Final Technical Report for Contract: DE-SC0008540

The following report is the final technical report for contract DE-SC0008540 described as follows:

Contract Title	Economic Model For a Return on Investment Analysis of United States Government High Performance Computing (HPC) Research and Development (R & D) Investment
Grantor	The U.S. Department of Energy's Office of Science
Funds Obligated	$410,000
Grantee	IDC Research Inc.
Grant Number	DE-SC0008540
Description	Developing An Assessment of How Investments In HPC and R&D Can Impact A Nation's Competitiveness and Level of Innovation

Reporting To:

Sponsoring Office

DoE Program Manager:
U.S. Department of Energy, Office of Science
Germantown Building, 1000 Independence Avenue, S.W., Washington DC 20585-1290

Administrator

Contracting Officer:
U.S. Department of Energy, Office of Acquisition and Assistance
9800 South Cass Avenue, Argonne, IL 60439

Reporting By:

Sponsoring Office

Principal Investigator:
Earl C. Joseph, PhD, VP, HPC Res. (612) 812-5798 ejoseph@idc.com,
Senior Technical Project Manager:
John Daly, DE-SC0008540 508 935-4643 jdaly@idc.com

IDC
Analyze the Future

SPECIAL STUDY

Creating Economic Models Showing the Relationship Between Investments in HPC and the Resulting Financial ROI and Innovation — and How It Can Impact a Nation's Competitiveness and Innovation

Earl C. Joseph, Ph.D.
Chirag Dekate, Ph.D.
Steve Conway

IDC OPINION

The authors want to thank the U.S. Department of Energy (DOE) for its insights and guidance on and funding of this research project. This study is based upon work funded by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, under award number DE-SC0008540.

This pilot study investigates how high-performance computing (HPC) investments can improve economic success and increase scientific innovation. This research is focused on the common good and should be useful to DOE, other government agencies, industry, and academia. The study has created two unique economic models and an innovation index:

- A *macroeconomic model* that depicts the way HPC investments result in economic advancements in the form of ROI in revenue (GDP), profits (and cost savings), and jobs.
- A *macroeconomic model* that depicts the way HPC investments result in basic and applied innovations, looking at variations by sector, industry, country, and organization size.
- A new *innovation index* that provides a means of measuring and comparing innovation levels.

Key findings of the pilot study include:

- IDC is able to collect the required data across a broad set of organizations, with enough detail to create the two economic models and the innovation index.
- Early results indicate very substantial returns for investments in HPC:
  - $356.5 on average in revenue per dollar of HPC invested
  - $38.7 on average of profits (or cost savings) per dollar of HPC invested
- The average number of years before returns started was 1.9 years.
- The average HPC investment per innovation was $3.1 million.

Note that an additional outcome of this research is an expansive list of HPC success stories (as shown in Appendix B: Full Descriptions of the ROI and Innovation Examples).

Filing Information: September 2013, IDC #243296, Volume: 1
Special Reports: Special Study

Global Headquarters: 5 Speen Street Framingham, MA 01701 USA P.508.872.8200 F.508.935.4015 www.idc.com

TABLE OF CONTENTS

P

TABLE OF CONTENTS — Continued

LIST OF TABLES

P

LIST OF FIGURES

P

IN THIS STUDY

Grant Contract References

Grant References

The authors want to thank DOE for its insights and guidance on and funding of this research project.

This study is based upon work funded by the U.S. Department of Energy Office of Science, Office of Advanced Scientific Computing Research, under award number DE-SC0008540.

The grant description is as follows: Developing an assessment of how investments in HPC and R&D can impact a nation's competitiveness and level of innovation.

Sponsoring Office

DOE Program Manager

Christine Chalk, 301-903-5152, christine.chalk@science.doe.gov, and Barbara J. Helland, 301-903-3127, barbara.helland@science.doe.gov, U.S. Department of Energy Office of Science, Germantown Building, 1000 Independence Avenue, S.W. Washington, D.C., 20585-1290.

Administrator/Contracting Officer

Warren Riley, 630-252-2485, warren.riley@ch.doe.gov, U.S. Department of Energy Office of Acquisition and Assistance, 9800 South Cass Avenue, Argonne, Illinois, 60439.

IDC Reporting

- **Principal investigator:** Earl C. Joseph, Ph.D., 612-812-5798, ejoseph@idc.com
- **Senior technical project manager:** John Daly, 508-935-4643, jdaly@idc.com

Why This Research Is Important to DOE

World scientific leadership and innovation leadership are becoming more dependent on the use of HPC/supercomputers every year. Economic leadership increasingly directly results from a nation's or an industry's or an enterprise's application of supercomputers in innovative and productive ways. Many countries/regions (such as China, Russia, Europe, and other Asian countries) are putting into place plans to gain leadership in innovation and economic progress by more broadly applying HPC/supercomputing across many different industries and segments. This research is focused on the common good and should be useful to DOE, other government agencies, industry, and academia.

EXECUTIVE SUMMARY

Project Overview

This pilot study investigates how HPC investments can improve economic success and increase scientific innovation. This research is focused on the common good and should be useful to DOE, other government agencies, industry, and academia. 208 ROI and innovation examples were collected and analyzed as part of the pilot study. This was made up of 67 financial ROI examples and 141 innovation examples. The sample had 67 basic research innovations and 74 applied research innovations. 79 of the innovations were in the public sector, and 62 were in the private sector.

The study was made up of 67 financial ROI examples and 141 innovation examples.

IDC created and tested four different versions of the economic models and four different innovation index versions for this study. The main reason for changing the model was to provide the ability to collect data from a broad set of organizations. For example, some organizations had difficulty in responding when asked for a specific number for what they spent on HPC or what they spent on the overall R&D program, so IDC created a number of ranges for the organizations to select from (instead of requiring them to provide a specific number).

The pilot study created and tested two economic models and a new innovation index:

- How HPC investments result in financial ROI — in terms of revenue/GDP, profits/cost savings, and job creation
- How HPC investments generate innovations — for basic and applied innovations in both the public and the private sectors
- The new innovation index that uses a scale from 1 to 10, with a mix of the importance of the innovation, how many organizations use the results, and how understandable the innovation is to the broader public

Two Examples of How HPC Provides ROI

The sections that follow provide first-hand perspectives on the value of HPC from two very different adopters, who graciously agreed to contribute their perspectives:

- The first perspective is from the National Oceanic and Atmospheric Administration (NOAA), a government organization that for many years has relied on HPC to produce its National Weather Service forecasts and other leading-edge work.
- The second perspective comes from PayPal, a successful global ecommerce company that adopted HPC not long ago for real-time detection of online fraud. IDC estimates that HPC adoption has saved PayPal about $700 million to date.

The National Oceanic and Atmospheric Administration: HPC as a Mission-Critical Tool

The following first-person perspective comes from David Michaud, deputy director, High Performance Computing and Communications (HPCC), NOAA.

The National Oceanic and Atmospheric Administration's mission is to understand and predict changes in the Earth's environment, from the depths of the ocean to the surface of the sun, and to conserve and manage coastal and marine resources. From daily weather forecasts, severe storm warnings, and climate monitoring to fisheries management, coastal restoration, and supporting marine commerce, NOAA's products and services support economic vitality and affect more than one-third of America's gross domestic product. HPC has long been indispensable for carrying out NOAA's mission.

At NOAA, high-performance computing is well integrated with the leadership and management of the typical enterprise IT services. Our CIO has two titles, CIO and director of High Performance Computing and Communications. So these subtleties are dealt with directly under a consistent perspective.

HPC is a critical tool needed for NOAA to carry out its mission on a daily basis. The extent of NOAA's HPC capability and capacity is directly linked to our mission performance. To ensure our HPC is acquired, implemented, and allocated for the maximal benefit to our organization, we have developed a well-established governance structure. The CIO and director of HPC chairs our High Performance Computing Board, which includes senior leaders from all of NOAA's major missions, such as weather services, satellite services, ocean services, fisheries services, and our applied research groups. These representatives who participate in overseeing HPC are generally non-technologists who are responsible for setting programmatic priorities within their mission areas. Beneath this group are the technologists responsible for acquiring HPCC systems and operating them on a daily basis, along with program managers in charge of allocating and prioritizing the HPC resources.

From the CIO's perspective, the highest priority is IT security, but HPC is a close second. Without HPC, we can't perform our mission work. NOAA's senior leaders recognize that HPC is a key component of our value chain for producing NOAA's weather reports and other products. Weather observations are very important, and HPC is important for turning all that data into meaningful information. We invest over a billion dollars each year in observations. HPC is crucial for leveraging those investments. The final crucial element is dissemination, getting the data out to the public.

At its core, HPC needs to be about enabling an organization's mission, not just putting together a computing program for technology's sake. There are important benefits, even for organizations that are just starting out in HPC. You can use a simple cluster to aggregate disparate applications within an organization and increase the performance of those applications. Within an HPC system, you can also aggregate all the data sources in one place and extract value from them in one place. These activities often reveal or create previously unknown dependencies between applications, which tighten relationships among groups within an organization. All of these benefits yield cost efficiencies. And while you shouldn't expect all the benefits to be realized immediately, even in the first year, you should expect to see important benefits with some targeted applications.

Thanks to HPC's strong connection to our CIO, at NOAA, we are able to provide HPC from an enterprise perspective. We started this integration of HPC into enterprise IT in 2006. Now, NOAA is looking to collaborate across federal agencies by providing HPC resources in a shared services paradigm. We're really good at knowing how to exploit midsize HPC systems cost effectively in a highly reliable manner, and we can help other organizations that have similar needs. One part of this challenge is working with a mix of experienced and less experienced users. In turn, we leverage other federal agencies' facilities such as the Department of Energy and National Science Foundation to understand how to exploit their larger, leadership-class HPC systems.

PayPal Exploits HPC for Fraud Detection

The following first-person perspective is from James Barrese, CTO, PayPal:

PayPal is a wholly owned subsidiary of eBay Inc., a company that in 2012 earned $14 billion in revenue and enabled more than $175 billion in commerce. Among other things, PayPal is responsible for detecting fraud across eBay and StubHub. By deploying HPC servers, storage, and software, PayPal is able to detect fraud in near real time before it hits consumers' credit cards. HPC has helped PayPal to catch substantial additional fraud.

PayPal's decision to use HPC goes back to our innovative technology roots. More than $5,200 is transacted every second across our platform, and it's our responsibility to not only deliver great customer experiences but keep our customers' financial information safe. As our company continued to rapidly grow, we saw both a challenge and an opportunity with handling fraud and risk while delivering more personalized shopping experiences. Our approach to these types of situations is to be very entrepreneurial and aggressive. We asked, "How can we solve this business problem? and our HPC team said, "Hey, this problem's been solved before, using HPC." This led to an HPC pilot and then to deployment.

I think we're in the most interesting time in my career — the pace of technology change is accelerating. There's a whole new S-curve with being able to leverage HPC and machine learning on a real-time basis. HPC helps us manage large, globally distributed systems. With HPC, we're able to do risk analysis in real time and detect fraud before it happens. We're also able to leverage HPC to further personalize the consumer experience. Say, you walk by a Jamba Juice and we know you like smoothies, so we deliver a coupon to you right then and there.

When you are evaluating whether or not HPC is right for your company, here are some tips:

- You need the right talent to be able to deliver results. I would urge people to start with clear objectives and have the right people on board. We had the nucleus of HPC talent and have been recruiting more HPC talent aggressively.

- Clearly understand that HPC is not a mass consumption technology where we enable everyone in our organization with it. This is a deep engineering function. It's custom built and includes writing software to solve cutting-edge problems.

- Think of HPC not as an IT function but as a competitive business advantage. There's a hard link between HPC and PayPal's top line and bottom line.

HPC is key to our company's strategy because PayPal is redefining an industry, and it's all driven by technology. HPC is enabling new business capabilities that are really product driven.

Key Research Findings

- IDC is able to collect the required data across a broad set of organizations, with enough detail to create the two economic models and the innovation index.

- Early results indicate very substantial returns for investments in HPC:

  - $356.5 on average in revenue per dollar of HPC invested

  - $38.7 on average of profits (or cost savings) per dollar of HPC invested

$356 dollars on average in revenue were generated per dollar of HPC invested.

- The average number of years before returns started was 1.9 years.

- The overall average HPC investment cost for jobs created was over $93,000.

