science now. High need to better understand the atomistic level to get the chemistry to the device.

ix) Would benefit from cross-discipline sharing, including cross-industry. For instance, sound waves from seismic modeling can inform sonograms.

x) NSF should fund more inter-disciplinary research.

xi) Make scientists simulate their algorithms. Real-world examples need engineers.

xii) U.S. scientists have great theory, but European dissertations often are more applied.

xiii) As a society we are at a good place to move science into engineering, but university graduates are not equipped.

xiv) The most important aspect of MS&A is the ability to optimize.

xv) Fuel efficiencies can only be achieved with HPC.

xvi) Federal certification agencies need better understanding of MS&A so as to reduce the costs of physical testing.

c) Question 3: How much do you think the sustained performance/realism of your key codes could be improved with sufficient effort?

i) Engineers run minimalistic models because they want to run on their laptop, so this drives gross errors. Removing barriers to scale will improve model fidelity, but these are cultural issues.

ii) One key element missing in government codes is ease of use. Figuring out how to use an application is a show-stopper

iii) 10X speedup in multithreading is needed. Video gamers and the financial community don’t need all this. Chip design and software is targeted at the mass market, which is not suitable for HPC. Really need three orders of magnitude improvement in key codes.

iv) Underinvestment abounds in making HPC specialists and computer scientists more productive. Must beat Moore’s Law or lose to the competition.

v) Performance capability has been nearly flat for 15 years because jobs are more complex.

vi) Need scientists and engineers in close proximity. The era of commodity computing made programmers lazy.

vii) Do graduates today know how to optimize algorithms?

d) If you were in charge of NCSA, what would you do?
i) Get the institutions to train and educate properly. Deliver education to working folks in the style of executive MBA degrees.

ii) Have smarter libraries and tools to make the users more productive.

iii) It’s easier to train an engineer in computing than to train a computer scientist to be an engineer.

iv) A fraction of non-US students in math and computer science are well qualified; however, they are low on experience.

v) NCSA can be a helpful aggregator.

vi) If computation is the third leg of science, then why doesn’t NSF boost the requirements?

vii) Change the thesis requirement.

viii) Workforce training needs an internship at an HPC center, followed by a practicum inside the intellectual property tent of a company.

ix) The old business model had 40-50% margins, which helped fund COTS code development. This broke. Fix it.

x) Strong scaling needs algorithms. Navier Stokes equations are 200 years old.

8) ADDITIONAL SURVEY and WORKSHOP FINDINGS – May 2012, NCSA Private Sector Program Annual Meeting, Urbana IL

a) Why is Multiphysics so Difficult?
   i) Need domain tools and various groups to work together
   ii) Post processing may be biggest challenge
   iii) Abaqus code base that’s 30 years old presents a legacy problem
   iv) Multi-vendor challenges now exists
   v) Academics often build single-discipline algorithm, making progress more difficult than it should be.
   vi) Let’s attack real physics, not just Navier-Stokes.
   vii) Melting and solidification encounter five different forms of physical phenomena, making it difficult for load distribution of the domain; 4 of these 5 prefer finite volume; solids prefer finite element methods.
   viii) Time-constrained development environments are key for industry; e.g., full body shake of a Caterpillar tractor the size of this room can only be determined by simulation, as no test machine will be built.

b) Is the Solver the Holy Grail?
   i) UI’s Bill Gropp - research interest is in the solver, but it's not enough to do the math and science right; load imbalance really kills codes at scale.
   ii) Slower individual steps may be better if they aid a faster workflow.
   iii) There is no black box for fluids (not even close for incompressible fluids). Tough to put these solvers in ISV codes.

9) ADDITIONAL SURVEY and WORKSHOP FINDINGS – June 2012, International Industrial Super-computing Workshop at NCSA, Urbana IL
a) Illinois Rocstar code (DOE funded at UIUC) can do far better simulations than there is data. How do they now improve the designs of experiments?
b) Key complexities: applications, algorithms, hardware, software
c) Most solvers have mathematical abstractions, not computational expressions.
d) Misunderstanding of mechanical computational codes result in conservative approaches.
e) HLRS (Stuttgart) receives $50M from EU for staff to help manufacturing SMEs.
f) For K at Japan’s Riken, they will write their own structures code. No plan to port ISV codes. To achieve 10% of peak, they must keep their code in cache and break the memory barrier.
g) Cray Supercomputing:
i) CAE (computer-aided engineering) drivers: extreme fidelity, design optimization, robust design
ii) ISVs buy into the need for scalability; ANSYS calls this mega simulation
iii) Difference from academic science is the demand for design optimization
iv) Early 1980s - vectorization was driven by Nastran code for automotive crash simulations
v) Parallel crash simulations with MPI came in 2003
vi) Five-year life cycle suggests improved scalability for CAE in 2013
vii) HPC centers play a vital role for proof of CAE scalability on varieties of platforms
viii) Is the business case for these proofs limited to companies that have large economic benefits?
ix) Fluent scales to 3000 cores on Cray in ESM mode and to 3000 cores on SGI. Cray’s CCM mode is comparable up to 1000 cores. LS-Dyna 2-car crash simulation scales to 1000 cores.
x) Implicit Abaqus still at about 64 cores; unsteady solution on 384 cores took 350 hours, generating terabytes of data, 1350 time steps, moving mesh, rotating boundary condition