  - 42 sites reported job creation. On average, 29.8 jobs were created from HPC investments at these sites.

  - A total of 1,251 jobs were created from the 42 financial ROI examples.

- The average HPC investment per innovation was $3.1 million.

  - Overall, approximately half a billion dollars in HPC investments were made to generate 141 of the innovations in the pilot study.

  - On average, the HPC investment was $3.1 million per innovation, with many under $1 million per innovation.

- Academic and industrial sites represented the bulk of the innovations in the sample. Most of the basic research innovations were in academia, while most of the applied research innovations were in industry.

- Government innovations averaged higher (7.0) on the innovation index scale, while innovations in industry averaged 5.7 and academic innovations averaged 3.9. The scale ranked the most important innovations as 10 and the least important as 1.

- The total HPC investments made for innovations in the sample population was $496.7 million, and the average HPC investment was $3.1 million per organization.

- Government sites tend to invest more per HPC innovation ($4.4 million), while academic sites invest less ($2.5 million) and industrial sites invest $3.8 million on average.

Note that an additional outcome of this research is an expansive list of HPC success stories (as shown in Appendix B: Full Descriptions of the ROI and Innovation Examples). These can be used to help explain the importance of HPC to funding bodies, key decision makers, and the broader public.

Cautions on Using the Pilot Study Data

While the data set is very rich and broad, it still isn't large enough to make direct correlations, especially when sorted into subsets. Additional research will be required to turn the correlations into causal relationships and to map the relationships more confidently by country and by industry.

One must be very careful in using the data for making key decisions. The pilot study data is large enough to show the major trends but isn't large enough to support detailed economic comparisons.

Note that there were two sites with considerably higher returns that were removed from the data set as outliers:

- $10,057 in revenue per dollar of HPC invested at a finance company:
    - If included, this raises the overall revenue ROI from $356.5 to $536.2.
- $1,350 in profits per HPC dollar invested at an oil and gas company:
    - If included, this raises the overall profit ROI from $38.7 to $65.5.

Proposed Future Next Steps

IDC recommends that this pilot study be used as the foundation for creating actual full-scale economic models for financial ROI and innovation from investments in HPC.

This new research could follow a three-year, three-phase approach:

- **Phase I (Year 1):** The goal is to create the actual ROI models with a full data set at least two times in Phase I, growing to at least eight times in size by Phase III. This is needed to create predictive models for a fuller understanding of the relationships, to provide enough data/analysis to start making predictive results, and to refine the models as needed.
    - Motivate a select set of nations to fund IDC to collect deeper data samples within their countries.
    - The goal is to have enough data to start making statistically sound correlations between industries, countries, and different sizes of organizations.

- **Phase II (Year 2):** Expand the data set by at least two times more, including more countries and industries that are harder to collect data from.

  - Motivate a larger set of nations for deeper data samples within their countries.

  - The goal is to have enough data to make strong, statistically sound correlations between industries, between countries, and between different sizes of organizations — and cross-correlations like industries by country and organization size by country.

  - At this phase, it should be possible to start the build out of the tie to larger general economic data sets.

- **Phase III (Year 3):** Expand the data set again by at least two times more and focus on hardening the predictive nature of the models by conducting additional research to "test" correlations. Refine the models as needed. It should also allow the fuller testing of trends and the impacts of changes made during the research period.

  - The goal is to conduct enough research to show both strong statistical correlation and causation between investments in HPC and the resulting ROI and innovation generation.

  - In this phase, the scenario testing should improve to be very robust and directly useful for making national and regional policy decisions.

  - At this phase, the ties to large general economic data sets should be robust enough to be able to hand off future tracking to one or more appropriate government organizations.

RESEARCH METHODOLOGY

Project Overview

This pilot study for the U.S. government describes and quantifies how increases in HPC investments by the federal government, industry, and academia can improve the nation's economic success and increase the nation's overall scientific innovation capability. The study employs a unique and innovative research approach that should provide strong insight to help improve DOE's and the nation's scientific and economic capabilities. Because the study is global in scope (although focusing most heavily on the United States), officials of other nations may also find it to be of interest. The research is focused on the common good and should be useful to DOE, other government agencies, industry, and academia.

This pilot study describes and quantifies how increases in HPC investments by the federal government, industry, and academia can improve a nation's success.

This pilot study used an approach based on traditional business market research methodologies. This approach greatly helped in the required collection of new data points and provided a faster path to creating the models and results. This is different from traditional economic modeling approaches, which typically use larger data sets and require far more time to conduct the analysis and create the models. The success of the current study is heavily dependent on the collection of in-depth industrial data around R&D/HPC investments and the resulting business returns in revenue, profits, and job creation.

The pilot study included the development of the economic models and collecting an initial data population to test the models. This study quantitatively showed the major economic impacts of HPC investments and allowed initial comparisons between the United States and a number of other innovative nations. These comparisons are intended not to rank nations but to show the differences in their investment strategies for HPC and R&D. The pilot study also reveals that these differences in strategy can be significant.

In future research phases, we plan to analyze and evaluate three different potential U.S. federal HPC funding levels and project their potential impacts on U.S. innovation and economic progress over a 10-year time frame. We also plan to compare how industry currently evaluates the financial returns from R&D investments and propose a new approach that would better reflect the real-world dynamics of competitiveness and innovation based on the results of the study.

Research Background Assumptions and Approach

- This research is based on a number of underlying assumptions (e.g., that industrial R&D is key to national economic success in important industries that exploit HPC).
  - The use of technical computers and in particular large technical computers can drive industrial R&D to be more successful in certain industries.
- This study only looks at organizations that conduct R&D and use computers as part of their R&D process such that advancing the organizations' use of computers for R&D can accelerate national economic growth. In other words, there are additional economic impacts from investments in HPC beyond the deliberately restricted scope of this study.
- This study focuses on collecting new base data that is specific to the topics being investigated.
  - Unlike previous studies that either use broad published data or use individual case studies, this research collects quantitative data from a broad set of industries/sectors for the creation of the two models.
- Innovation for this study is defined by what experts within each sector judge as valuable advances.
  - Organizations surveyed were asked to identify their major innovations and their counterparts' and competitors' top innovations.
  - This approach goes far beyond counting only patents, papers, and other indirect measurements of the economic impact of R&D investments.

This study focuses on collecting new base data that is specific to the topics being investigated.

Project Goals (as Stated in IDC's Original Proposal to DOE)

Stated goals: The project will develop two models and a study that shows quantitatively how the federal government — by increasing investment in HPC — can improve the nation's economic success and increase national innovation. The study will document contributions of supercomputing to scientific advancement and economic growth — showing why it is important to maintain and grow leadership in this critical area.

The project will develop two models and a study that shows quantitatively how the federal government — by increasing investment in HPC — can improve the nation's economic success and increase national innovation.

The study will include two research-driven macroeconomic models that depict the way HPC investments at a federal level result in national economic advancements and innovation levels. The models will increase the understanding of the role of HPC in national economic viability and innovation and will provide information that will partially test the following key interrelated hypotheses:

- An increase in federal investment in HPC will increase the competitiveness, productivity, and innovativeness of U.S. scientists, engineers, researchers, and companies by 2020.
- This increase in science and engineering effectiveness will in turn lead to greater scientific advancement and economic growth and more higher-paying jobs.
- Increased investment in HPC at the federal level will in turn increase the productivity of U.S. scientists and engineers by enlarging and deepening the use of HPC computing throughout U.S. industry, academia, and government agencies.
- Changes in other countries' relative investments in HPC directly impact the United States' long-term economic success.

Original Research Approach and Phases

- **Step 1 — collect relevant data from a broad set of resources, including:**
  - Case study examples across a broad set of industries, agencies, and educational institutions that depict historical trends and will inform the development of both the economic model and the innovation index. (The focus of this data collection is to obtain consistent quantitative data that can be used to create the economic impact ROI model.)
  - Meaningful data from IDC's existing internal data on HPC usage by industrial organizations around the world (This includes data mining of IDC data structures as well as external data sources. It is expected that the majority of the model data will come from the collected primary research data, as it has to be tied directly to individual products/services.)
- **Step 2 — develop macroeconomic models that demonstrate the relationships between HPC investments and overall national economic growth.** The models will incorporate a wide-angle macroeconomic view of the ROI chain of interrelationships, clearly showing the value of HPC to economic growth. The ROI model will use metrics based on the new collected primary data and is expected to include parameters such as HPC investments, R&D, revenue

resulting from these products/services, profits resulting from these products/services, time delays from R&D to dollar returns, amounts of HPC invested on these products/services, and changes in HPC investments as well as how these impacted revenue and profit returns. Data for the model will include specific ROI data on individual products and services that were developed by each company to properly correlate the data with the specific investments in HPC. This will also provide supplemental information on the typical ratio of HPC-to-R&D investments by different industries.

- **Step 3 — create an innovation index that identifies key metrics of national scientific innovativeness and then compare at least three different countries with the United States.** The team will collect multiple measures of innovation to develop the most useful innovation index (as indentified by experts within each domain/sector). This will be based on new data collected by the surveys conducted for the study. The index will likely vary by sector and by industry.

- **Step 4 — analyze and evaluate three different potential federal HPC funding levels and project their potential impacts on U.S. innovation and economy viability.** IDC will create three potential levels of changes in federal HPC funding and will then apply the models to create an estimated impact of each level. The ROI impact to the economy will be based on looking at the number of companies within the overall U.S. economy that use R&D and also use computers to support their R&D activities. (Note: This will be a key portion of the future study phases).

- **Step 5 — IDC will have a team of external experts validate the research results and models.** The experts can be identified by DOE or others. (Note: This was only partly completed in the pilot but is planned to be more robust in future research).

- **Step 6 — communicate and disseminate the data, analysis, and ideas across the broader U.S. HPC community including all stakeholders, investors, congressional staff, users, and vendors.** This includes showing the ROI examples, the benefits of HPC, and the potential impacts of expanded HPC investments.

Note: DOE has ownership rights for use of the models and data inputs/outputs. IDC also retains ownership rights to the models and data for its use in future studies and/or projects.

Changes Made to the Research During the Pilot Study

IDC created and tested four different versions of the economic models and four different innovation index versions for this study. The main reason for changing the models was to provide the ability to collect data from a broad set of organizations. For example, some organizations had difficulty in responding when asked for a specific number for what they spent on HPC or what they spent on the overall R&D program, so IDC created a number of ranges for the organizations to select from (instead of requiring them to provide a specific number).

IDC created and tested four different versions of the economic models and four different innovation index versions for this study.

The pilot study created and tested two economic models and a new innovation index:

- How HPC investments result in financial ROI — in terms of revenue/GDP, profits/cost savings, and job creation
- How HPC investments generate innovations — for basic and applied innovations in both the public and the private sectors.
- The new innovation index that uses a scale from 1 to 10, with a mix of the importance of the innovation, how many organizations use the results, and how understandable the innovation is to the broader public

Notes:

- The pilot study succeeded in developing an approach to collect highly private data in a form that organizations are willing to share. It required more effort than originally expected and frequently required revisiting sites to obtain more accurate data.
- Representatives of some countries wanted to first see the results and how the data is used before deciding whether to participate on a larger scale (including Germany, Japan, and some sites in France). After IDC reviewed the pilot study results with them, most of these officials are more enthusiastic about participating in future phases.

Data Retention Plans

IDC is putting into place a multifaceted approach to ensure full data retention:

- IDC will maintain copies of the data for no less than five years.
- IDC will provide DOE with 100% of the base data for DOE use and retention.
- IDC will distribute the full data set and report to approximately 4,500 members of the broader HPC community around the world.

The Pilot Study Sample Demographics

Sample demographics:

- A total of 208 case study examples of ROI and innovations were collected as part of the pilot study:
  - 67 financial ROI examples
  - 141 innovation examples
- In addition, a very large number of microsurveys were conducted to obtain key ratios needed to eventually apply the results to large economic data sets:

A total of 208 case study examples of ROI and innovations were collected as part of the study

- Over 30,000 scientists and engineers were contacted, with over 1,500 completing the microsurveys.

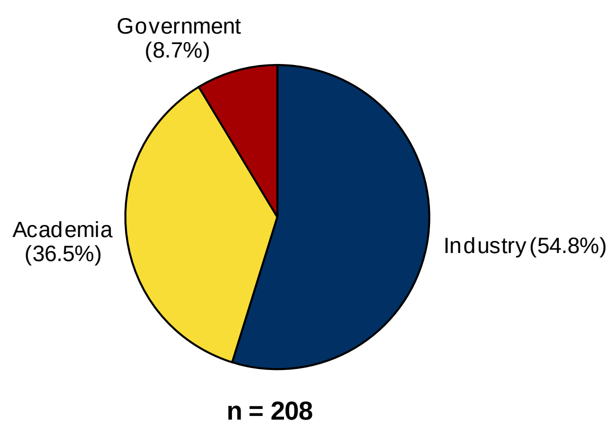
Figure 1 shows the overall mix for the 208 examples in the study. Figure 2 shows the mix for the 67 financial ROI examples in the study, and Figure 3 shows the mix for the 141 innovation examples.

Most of the financial ROI examples are from industry as shown in Figure 2.

Academia represents a major portion of the reported innovations, followed by industry, as shown in Figure 3.

FIGURE 1

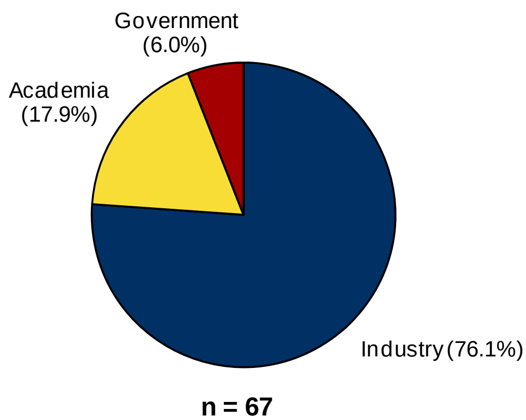
Overall Sample Mix



Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

FIGURE 2

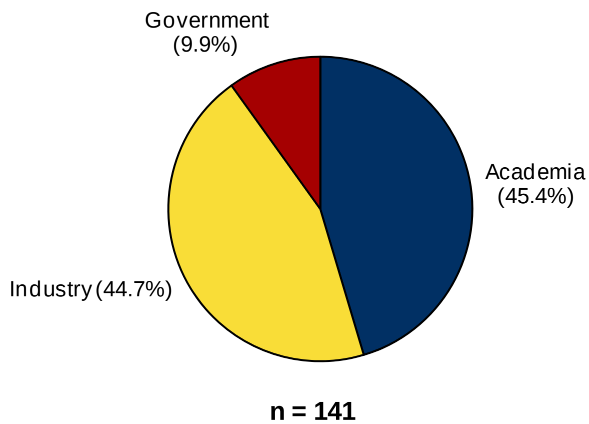
Overall Sample Mix: Financial ROI Examples



Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

FIGURE 3

Overall Sample Mix: Innovation ROI Examples



Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

Survey Methodology to Create Ties to Broader Economic Reports

This pilot study started the process of collecting broader data on organizations within major economic sectors that conduct R&D and that use HPC in their R&D operations. Attempts were made to contact over 30,000 scientists and engineers in order to start creating the required profiles by industry sector. IDC projects that 5–10 times more surveys will be required to create a data set that is useful for economic analysis.

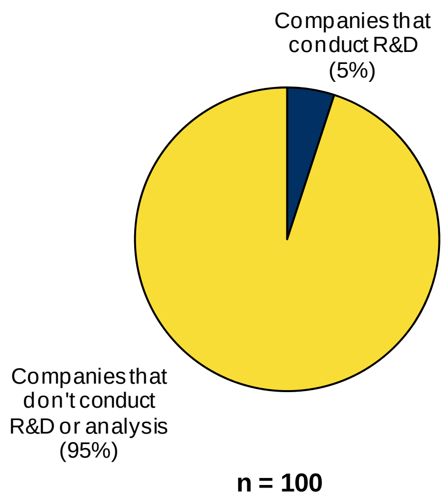
Over 30,000 scientists and engineers were attempted to be contacted in order to start creating the required profiles by industry sector.

Figure 4 shows the types of ratios needed for each industry sector. Figure 5 shows the ratios needed to then model the potential impact of increased HPC usage by industry sector.

The plan was to determine the key metrics that will be needed during the pilot study and then collect additional data in future phases. The pilot study was successful in determining which metrics can be collected and in starting the collection process.

FIGURE 4

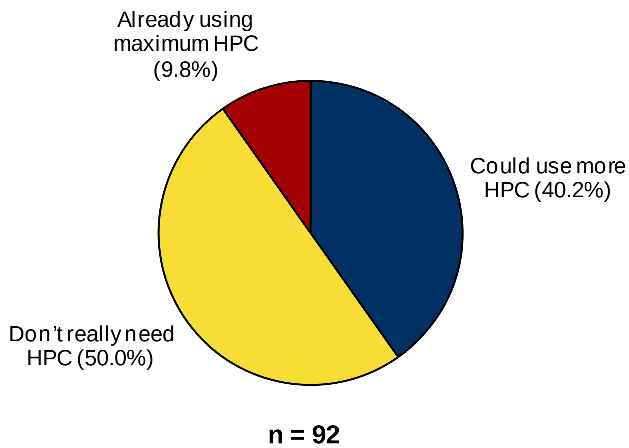
Example of Broad Metrics Required for Each Sector



Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

FIGURE 5

Example of Metrics Required for Each Sector



Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

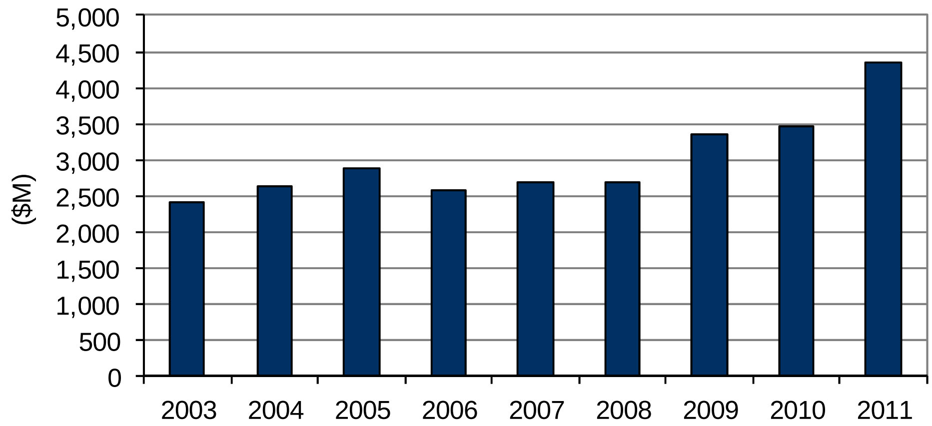
SITUATION OVERVIEW

Major Supercomputer Trends

Figure 6 shows worldwide supercomputer server sales from 2003 to 2011. There was a major slowdown in 2006–2008 and then a major growth spurt from 2008 to 2011 and in 2012.

FIGURE 6

Worldwide Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

How Nations Are Using HPC to Improve Their Competitiveness

Figure 7 shows the growth in supercomputer server sales in the United States, which have been strong but relatively flat from 2003 to 2008, with stronger growth to 2011. The United States is the largest market for supercomputers but has had dramatically lower growth over the past half a decade compared with countries like China, Germany, and France.

Figure 8 shows growth in supercomputer server sales in China, which have been extremely strong over the past 10 years, showing higher growth rates than any other country.

Growth in supercomputer server sales in China has been extremely strong over the past 10 years.

Figure 9 shows growth in supercomputer server sales in Germany, which have been consistently strong over the past 10 years.

Figure 10 shows growth in supercomputer server sales in France, which have been consistently strong over the past 10 years, with growth rate patterns similar to Germany.

Japan has had the most erratic growth rates. Note that single large supercomputer purchases can have a major impact on a single country. Figure 11 shows Japan's supercomputer purchases from 2003 to 2011. In 2012, there was a single sale at RIKEN for around $500 million, which is greater than all yearly supercomputer sales in Japan.

Japan has had the most erratic growth rates.

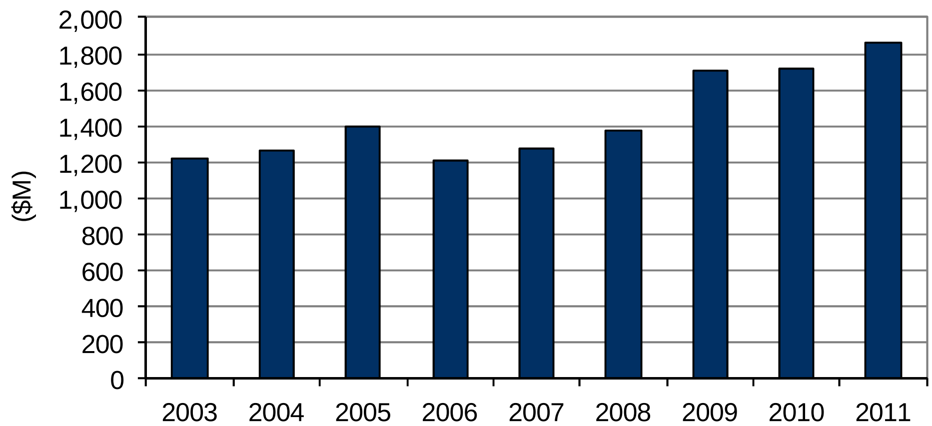
It is critical to the future of the United States to maintain a lead in HPC and supercomputer investments. It is clear that many other nations recognize the

It is critical to the future of the United States to maintain a lead in HPC and supercomputer investments.

importance of HPC investments to their future economic, scientific, and innovation success. IDC expects that more nations will decide to invest in R&D and HPC to help grow their national economies and generate stronger job creation.

FIGURE 7

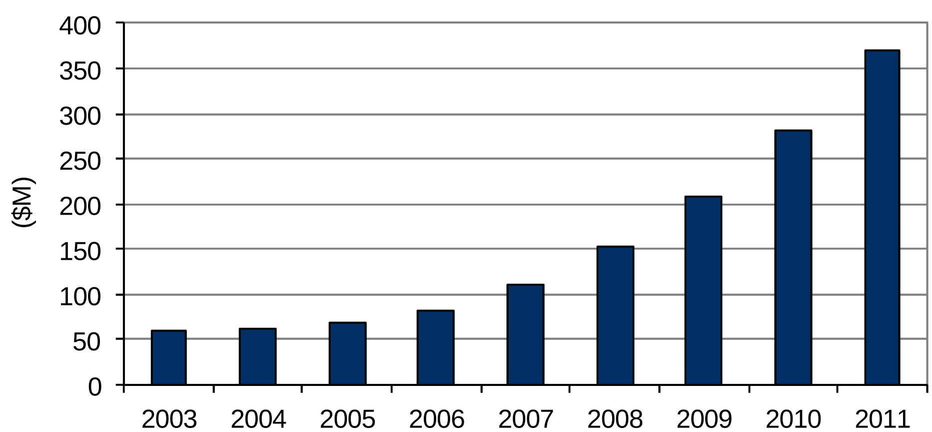
U.S. Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

FIGURE 8

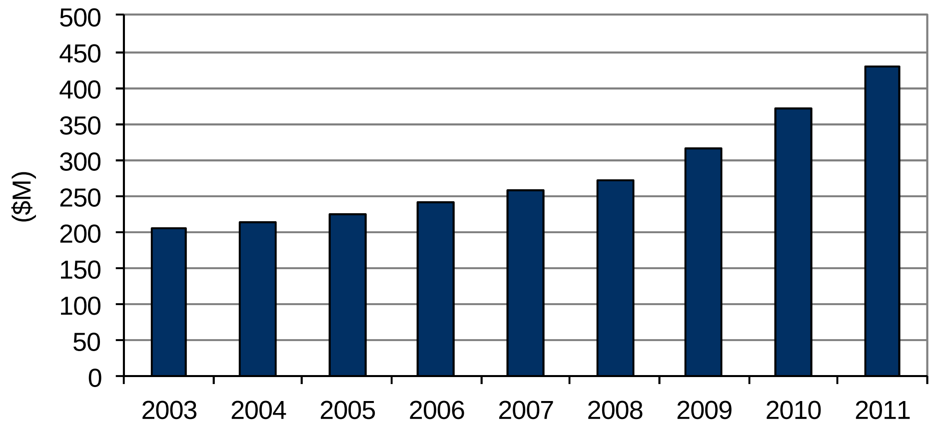
China Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