7) PEOPLE WORKING ON THE PROJECT

a) PI: Merle Giles, NCSA’s Associate Director responsible for Economic and Private Sector partnerships and development
b) NCSA staff:
i) Dr. Seid Koric, Senior Technical Coordinator and Lead, Finite Element Analysis and Structural Mechanics for Private Sector Program and XSEDE;
ii) Dr. Ahmed Taha, Lead, Computational Fluid Dynamics for Private Sector Program and XSEDE
iii) Evan Burness, current PSP project manager, 2010 student project team manager
c) MS-Tech Management program project team at U of I:
i) Evan Burness (project manager), Durham NC,
ii) Majed Alotaibi, Riyadh, Saudi Arabia
iii) Jose Arbelaez, Bogota, Colombia
iv) Jae Joon Hyun, South Korea
v) Iru Jun, Seoul, South Korea
vi) Jin Hyuk Mun, South Korea
vii) Jordan Schafer, Ed.M. degree candidate
d) Faculty consultants at U of I:
i) Professor Robert Haber  
ii) Professor Philippe Geubelle  
iii) Professor Pratap Vanka  
iv) Professor Arif Masud  
e) IDC staff  
i) Earl Joseph, PhD  
ii) Steve Conway  
iii) Chirag Dekate, PhD

8) OPPORTUNITIES FOR TRAINING, DEVELOPMENT, AND MENTORING

a) Findings from this SBE&S study have been shared with attendees at NCSA’s PSP annual meeting and at a September 2012 IDC HPC User Forum meeting in Dearborn MI.

b) Expertise and understanding gained from this study are currently being shared with the industrial partners and small and medium supply-chain manufacturers involved in NDEMC (National Digital Engineering and Manufacturing Consortium), a public-private partnership that includes the following agencies and major companies:
   i) U.S. Department of Commerce’s Economic Development Administration  
   ii) NSF, NIST, NASA, White House OSTP  
   iii) GE, John Deere, Lockheed Martin, Procter & Gamble

c) NCSA intends to publish a document summarizing the results of these NSF-funded results.

d) Industry participants, specifically NCSA PSP partners, will be involved in an upcoming NSF SI² workshop in May 2013, providing academic/industry scientist collaboration and insight into demand for industrial petascale applications.

Appendix A: IDC SPECIAL REPORT - NSF/NCSA Special Investigation of HPC Applications Used in Industry: Their Usage and Needs and the State of the Art in the Science Underlying the Algorithms, August 2012

Appendix B: MS-Technology Management Report on FACTORS GOVERNING THE USE OF COMPUTATIONAL & SIMULATION MODELING WITHIN THE AMERICAN MANUFACTURING INDUSTRY, July 2010


IDC NSF/NCSA Special Investigation of HPC Applications Used in Industry: Their Usage and Needs and the State of the Art in the Science Underlying the Algorithms, August 2012, Tables 8-14

Ibid. Table 22

Ibid. Table 20

Ibid. Table 21

Ibid. Tables 26-31

MS-Technology Management Report on Factors Governing the Use of Computational and Simulation Modeling Within the American Manufacturing Industry, July 2010, page 19, question #7

Hofstetter, Keven, Caterpillar Inc. manager of virtual product development, (see section 5.a.)

Section 7.a. focus group findings

Simulia presentation at ISC'09, Hamburg, Germany

Giles, Merle, August 31, 2010 personal notes from NDEMC organizational meeting and Supply Chain Summit, Chicago, 58 attendees including Argonne Lab, NCSA, Purdue, NSF, OSTP, GE, IBM, Boeing, Lockheed Martin, Rockwell, Caterpillar, Deere, P&G, ANSYS, CD-adapco, Accelrys, Llamasoft, LSTC, Polymer Ohio, Illinois Manufacturing Extension Partnership, Illinois Valley Plastics, Premix