FIGURE 9

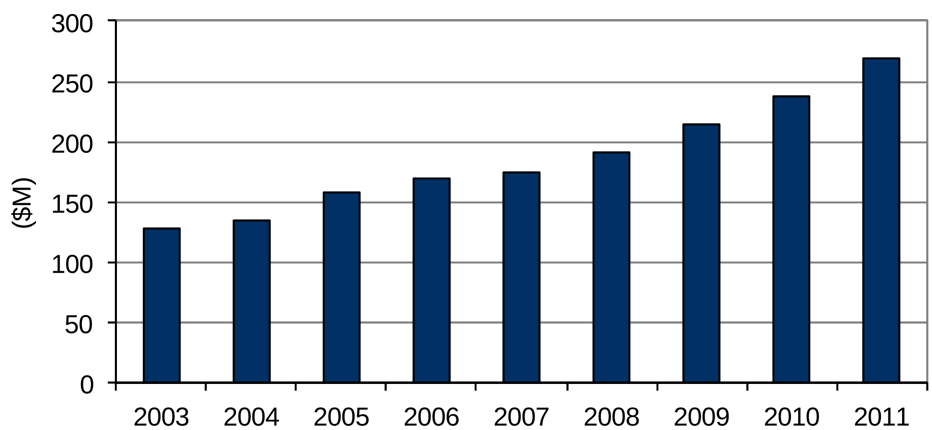
Germany Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

FIGURE 10

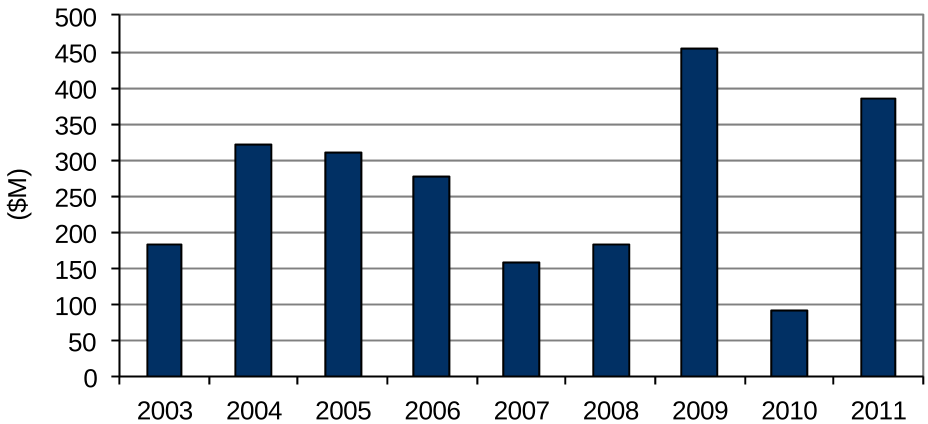
France Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

FIGURE 11

Japan Supercomputer Server Sales, 2003–2011



Source: IDC, 2013

KEY FINDINGS: FINANCIAL ROI

The main economic model IDC developed for the pilot study explores the relationship between HPC investments and the resulting returns in revenue, profits, and job creation. Future versions of the model will collect data on investments in total R&D more accurately and broadly to also show their ROI ratios.

Note that the results will change as more data points are added to the models.

Note that there were two sites with considerably higher returns that were removed from the data set as outliers:

- $10,057 in revenue per dollar of HPC invested at a finance company:
  - If included, this raises the overall revenue ROI from $356.5 to $536.2.
- $1,350 in profits per HPC dollar invested at an oil and gas company:
  - If included, this raises the overall profit ROI from $38.7 to $65.5.

Table 1 provides an overview of the results of the pilot study:

- For every HPC dollar invested, the sites in the study generated $356.5 in revenue on average.
- For every HPC dollar invested, the sites in the study generated $38.7 in profits on average.
- On average, it took almost two years from the end of the HPC investment before returns started.

For every HPC dollar invested, the sites in the study generated $38.7 in profits on average.

TABLE 1

Primary Financial ROI Results

Area	Pilot Model Results
Revenue ROI per HPC dollar invested ($)	356.5
Profit/cost-saving ROI per HPC dollar Invested ($)	38.7
Average number of years before returns start	1.9
Number of financial ROI examples in the pilot	67
Employee growth in the pilot sample	1,169

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Financial ROI Findings by Major Sector

Table 2 shows the differences in financial returns between industry, government, and academia. In industry, revenue per HPC dollar invested averaged $462.4 in the study sample. In government, it was much lower at $9.2 in revenue per HPC dollar invested, and in academia, it was $37.4 in revenue per HPC dollar invested.

In industry, revenue per HPC dollar invested averaged $462.5.

TABLE 2

Financial ROI Results by Major Sector: Primary ROI Findings

Sector	Number of Examples	Revenue per HPC Dollar Invested ($)	Profit per HPC Dollar Invested ($)
Academia	12	37.4	70.8
Government	4	9.2	3.9
Industry	51	462.4	36.4
Overall	67	356.5	38.7

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

With regard to ROI in relation to profits, industry averaged $36.4 in profits per HPC dollar invested, government $3.9, and academia the highest at $70.8.

Note that the industry sample was the largest at 51 examples in the pilot study, so the results for industry are statistically stronger.

Table 3 shows additional data on financial ROI returns by major sector:

- The average time between completing the HPC work and seeing financial returns was consistent for industry and academia (1.8–1.9 years) but lower for government (1.4 years).
- Employee growth/job creation was clearly driven by industry with 1,157 jobs created at the sample sites.
- Overall, for each job created, on average $93,000 was invested in HPC. Note that this number will likely change as we expand the sample size in future research.
- The HPC investment data per new employee for government is empty in Table 3 as there wasn't any relevant data in this subdata cut. Additional sample size can potentially help address this and allow for more representative coverage across the other sectors.

For each job created, on average $93,000 was invested in HPC.

TABLE 3

Financial ROI Results by Major Sector: Additional ROI Findings

	Number of Examples	Years Before First Return	Employee Growth	Total HPC Investment ($M)	HPC Investment per New Employee ($M)
Academia	12	1.8	2	11.7	0.050
Government	4	1.4	10	175.0	
Industry	51	1.9	1,157	192.3	0.095
Overall	67	1.9	1,169	379.0	0.093

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Financial ROI Findings by Industry

With regard to ROI by industry, finance and oil and gas were very high as expected, and transportation logistics was the highest at $1,793.2 in revenue returned for each HPC dollar invested. The lowest returns were in defense, government, telecom, and retail.

The highest profit returns were in industrial sites and in oil and gas. The lowest profit returns were in life sciences, government, and defense.

As noted previously, there were two sites with considerably higher returns that were removed from the data set as outliers:

- $10,057 in revenue per dollar of HPC invested at a finance company:
  - If included, this raises the overall revenue ROI from $356.5 to $536.2.
- $1,350 in profits per HPC dollar invested at an oil and gas company:
  - If included, this raises the overall profit ROI from $38.7 to $65.5.

In Table 4, employee growth data for academia, defense, research, and a few other segments isn't shown as there was no data in the subdata cut. Increased sample size can potentially help provide more representative coverage.

TABLE 4

Financial ROI Results by Industry Segment

	Number of Examples	Revenue per HPC Dollar Invested ($)	Profit per HPC Dollar Invested ($)	Years Before First Return	Employee Growth
Academia	3	31.3		1.0	
Defense	1	5.3	5.3	2.0	
Financial	21	508.4	47.7	1.8	532
Government	2	11.1	4.1	1.5	10
Industrial	1	357.1	100.0	1.5	
Insurance	1	71.4	7.1	2.0	5
Life sciences	3	59.9	3.0	1.8	
Manufacturing	13	28.1	21.6	1.3	31
Oil and gas	5	467.2	86.0	4.3	20
Research	6	35.1	35.1	2.0	
Retail	3	23.8	12.3	1.5	49
Telecommunications	1	10.0	10.0	3.0	210
Transportation	4	1,793.2	22.1	1.5	310
Environmental safety	3	30.0	107.4	1.3	2
Overall	67	356.5	38.7	1.9	1,169

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Financial ROI Findings by Country

Table 5 shows the differences in financial returns between countries. Because countries invest in HPC and R&D using different strategies, the data collected for this study does not justify direct "apples to apples" comparisons or rankings of countries. The ability to support such direct rankings is not a goal of this study. Nevertheless, comparing ROI results by country is useful as a way of evaluating the efficacy of the countries' various investment strategies related to HPC. Even for these comparisons, however, more data is required, especially for countries/regions other than the United States and the United Kingdom (both of which are more strongly represented in the pilot study). A key objective of subsequent phases of this study is to capture many more data points in all targeted countries/regions, including all of the countries shown in Table 5 and a growing number of other countries. In the limited data sample shown in Table 5, France clearly has the strongest revenue and profit performance, followed by the United States and the United Kingdom. China and Italy currently have the lowest ROI ratios, but these numbers may change substantially with the addition of more data points in subsequent phases of the study.

TABLE 5

Financial ROI Results by Country

	Number of Examples	Revenue per HPC Dollar Invested ($)	Profit per HPC Dollar Invested ($)	Years Before First Return	Employee Growth
China	3	8.7	5.4	1.3	30
France	4	621.7	125.0	5.1	
United Kingdom	31	366.5	26.7	1.6	896
United States	27	373.3	49.8	1.8	243
Italy	2	10.0	7.5	1.0	
Overall	67	356.5	38.7	1.9	1,169

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Financial ROI Findings by Organization Size

With regard to ROI by organization size, medium-sized companies tend to have the top ROI ratios and smaller companies the lowest ratios (see Table 6). But if the two outliers were included, the larger companies' ROI ratios would be more similar to those of medium-sized companies.

TABLE 6

Financial ROI Results by Organization Size

	Number of Examples	Revenue per HPC Dollar Invested ($)	Profit per HPC Dollar Invested ($)	Years Before First Return	Employee Growth
Large	39	373.1	37.0	2.1	1,084
Medium sized	12	661.4	62.0	2.1	57
Small	16	31.2	18.8	1.0	28
Overall	67	356.5	38.7	1.9	1,169

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Financial ROI Findings by Research Success Area

Table 7 compares the pilot study ROI examples based on the primary of type research performed. Creating better products and new approaches clearly provide the highest ROI ratios, while cost savings and helping research programs had the lowest returns. Some fields in Table 7 are empty as there was no data in those subdata cuts. Increased sample size will help provide more comprehensive coverage.

TABLE 7

Financial ROI Results by Primary Success Area

	Number of Examples	Revenue per HPC Dollar Invested ($)	Profit per HPC Dollar Invested ($)	Years Before First Return
Created new approach	27	408.2	39.4	1.7
Better products	24	479.6	34.2	2.1
Discovered something new	8	259.2	81.0	2.1
Helped society	5	50.1	14.1	1.4
Cost savings	2	13.4	13.4	1.5
Helped research program	1	2.8		
Overall	67	356.5	38.7	1.9

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Job Creation from the Financial ROI Examples in the Study

42 of the sites participating in the pilot study reported that their HPC investments led to job creation. At these 42 sites, a total of 1,251 jobs were created, with an average of 29.8 jobs created per site (as shown in Table 8). In future phases, more emphasis will be placed on finding job creation examples. (Refer to the data presented in Tables 9–12 for splits in the job creation examples by category.)

At these 42 sites, a total of 1,251 jobs were created.

TABLE 8

Jobs Created via Financial ROI Examples: HPC Projects

	Value
Number of examples that reported job creation	42
Average job creation at these sites	29.80
Total jobs created in these examples	1,251

n = 42

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Job Creation via Financial ROI Examples by Major Sector

Table 9 shows the job creation mix by major sector for the 42 sites that said they created jobs. As might be expected, most jobs were created in the industrial sector. The average HPC investment associated with creating a new job was about $107,000 for these sites that reported financial ROI gains from their HPC investments. (For all sites that created jobs, including those that reported financial ROI and those that created innovations without quantifiable financial ROI, the average cost to create a job was over $93,000.)

TABLE 9

Jobs Created via Financial ROI Examples by Major Sector

	Employee Growth	Average HPC Investment per New Employee ($000)
Academia	6	137.5
Government	20	
Industry	1,225	103.6
Overall	1,251	106.7

n = 42

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Job Creation via Financial ROI Examples by Industry

In comparing the different industries, we find that finance, transportation, and telecom created the most jobs in the pilot study sample of sites (see Table 10). In Table 10, some fields are empty as a result of lack of data in the subdata cut resulting from the limited sample size in the pilot study.

Finance, transportation, and telecom created the most jobs in the pilot study.

TABLE 10

Jobs Created via Financial ROI Examples by Industry Segment

	Employee Growth	Average HPC Investment per New Employee ($000)
Academia	2	312.5
Financial	565	110.0
Government	20	
Insurance	5	175.0
Manufacturing	46	20.2
Oil and gas	40	
Retail	49	151.5
Telecommunications	210	178.6
Transportation	310	97.3
Environmental safety	4	50.0
Overall	1,251	106.7

n = 42

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Job Creation via Financial ROI Examples by Country

Table 11 shows the mix of job creation by country. The United Kingdom and the United States reported the most jobs created, aided by the fact that a large majority of the surveyed sites were in these two geographies. There was insufficient data in the China subdata cut for average HPC investment per job created, resulting in that field remaining empty.

TABLE 11

Jobs Created via Financial ROI Examples by Country

	Employee Growth	Average HPC Investment per New Employee ($000)
China	60	
United Kingdom	904	102.6
United States	287	112.9
Overall	1,251	106.7

n = 42

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Job Creation via Financial ROI Examples by Organization Size

A comparison of job creation by organization size shows that large organizations created the most jobs (Table 12).

TABLE 12

Jobs Created via Financial ROI Examples by Organization Size

	Employee Growth	Average HPC Investment per New Employee ($000)
Large	1,115	125.6
Medium sized	87	67.5
Small	45	13.2
Overall	1,251	106.7

n = 42

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

KEY FINDINGS: INNOVATION ROI

Table 13 shows the number of innovations in the sample by the primary area of innovation. The sample had 67 basic research innovations and 74 applied research innovations. 79 of the innovations were in the public sector, and 62 were in the private sector.

The sample had 67 basic research innovations and 74 applied research innovations.

TABLE 13

Primary Innovation ROI Results

Area	Pilot Model Results
Total number of innovation ROI examples	141
Number of basic research innovation ROI examples	67
Number of applied research innovation ROI examples	74
Number of public sector innovations	79
Number of private sector innovations	62

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Innovation ROI Findings by Major Sector

208 ROI and innovation examples were collected and analyzed as part of the pilot study. This was made up of 67 financial ROI examples and 141 innovation examples (refer back to Figure 1). Of the 141 innovations, 67 were basic research innovations and 74 were applied research innovations. 79 of the innovations were in the public sector, and 62 were in the private sector.

The study was made up of 67 financial ROI examples and 141 innovation examples.

As Table 14 illustrates, the bulk of the innovations collected by the survey were from academic and industrial sites. Most of the basic research innovations were in academia, while most of the applied research innovations were in industry.

Government innovations were higher on the index scale that rated the significance of the innovations (averaging 7.0 out of a possible 10), while innovations in industry were ranked lower (averaging 5.7), and academia innovations averaged 3.9.

TABLE 14

Innovation ROI Results by Major Sector

	Number of Basic Innovations	Number of Applied Innovations	Total Innovations	Average Innovation Level
Academia	49	15	64	3.9
Government	8	6	14	7.0
Industry	10	53	63	5.7
Total	67	74	141	5.0

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Innovation ROI Findings by Industry

Table 15 shows the mix of innovations by market segment. The sample had a larger number of innovation examples in general research, manufacturing, academia, finance, life sciences, and oil and gas than in other segments.

There was a broad range in the innovation index ratings across the different segments, ranging from an average high of 8.0 to a low of 2.0. There was insufficient data in some subdata cuts, resulting in some fields remaining empty.

TABLE 15

Innovation ROI Results by Industry Segment

	Number of Basic Innovations	Number of Applied Innovations	Total Innovations	Average Innovation Level
Academia	8	11	19	4.3
Defense	1	3	4	6.0
Financial	3	8	11	5.8
Government	4	3	7	7.6
Life sciences	4	7	11	6.0
Manufacturing	6	27	33	5.8
Oil and gas		6	6	8.0
Research	38	4	42	3.4
Telecommunications		1	1	4.0
Transportation		1	1	4.0
Digital content creation		2	2	2.0
Agriculture		1	1	5.0
Environmental safety	3		3	6.3
Total	67	74	141	5.0

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Innovation ROI Findings by Country

Table 16 shows the innovations by country. The bulk of the innovations collected for the pilot study was from the United States and the United Kingdom. The United Kingdom sites rated their innovations lower than sites in other countries did, on average. Some fields in Table 16 are empty as there was no data available for those subdata cuts.

TABLE 16

Innovation ROI Results by Country

	Number of Basic Innovations	Number of Applied Innovations	Total Innovations	Average Innovation Level
China	3	10	13	6.8
France		4	4	8.5
India	1		1	8.0
United Kingdom	43	7	50	3.5
United States	20	48	68	5.6
Italy		5	5	4.0
Total	67	74	141	5.0

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

Innovation ROI Findings by Research Success Area

IDC split the 141 innovations collected in the survey into general categories based on the longer descriptions of the innovations supplied by the sites. The categories correspond to the major areas of success (i.e., benefits) of the research carried out by the sites with the help of HPC. Table 17 shows that the top 2 benefits were creating better products and creating something new. The next most common innovation success area was discovering something new, followed by helping society.

TABLE 17

Innovation ROI Results by Primary Success Area

	Number of Basic Innovations	Number of Applied Innovations	Total Innovations	Average Innovation Level
Better products	10	44	54	6.0
Created new approach	34	6	40	3.5
Discovered something new	10	10	20	5.2
Helped society	7	4	11	6.0
Cost savings	2	4	6	3.8
Helped research program	1	4	5	2.8
Major breakthrough	3	2	5	6.4
Total	67	74	141	5.0

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Average HPC Investments Required for the Innovations in the Study

Overall, approximately $500 million in HPC investments were made to generate the 141 innovations in the pilot study. On average, the HPC investment was $3.1 million per innovation, with many investments under $1 million per innovation.

HPC Investment per Innovation by Major Sector

Table 18 shows the mix by major sectors and that the average HPC investment was $3.1 million per organization. In the data collected for the pilot study, government sites tended to invest more in HPC per innovation ($4.4 million), while academic sites invested much less ($2.5 million) and industrial sites invested $3.8 million on average.

Government sites tend to invest more in HPC per innovation ($4.4 million).

TABLE 18

HPC Investment per Innovation by Major Sector

	Total HPC Investment ($M)	Average HPC Investment per Innovation ($M)
Academia	146.1	2.5
Government	218.3	4.4
Industry	132.3	3.8
Total	496.7	3.1

n = 105 innovations had clear data on HPC investments

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

HPC Investment per Innovation by Industry

Table 19 shows the variation in HPC investments, in total and per project, by market/industry segment. Oil and gas tended to make larger average HPC investments per innovation ($12.0 million), as did government sites ($7.9 million). The lowest average investments were in general research ($0.1 million), transportation ($0.2 million), and environment safety ($0.9 million).

Oil and gas tended to make larger average HPC investments (averaging $12.0 million), as did government sites (averaging $7.9 million) per project.

TABLE 19

HPC Investment per Innovation by Industry Segment

	Total HPC Investment ($M)	Average HPC Investment per Innovation ($M)
Academia	137.8	8.1
Defense	173.3	1.1
Financial	31.1	2.8
Government	39.6	7.9
Life sciences	18.0	2.6
Manufacturing	26.8	2.1
Oil and gas	60.1	12.0
Research	4.1	0.1
Transportation	0.2	0.2
Digital content creation	2.8	2.8
Environmental safety	2.8	0.9
Total	496.7	3.1

n = 105 innovations had clear data on HPC investments

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

HPC Investment per Innovation by Country

Table 20 shows the differences by country in HPC investments per innovation project. In the examples collected for the pilot study, France and China invested much more (almost 10 times higher) than the United States and the United Kingdom per project. Because the number of innovations collected in the pilot study for France (4) and China (13) was much smaller than for the United States (68) and the United Kingdom (50), the average investment amounts for France and China may change substantially as more examples are collected in later phases of the ROI study.

TABLE 20

HPC Investment per Innovation by Country

	Total HPC Investment ($M)	Average HPC Investment per Innovation ($M)
China	132.0	12.0
France	51.2	17.1
United Kingdom	52.6	1.1
United States	260.5	2.4
Italy	0.5	0.1
Total	496.7	3.1

n = 105 innovations had clear data on HPC investments

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

HPC Investment per Innovation by Organization Size

Not surprisingly, large organizations invested more on average ($9.5 million) than did medium-sized organizations ($4.4 million) or smaller organizations ($3.3 million) (see Table 21).

TABLE 21

HPC Investment per Innovation by Organization Size

	Total HPC Investment ($M)	Average HPC Investment per Innovation ($M)
Large	330.7	9.5
Medium sized	52.3	4.4
Small	77.0	3.3
Size not listed	36.7	0.7
Total	496.7	3.1

n = 105 innovations had clear data on HPC investments

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

KEY FINDINGS: A NEW INNOVATION INDEX

Index Development Methodology and Objectives

IDC created and tested four versions of the innovation index to produce an index that met the targeted criteria for a study on ROI related to investments in HPC and R&D. IDC required that the index meet two key criteria, in addition to being a useful scale on innovation:

Four different innovation indexes were created and tested.

- The index scale must be easy to understand and use by those organizations creating the innovation (and by external experts that review the ranking of the innovations).
- The index scale must represent the breadth of innovation types:
  - As a group, the surveyed organizations need to use the whole scale and not just a small portion of the index.

The Innovation Index

The two types of innovations studied included:

- **Basic research:** Major discoveries and pioneering breakthroughs
- **Applied research:** Incremental innovations and process improvements

Table 22 shows the final innovation index (version 4) that was used in the pilot study. The index level of 10 represents one of the top 2–3 innovations in the field over the past decade, while at the other end of the scale, level 1 represents an innovation that is only understood by experts in the innovator's field.

TABLE 22

HPC Innovation Index Scale Definitions

Innovation Index Scale	Innovation Level Description
10	One of the top 2–3 innovations in the past decade
9	One of the top 5 innovations in the past decade
8	One of the top 10 innovations in the past decade
7	One of the top 25 innovations in the past decade
6	One of the top 50 innovations in the past decade
5	It had a major impact and is useful to many organizations
4	A minor innovation that is useful to many organizations
3	A minor innovation or only useful to 2–3 organizations
2	A minor innovation or only useful to 1 organization
1	An innovation that is recognized only by experts in the field

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

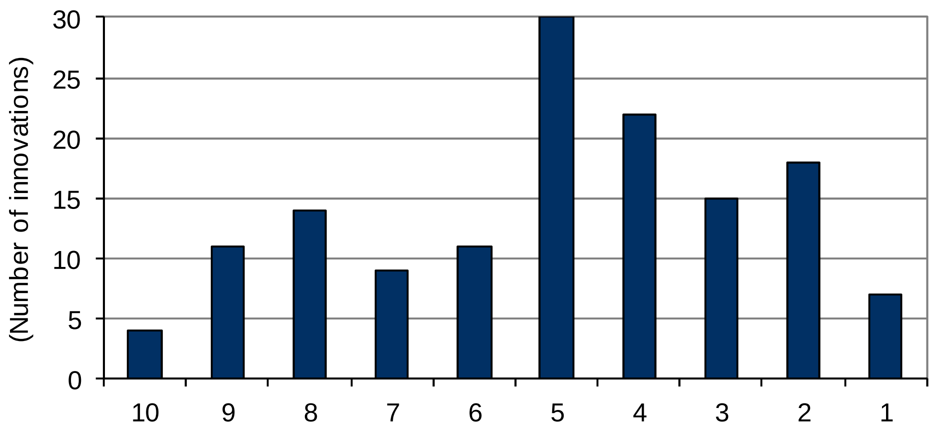
Figure 12 and Table 23 show the spread of the 141 innovations in the sample across the innovation index scale. The organizations in the sample used the full range of the index scale. The usage was fairly even, with a large number of 5s and 4s, and relatively fewer 10s.

The organizations in the sample used the full range of the index scale.

Overall, the average score was 5.0.

FIGURE 12

HPC Innovation Index Scale Results



n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

TABLE 23

HPC Innovation Index Scale Results

Innovation Index Scale	Pilot Model Results
10	4
9	11
8	14
7	9
6	11
5	30
4	22
3	15
2	18
1	7

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

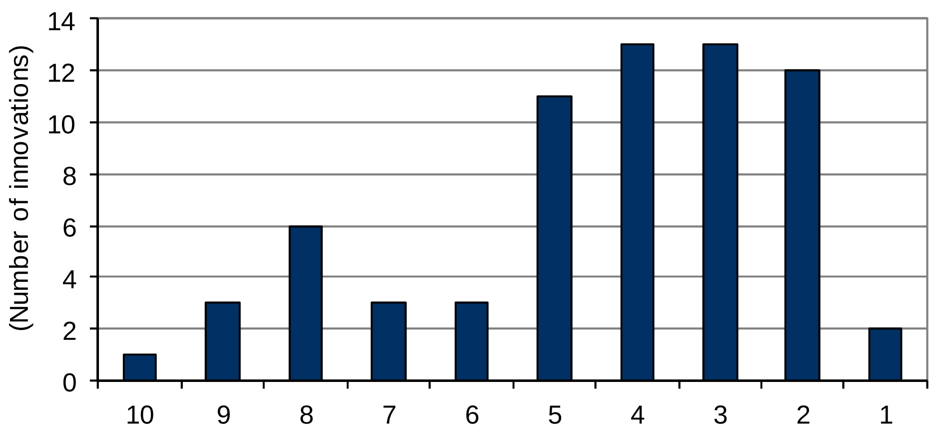
Basic Research Innovations

Figure 13 and Table 24A show the ranking frequencies of the 67 basic research innovations on the innovation index scale. A large portion of the basic research projects in the sample were rated from 2 to 5, and the vast majority were ranked under 6.

The average index score for basic research innovations was 4.4.

FIGURE 13

HPC Innovation Index Scale Results: Basic Innovations



n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

TABLE 24A

HPC Innovation Index Scale Results: Basic Innovations

Innovation Index Scale	Pilot Model Results
10	1
9	3
8	6
7	3
6	3
5	11
4	13
3	13
2	12
1	2

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

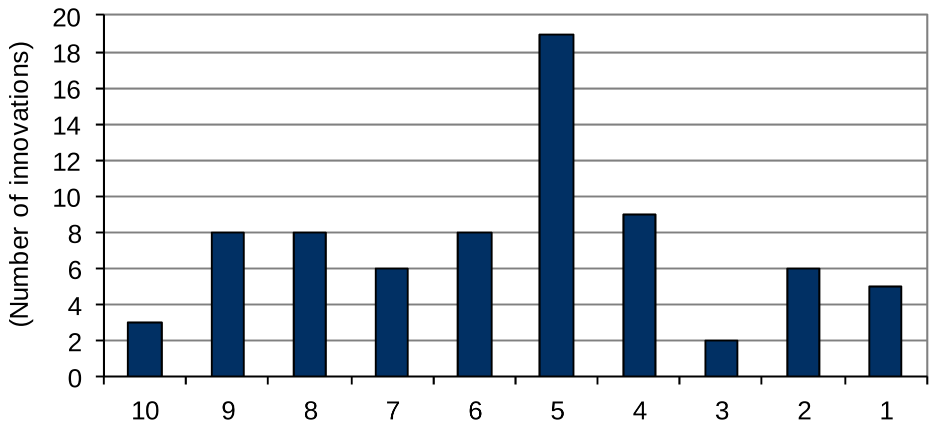
Applied Innovations

Figure 14 and Table 24B show the 74 applied research innovations. Applied research innovations were more evenly spread across the index scale, with many at the midrange ranking of 5.

The average index score for applied research innovations was 5.5.

FIGURE 14

HPC Innovation Index Scale Results: Applied Innovations



n = 74

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

TABLE 24B

HPC Innovation Index Scale Results: Applied Innovations

Innovation Index Scale	Pilot Model Results
10	3
9	8
8	8
7	6
6	8
5	19
4	9
3	2
2	6
1	5

n = 74

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

Innovations in the Public Sector

Table 25 shows innovation index rankings for the two types of innovations for the public sector organizations participating in the pilot study.

TABLE 25

Innovations in the Public Sector

	Public Sector Responses
Applied	
Innovation level	
10	1
9	2
8	2
7	3
6	1
5	8
4	2
2	1
1	2
Subtotal	22
Basic	
Innovation level	
10	1
9	1
8	4
7	3
6	2
5	8
4	12
3	12
2	12
1	2
Subtotal	57
Total	79

n = 79

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

Innovations in the Private Sector

Table 26 shows innovation index rankings for the two types of innovations for the private sector organizations contributing to the pilot study.

TABLE 26

Innovations in the Private Sector

	Private Sector Responses
Applied	
Innovation level	
10	2
9	6
8	6
7	3
6	7
5	11
4	7
3	2
2	5
1	3
Subtotal	52
Basic	
Innovation level	
9	2
8	2
6	1
5	3
4	1
3	1
Subtotal	10
Total	62

n = 62

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

TYING THE RESULTS TO EXISTING ECONOMIC DATA SETS

Research Results by Industry Sector

IDC's plan is to determine the key HPC ROI metrics during the pilot study and then apply these metrics to increasingly large data sets collected in future phases of the study. The pilot study succeeded in identifying key metrics that create the basis for tying the study findings to broader economic data measures, such as Standard Industrial Classification (SIC) codes and GDP reports, in later phases of the study. Refer to Table 27 for the key metrics (also refer back to Figures 4 and 5).

Research Still Required

Table 27 provides an example of how IDC plans to tie key HPC ROI metrics to broader economic data in future phases of the study. The broader data categories in this example are SIC codes that are used by government agencies in the United States and the United Kingdom (e.g., Companies House) for classifying industries. IDC provides Table 27 (with no data in the key metrics columns) for illustration purposes only, to show how real key metrics data might be linked to one or more widely recognized classification systems for broader economic data.

IDC estimates that it will take about two years to capture enough data to create confident linkages to broader economic data. For the pilot study, IDC reached out to more than 30,000 scientists and engineers across the world. Over 1,500 of these individuals completed a minisurvey focused on the key metrics shown in Table 27. We project that IDC will need to reach out to 5–10 times this number of scientists and engineers in subsequent phases of the study in order to create the basis for confident linkages to broader economic data.

Table 27 shows the data type required to map the ROI and innovation results to broader economic data sets.

TABLE 27

Example Data Needed to Tie to Broader Economic Reports

Industry	SIC Code	Percentage That Conduct R&D	Percentage That Also Use HPC in Their R&D	Percentage That Could Use More HPC	Percentage Using Maximum Amount of HPC
Agricultural production — crops	100				
Agricultural production — livestock and animal specialties	200				
Forestry	800				
Fishing, hunting, and trapping	900				
Metal mining	1000				

TABLE 27

Example Data Needed to Tie to Broader Economic Reports

Industry	SIC Code	Percentage That Conduct R&D	Percentage That Also Use HPC in Their R&D	Percentage That Could Use More HPC	Percentage Using Maximum Amount of HPC
Gold and silver ores	1040				
Crude petroleum and natural gas	1311				
Drilling oil and gas wells	1381				
Oil and gas field exploration services	1382				
General building contractors	1520				
Food and kindred products	2000				
Meat packing plants	2011				
Sausages and other prepared meat	2013				
Poultry slaughtering and processing	2015				
Dairy products	2020				
Ice cream and frozen desserts	2024				
Bakery products	2050				
Cookies and crackers	2052				
Tobacco products	2100				
Cigarettes	2111				
Textile mill products	2200				
Mobile homes	2451				
Household furniture	2510				
Office furniture	2520				
Pulp mills	2611				

Note: The Standard Industrial Classification (abbreviated SIC) classifies industries with a four-digit code. Established in the United States in 1937, government agencies use it to classify industry areas. The SIC codes group entries into progressively broader industry classifications: industry group, major group, and division. The first three digits of the SIC code indicate the industry group, and the first two digits indicate the major group. Each division encompasses a range of SIC codes. For example, entries from 2000 to 3999 represent the division Manufacturing. We use the three-digit industry group throughout this document.

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

FUTURE OUTLOOK

Using the Economic Models

The Multiphase Study

The IDC study on ROI related to HPC investments, consisting of the completed pilot phase and subsequent planned phases, is designed to capture a large, representative set of quantitative and qualitative data to show how investments in HPC create economic value — in the form of scientific and engineering innovations, along with additional corporate revenue, profits, and jobs. In the pilot phase of the study, funded by a grant from the U.S. Department of Energy's Office of Science and National Nuclear Security Administration, IDC created and tested economic models for the relationship between HPC investments and ROI and populated the models with more than 200 real-world ROI examples. Subsequent phases of the study, assuming they are funded, will refine the model and increasingly expand the data set in size and geographical scope.

A key objective of the multiphase study is to provide the tools (i.e., the economic models) and sufficient data to identify the economic value of HPC investments, both in general and in relation to specific industries and other market segments, as well as by global geography. A related objective is to enable funders and buyers of HPC resources to make more informed investment decisions, including (at the conclusion of the study) highly specific, predictive assessments of the likely benefits of certain investment levels, along with the time frames for realizing these benefits.

Today, it is widely assumed, based on a relatively small number of shining public examples, that HPC investments produce substantial financial and other returns, HPC-based modeling and simulation is widely recognized as the third branch of the scientific method, complementing theory and physical experimentation. In addition, a prior IDC study carried out for the Council on Competitiveness anecdotally established that the use of HPC is indispensable for 97% of the firms that have adopted this technology. But the present study is the first, to our knowledge, that set out to create reliable quantitative measures of HPC ROI, both on a very broad scope and with market-segment specificity for predictive purposes. Assuming that the remaining phases of the study are funded, in a little more than three years, HPC funders, buyers, and other interested parties should have a powerful, flexible new resource for helping them determine how much, and where, to invest in HPC in relation to other budget priorities.

The Pilot Study

The pilot study data set, in the form of Excel sheets and a pivot table, provides the ability to sort and analyze the data set in many ways. In addition, the "raw data" sheet allows deeper exploration of the full data set. The data can be sorted by industry, country, organization size, research areas, and other parameters.

The pilot-phase data set is much richer and broader for some parameters than for others. It still isn't large and balanced enough to support direct correlations and confident conclusions, especially when sorted into subsets. Substantially more

While the data set is very rich and broad, it still isn't large enough to make direct correlations, especially when sorted into subsets.

research is required to show and test the actual correlations, and many more data points are needed to pin down the relationships by country and by industry. In the meantime, the pilot study data is large enough to show the major trends but isn't large enough to support detailed economic comparisons.

Dissemination Plan

IDC, after approval by DOE, will present the results of the pilot study to a broad set of audiences, including at the annual supercomputing and ISC conferences, HPC User Forums, and other events. IDC will distribute the study and full data set to over 4,000 individuals in the HPC community. In addition, the study and data will be posted at a Web site (**www.hpcuserforum.com/ROI**). IDC will also be open to providing information to others on request.

The data will be maintained by IDC for at least five years. IDC will provide DOE with the full data set.

Proposed Future Research Possibilities

The Just-Completed Pilot Study

The goal was to design, create, and test multiple versions of the models, to populate them with a base set of "test" data, and to conduct a large set of short surveys to start developing ratios in order to apply the model results to existing economic data reports.

Next Steps: Creating Full-Scale Economic Models

IDC recommends that this pilot study be used as the foundation for creating actual full-scale economic models for financial ROI and innovation from investments in HPC.

This new research could follow a three-year, three-phase approach:

- **Phase I (Year 1):** The goal is to create the actual ROI models with a full data set at least two times in Phase I, growing to at least eight times in size by Phase III. This is needed to create predictive models for a fuller understanding of the relationships, to provide enough data/analysis to start making predictive results, and to refine the models as needed.

    - Motivate a select set of nations to fund IDC to collect deeper data samples within their countries.

    - The goal is to have enough data to start making statistically sound correlations between industries, countries, and different sizes of organizations.

- **Phase II (Year 2):** Expand the data set by at least two times more, including more countries and industries that are harder to collect data from, and refine the models as needed.

- Motivate a larger set of nations for deeper data samples within their countries.

- The goal is to have enough data to make strong, statistically sound correlations between industries, countries, and different sizes of organizations — and cross-correlations like industries by country and organization size by country.

- At this phase, it should be possible to start the build out of the tie to larger general economic data sets.

- **Phase III (Year 3):** Expand the data set by at least two times more and focus on hardening the predictive nature of the models by conducting additional research to "test" correlations. Refine the models as needed. It should also allow the fuller testing of trends and the impacts of changes made during the research period.

  - The goal is to conduct enough research to show both strong statistical correlation and causation between investments in HPC and the resulting ROI and innovation generation.

  - In this phase, the scenario testing should improve to be very robust and directly useful for making national policy decisions.

  - At this phase, ties to larger general economic data sets should be robust enough to be able to hand off to other government organizations.

APPENDIX A: SHORT DESCRIPTIONS OF THE 208 EXAMPLES

Refer to Tables 28 and 29 for the examples listed by the primary area of impact (refer to Table 28 for those examples that had a financial ROI, and refer to Table 29 for those examples that resulted in an innovation). The different areas of impact are:

- Better products

- Major scientific breakthrough

- Cost savings

- Created a new approach

- Discovered something new

- Helped society

- Supported research programs

Primary Research Areas of the Financial ROI Examples

Table 28 shows the primary research area for the 67 financial ROI examples in the pilot study. The ROI examples are sorted by the primary research area, and a short description of each ROI example is included. The longer descriptions are in Appendix B: Full Descriptions of the ROI and Innovation Examples.

TABLE 28

Financial ROI Examples by Primary Area of ROI

Primary ROI Area	Short Description	Primary ROI Area	Short Description
Better products	• $630,000 through financial simulation optimizations • Innovative engineering in race car design • $1.75 million revenue through manufacturing simulations • Innovative engineering in plastics manufacturing • $4 million savings in gas compressor system development • Improved tropical cyclone forecasting • $120 million savings through cancer simulations • $175 million through improved financial analytics • $87.5 million through financial analytics • High-density seismic data processing • $200 million revenue through improved insurance simulations • $175 million through optimized automotive simulations • Electroplating for processor development • $710 million savings through fraud detection	Created a new approach	• $380,000 through innovative data analytics • $380,000 through optimized data analytics • $1.75 million through improvements in rendering algorithms • $3.75 million through financial data analytics • $6.5 million profits through financial system improvements • $62.5 million in catastrophic risk analytics • $370 million revenue through improved financial analytics • $17.5 billion through integrated analytics • $17.5 million revenue through insurance analytics • $62.5 million through financial analytics • $37.5 million through booking system optimizations • $67.5 million through retail analytics • $3.75 billion through financial analytics • $175 million through high-performance financial analytics

TABLE 28

Financial ROI Examples by Primary Area of ROI

Primary ROI Area	Short Description	Primary ROI Area	Short Description
	• Maximizing oil reservoir production • $3 million savings — Quantum Monte Carlo simulations • Depth imaging on Block 32 in the Gulf of Guinea • $300 million revenue through in silico clinical trials • $875 million in healthcare insurance analytics • Advanced aerodynamics simulation of wing structures • $3 million revenue through plastics simulations • $0.63 million revenue through optical encoder design • $4.62 billion through transportation analytics • $67.5 million through transportation analytics		• $185 million through real-time investment decisions • $375 million through improved performance in finance • $1.75 billion through enhanced financial analytics • $1.75 billion through simulation-driven portfolio management • $17.5 million revenue through insurance analytics • $8.5 million through innovative financial infrastructure • $82.5 million through integrated financial analytics • $375 million from pricing and tariff analytics • $175 million revenue through enhanced financial analytics • $1.54 billion through debt analytics simulation improvement • $1.54 billion through improved financial analytics • $17.5 billion through consolidation analytics • $1.75 million revenue through simulations

TABLE 28

Financial ROI Examples by Primary Area of ROI

Primary ROI Area	Short Description	Primary ROI Area	Short Description
Cost savings	• Speed of materials science code quadrupled • $935 million savings from armor/antiarmor simulations		
Helped research program	• Enabled $17.5 million in grants		
Helped society	• Building industrial computing environments • Speed of materials science simulations • Enabled $8 million in grants and 62 publications • $8.25 million through NLP simulations • $6.25 billion through improved reservoir simulations	Discovered something new	• $3.75 million through blast impact predictions • Step change in turbulent combustion simulation • Turbocharged genomics simulations • Innovative earthquake prediction using simulations • Simulation innovations in storm surge predictions • Atmospheric chemistry simulations enhanced • Usability of oceanography code enhanced • Analyzing seismic profiles of a region of the Gulf of Mexico

n = 67

Source: IDC's *Return on High-Performance Computing Investment Survey*, 2013

Primary Research Areas of the Innovation Examples

Table 29 shows the primary research area for the 141 innovations in the pilot study. The innovations are sorted by the primary research area, and a short description of each innovation is included. The longer descriptions are in Appendix B: Full Descriptions of the ROI and Innovation Examples.

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
Better products	• HPC-powered testing and engineering • Innovative EcoBoost engine and NVH simulations • Enhanced surfactant simulations • Improved financial analytics accuracy • $630,000 through financial simulation optimizations • Innovative engineering in race car design • $1.75 million revenue through manufacturing simulations • Innovative engineering in plastics manufacturing • $4 million savings in gas compressor system development • Improved tropical cyclone forecasting • $120 million savings through cancer simulations • $175 million through improved financial analytics • $87.5 million through financial analytics • High-density seismic data processing • $200 million revenue through improved insurance simulations • $175 million through optimized automotive simulations • Electroplating for processor development • $710 million savings through fraud detection	Created new approach	• Multiphase CFD code developed • Improved sorting simulations • Innovative, data-intensive, tiled visualizations • Massive remote-batch visualizations • Materials science code enhanced • Materials science codes (CASTEP, ONETEP) enhanced • VASP materials science code optimized • CRYSTAL materials modeling code enhanced • Magnetic modeling code (MicroMag) parallelized • HELIUM code enhanced • Molecular dynamics code enhanced • RMT method for Schrödinger equation solutions • Scalability and efficiency of Navier–Stokes codes • Materials science code (CONQUEST) • Multiphase CFD code (Fluidity) • Magnetic plasma turbulence modeling code • Improved I/O for larger voxel-based simulations • Turbulent fluid flow simulations improved by 6 times

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
	• Maximizing oil reservoir production • $3 million savings — Quantum Monte Carlo simulations • Depth imaging on Block 32 in the Gulf of Guinea • $300 million revenue through in silico clinical trials • $875 million in healthcare insurance analytics • HPC-powered testing and engineering • Innovative EcoBoost engine and NVH simulations • Advanced aerodynamics simulation of wing structures • Enhanced surfactant simulations • $3 million revenue through plastics simulations • Improved financial analytics accuracy • Advanced plant genomics • Reduced fuel consumption and carbon emission in trucks • Improved aerodynamic efficiency of gas turbines • Innovative tire simulations • Simulating viability of biomedical devices • Advanced optical materials manufacturing • Innovative aluminum manufacturing • 10–100 times faster design times		• Turbulent fluid flow simulations improved • Aeroacoustics and geophysical fluid dynamics solvers • Electron-molecule scattering • Quantum mechanical simulation package • Simulations of catalytic chemistry 8 times faster • Speed and scalability of key materials science code • Aeroacoustics and geophysical fluid dynamics solvers • Performance of molecular dynamics application 20 times • Fluid turbulence simulations enhanced up to 40 times • Identified sodium nitrite's role in causing cancer • Turbulent fluid flow simulations • $50,000 savings through increased productivity • Enabled 30-year econometric simulations • Enabled global financial simulations • Improved financial analytics capability • Improved Web-profiling analytics • Innovative stream computing for econometrics • Innovative fund management simulations

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
	• Turbocharged golf club simulations • Improved manufacturing and packaging • Simulation-driven product packaging development • Advanced modeling of cellular networks and telephones • 1.26% efficiency improvements in wind turbine • 66% reduction in design cycles • 100% sales growth of racing wheels • Enabled testing simulation framework • Innovative simulations to enhance bottle manufacturing • Innovative financial analytics simulations • 4% improved wing-lift in aerospace • Scalable animation rendering • Improved aerodynamic simulation fidelity • Application acceleration on heterogeneous computers • Improved in silico modeling of genetic and clinical trials • Simulating Sun's corona and space weather • Acceleration of crash analysis for car design		• Acceleration of movie rendering with 3 times improved efficiency • $750 million through yield improvements • Leading-edge astrophysical research • Software platform for accessible HPC

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
		Discovered something new	• $3.75 million through blast impact predictions • Step change in turbulent combustion simulation • Turbocharged genomics simulations • Innovative earthquake prediction using simulations • Simulation innovations in storm surge predictions • Atmospheric chemistry simulations enhanced • Usability of oceanography code enhanced • Innovative cloud-based drug discovery • 100 times speedup in stem cell simulations • Innovative healthcare analytics • Improved data-intensive financial analytics • Oceanography code (NEMO) enhanced • $600 million revenue from genome sequencing • Geodynamic thermal convection simulations enhanced • Innovative chromodynamics simulations • Innovative genomic data analytics • Developed innovative 3D CT system • Simulation-based blood supply management • Innovative research in atmospheric sciences

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
			• Data mining innovations in genomics
Cost savings	• Speed of materials science code quadrupled • $935 million savings from armor/antiarmor simulations • Computational fluid dynamics solver • Improved simulating efficiencies • Innovative systems for financial simulations • Multiscale CFD for fluid catalytic cracking		
Helped society	• Building industrial computing environments • Speed of materials science simulations • Enabled $8 million in grants and 62 publications • Innovative evidence-based medical analytics • Heart modeling application 20 times faster • Geological simulations for city planning • Innovation in systems engineering • Innovative systems for satellite data analysis • Improved macromolecular simulations • Innovative public cloud platform • Data-intensive computing for high-energy physics	Helped research program	• Enabled $17.5 million in grants • Improved simulation times by 4 times • Innovative framework for HPC dissemination • Innovative codesign of modular HPC • Advanced HPC for scientific research

TABLE 29

Innovation Examples by Primary Area of Innovation

Primary Innovation Area	Short Description	Primary Innovation Area	Short Description
		Major breakthrough	• Simulation innovations in missile designs • Improved certification of airworthiness of UAVs • Improved prototyping of direct energy weapons • 175 times speedup in Hepatitis C simulation • Breakthrough brain research

n = 141

Source: IDC's *Return on High-Performance Computing Investment Survey,* 2013

APPENDIX B: FULL DESCRIPTIONS OF THE ROI AND INNOVATION EXAMPLES

This is a list of the individual examples collected and used in the pilot study. Note that a few may be repeated, as some organizations provided both an innovation and a financial ROI example on the same project.

The full list:

- Through faster computations (more than 175 times speedup), a better understanding of networks of coordinated amino acid variation may enable the discovery of new therapeutic targets for the hepatitis C virus (HCV). Assuming an estimated survival of 40 years, the annual healthcare costs for the affected U.S. population with chronic hepatitis C may be as high as $9 billion.

- The Continuous Casting Consortium has developed groundbreaking numerical models of the continuous casting of steel. A 1% reduction in yield loss could save about $400 million per year, along with energy savings of about $350 million per year during reheating.

- C-DAC's CHReME software has been deployed at HPC sites within India and as far away as Armenia and Tanzania. The software is reducing costs by tens of thousands of dollars while improving product cycle times and increasing productivity in critical projects.

- Swift uses HPC to develop prototype vehicles for the automotive and aerospace industries much faster than these industries develop production vehicles. The prototypes can save auto/aerospace firms tens to hundreds of millions of dollars in the design cycle. HPC enables Swift to explore aerodynamic design spaces 10–100 times faster than before.

- UTRC sped up a United Technologies computational fluid dynamics (CFD) code, enabling for the first time the simulation of realistic fuel spray–related applications. UTRC has seen reductions in design cycles of at least 66%.
- Researchers at the centers performed clinical trials in silico. Simulation of therapy equipment resulted in savings of more than a $100 million in infrastructure costs and over $12 million in research costs while enabling rapid advancements in cancer care research.
- BGI helped many customers in plant and animal genome, RNA sequencing, microorganism genome, metagenome sequencing, complex disease research, epigenomics, proteomics, drug discovery, and so forth.
- Developed add-on aerodynamic components called the SmartTruck UnderTray System that will substantially reduce the fuel consumption and carbon emissions of Class 8 long-haul trucks.
- Improved virtual prototyping of direct energy weapons.
- New version of Coupled Ocean/Atmosphere Mesoscale Prediction System (COAMPS), and improved tropical cyclone forecasting.
- Improved aerodynamic efficiency of gas turbines.
- Improved certification process for airworthiness of small UAVs.
- Improved performance of clinical trials in silico.
- 50,000-core cloud-based drug discovery on Amazon Web Services.
- Zipp Speed Weaponry CFD software brought pioneering new racing wheels to market, enabling 100% sales growth.
- eCommerce fraud detection and real-time stream analytics functionality for behavioral and situational modeling.
- Developing gas compressor systems for supersonic flight applications — IT savings from HPC.
- NCAR, as an FFRDC, performs basic research in atmospheric and related sciences. The HPC investments made by NCAR serve the scientific R&D projects and collaborative research projects of NCAR, with member universities for both numbering in the hundreds.
- This project enabled underwriting to aggregate the insured's life expectancy to enable us to predict their life expectancy for people who have impaired medical conditions. Going forward, the company can use the application to medically underwrite people rather than using skilled underwriters.

- Our HPC supports faculty research, which includes publications and grants. In 2011, our HPC resource supported about $8,000,000 in grants and 62 publications. (Self-reporting by faculty members.)

- This project monitored the increase in disposable income against the RPI/CPI indexes to review the trend over a period of 30 years to understand how social-economic factors influence incomes for the average consumer and their reach into the social infrastructure of the country. We were able to match the trends to suggested GDP shifts and movements successfully.

- We built an online data warehouse that was constantly updated to give real-time (ODS) reports and MIS data. These reports took all the data from our front-end application and processed large quantities of data to enable decisions to be made on investments in real time.

- We have taken over and merged with a number of banking entities where we have had to merge data. There has been a need to consolidate this data and run analytics over it to understand trends, exposure, and root debt risk.

- The project was to analyze our charging and cost models to see where savings could be made. It had to analyze each cost and compare this against industrial averages and rates to produce a mean cost. This enabled us to review our processes and implement savings to enhance the profit margin. Overall, our profits grew as we didn't increase our costs but looked at savings, which improved the bottom-line balance sheet.

- This project was to analyze our clients' transportation needs and trends to see whether cost savings could be made in the distribution of our trucks to be closer to clients and their logistical requirements.

- This project was set up to track customer activity with regard to their Internet usage. We wanted to track the sites they visited, the amount of bandwidth/data used, and the time it takes to render pages. By understanding this, we are able to recommend to other companies that they should be thinking about developing mobile-configured Web sites to reduce the amount of data that is used from mobile Web browsing. This would be seen as a bonus to the customer and give an advantage to the company.

- The project was to track and process all share prices for London, Frankfurt, Tokyo, and New York over a 12-month period for trend analysis. We would then overlay economic data to view enhanced trends.

- We were able to project areas of the borough where buildings were at the risk of subsidence due to the changes in climate and other factors such as construction and highway changes.

- The project was part of our course for medical students last year when they had to analyze the general public's diet against the cancer death rate in the United Kingdom. The objective was to connect the nation's diet to cancer conditions. The final results demonstrated the links between processed foods and certain stomach cancers. In particular, food with high fat content that also contained sodium nitrites increased the likelihood of cancer by up to 15%.

- Developed an enterprise system of mass optimization of full automotive vehicle design using high-performance computing and state-of-the-art optimization and automation processes and tools. The optimization includes all the major CAE simulation of vehicle performances, crashworthiness, dynamics, durability, and noise and vibration so that the best design is generated with the least mass and cost while satisfying all the performance targets.

- The project/solution involved the implementation of a customer loan/credit portfolio management application running on an HPC platform that rated customer financial position changes to provide proactive management of potential defaults on loans. This system assigned risk scores to the monthly/weekly changes in customer financial position based on external data reporting feeds (credit bureau's or other regulatory sources').

- Using HPC to build a Big Data — data mart, we can look at a profit-and-loss statement by line of business and by insurance coverage. This will allow us to attack and recoup premiums for our least profitable lines of business.

- The use of a shared HPC environment housed at a university to complete mandatory testing simulation work.

- This project was created to help reduce costs related to simulation and high math/computational algorithms. This study was also for the purpose of make/buy decisions.

- We undertook a project to measure the change in the population's blood types over the past 50 years. This was a directive and supported by the Department of Health to understand how blood groups have changed and how we ensure we keep adequate supplies through donor schemes. Currently, there is a shortage of blood in the rarest groups, and this puts the whole of the health service at risk in terms of operations and the event blood is required. The output of the study is driving the NHS to concentrate on keeping blood stocks highest on the most common groups but adequate supply on the rarest.

- Increase data process efficiency, increase accuracy of financial analyses, improve data mining, and create a culture that promotes collaboration, agility, and high performance for real-time financial reporting.

- Wells Fargo Services Grid is a multifunction grid used basically for distributed compute needs for multiple applications/LOB as well as a cache grid for low-latency applications. It is made up of heterogeneous hardware running either the DataSynapse grid software or the Oracle Coherence software product for cache. Future implementations will include grid services for standard middleware products like WebLogic, WebSphere, and JBoss. This has promoted the use of offloading high and large compute needs for multiple applications. The compute nodes communicate mostly through multicast and can be master and slaves for multiple units of work at the same time. This distributed methodology has enabled the use of the same set of compute resources based on the on-and-off processing cycles for different applications during the day.

- HPC was used to render various high-image content faster, thereby reducing time to market and production cost by approximately $150,000.

- Consolidation of four diverse and separate financial systems into company-standard software — Oracle Financing.

- Implementation of new stock market–forecasting software utilized to predict share price based on mathematical algorithms and historical data for retirement planning mutual funds. Implementation of multiple servers required in addition to the development of an interface for daily stock/mutual fund data downloads along with security protocols and software.

- New vehicle platforms based on existing architectures typically require additional weight, reducing crash and vehicle dynamic performance. HPC was utilized by L&L to leverage technical expertise to create design solutions, enlightening options to the customer such as maintaining performance at a lower weight, increasing performance at a slightly higher weight, and maintaining performance at a lower cost.

- We design optical encoders for the automotive, medical, military, and avionics industries. Optical encoders are devices that allow for menu selection of user interface panels. Instead of using mechanical contacts to perform the selection, it uses optics.

- Integrated/evolved the GPU library "QUDA" developed in collaboration with JLab, developed the Lattice QCD application and framework "Chroma", and deployed a cluster of GPU-accelerated nodes to yield performance per dollar 10 times as high as a conventional InfiniBand cluster, enabling a significant expansion in the reach of the calculations.

- Develop a practical industrial 3D computed tomography (CT) system from R&D work done at the corporate level. Our customer, the USAF, had a requirement to measure and verify the location of certain embedded objects. The objects were part of an aircraft wing. Due to the design and type of materials used, no conventional techniques would provide the necessary information. The R&D group had been working on a 3D computed tomography system for another project when it was proposed to use the CT system of our customer. At the time, the computational requirements were very high and were only accomplished via a high-end VME-based Unix system. The other large challenge was to move a system from the R&D phase to production. At the conclusion of the project, we achieved, what was, I believe at the time, the only in-production large-scale industry CT system that also was able to meet the requirements of our customer.

- We developed an industry-leading risk management program based on principles, practices, and models that provides our associates with the knowledge necessary to adopt and an ability to measure risk, identify profitable opportunities, and make sound financial decisions.

- Deployment of multitile display for extremely high-resolution images.

- Developed advanced simulation tools for the prediction of blast effects on a vehicle and its occupants considering the penetration of the blast wave through solid structure, the trajectory of any object inside, and the impact on its occupants' body by the transmitted shock waves.

- The project was to create an application to aggregate all risks written to calculate the exposure in the event of a major catastrophe such as Hurricane Sandy or 9/11. It overlaid the areas where risks were written to total exposure for the area. In the event of any occurrence, we could immediately calculate our cost to the balance sheet.

- We needed to have a way in which we could have online reports provided to us instantly, at any time of the day or night. The reason for this is we are operating in markets overseas and around the world, which reduces the overall operational window upon which we can update the data warehouse and cubes. We receive large amounts of information from clients and process this centrally and needed a fast processing system to give us the data we require. We chose an HPC solution to give us the performance and processing power we required.

- Use of HPC in science and engineering research.

- Built a supercomputer that is housed in a customized, modular, codesigned facility.

- We had a review by PWC to look into our financial reporting for both compliance and regulatory returns. Their analysis detailed gaps in our data warehouse whereby areas of the business were being missed in reporting, leading us to inaccurately report the current state of the business to shareholders, Lloyd's of London, and the FSA. This highlighted a bigger problem as Solvency II is the major regulatory project within the company where inaccurate data could have an effect on our ability to trade. The recommendation was to build a new data warehouse from the ground up to ensure all areas of the business were covered. The project is delivering a fast-performing and reliable data warehouse, enabling us to deliver the reports required.

- We needed to consolidate all our reinsurance transactions, accruals, and collections for both inward and outward transactions into a single application and database. This involved migrating all legacy data and processing them to give a true financial position. These volumes of data were huge, with ledger entries in the tens of millions. Each entry had to be processed and reconciled to ensure the correct values reconciled. The end accomplishment was a consolidated database with all data in the same place, which then can be reported on. This led to cost reductions in overall systems costs and staff costs.

- The project was set up to review the costs of bringing goods to market from supplier to eventual sale. The major driver was to bring down costs from supplier to warehouse to review the transportation and time to market. We found that if we tracked all goods sold throughout the group on an ongoing basis, we would be able to preempt trends when orders were needed to be made and make the most of transportation. Currently, transportation costs are increasing dramatically, and if we can make solid predictions when transportation is needed, we could improve the use of transport and combine pickups and delivery from multiple suppliers in the same logistical area. The result has been reduced costs, improved profit, and improved customer service as we find our shelves in supermarkets and retail outlets are never sold out.

- We built an application that allowed us to quickly build applications for HPC throughout the university. It was based on .NET and allowed very quick build and deployment as it built the required code in the background. This design framework was an innovation as it makes it possible for non-IT programmers to build applications that fit exactly to their projects and is used throughout the university.

- Due to the falling sales of nonfood items, there was a need to review the strategy and the items that sell and the ones that do not. We undertook a project to analyze everything that was sold over a period of 24 months to target the sales that are successful. The main driver for this was to halt the decline in sales on nonfood items and target growth in areas that are successful. The accomplishment was a turnaround in sales that brought the nonfood item areas back into profit and an increase in sales.

- The request came from our new owners, requiring faster reconciliation moving weekly job runs at the weekend to overnight runs. This has had significant impact on system requirements and business operations and changed the way in which trading tickets are processed. Although the front-end systems didn't change, the underlying database repositories moved from incremental updates to online data stores. This enables the data to be picked up on the overnight runs rather than weekends, and in turn, it didn't require the large operation windows required to process a week's worth of data. We are now able to run nightly runs using our HPC hardware and software to satisfy this request.

- The system analyzes how efficient the design of each car is against the engine power in PS. This information is gathered over at least 10,000 hours of wind tunnel and test circuit data. Each part of the car and engine is tested for efficiency to ensure optimum performance. We developed CMS to report back to designers the required data so any part that has not performed will be redesigned and retested. HPC is used to process and report back the required information.

- Throughout the organization, we massed large volumes of data being stored in many databases across a huge number of servers. Budget was made available to consolidate these databases into a number of large, clustered HPC environments to reduce costs and improve operational uptime. In addition, there was an issue with space as some database servers had large volumes of free space whilst others were becoming at risk of being overloaded. All these servers had direct-attached storage and were not connected to any SAN, etc. The decision to consolidate brought speed and resiliency benefits together with cost savings together with application availability and the ability to grow without restriction. The small number of HPC servers was attached to a new SAN that allowed disk space to be provisioned as the need arises.

- Following the purchase of Northern Rock bank from the HM government, one of the major projects was to integrate the mortgage and bank accounts and business into the Virgin core systems. As this was a high-profile acquisition, there was focus from both shareholders and the general public. Although the purchase had been announced 18 months before, the integration project had a life cycle of 12 months to be fully integrated. The volumes of data were huge, and the most important deliverable was that the integration of data should not have any performance effect on the existing systems. We were able to use the current HPC systems that have been proven to scale and perform adequately. The result was a seamless integration into the core systems with regard to data, although we did experience some data integrity issues.

- Amlin purchased Fortis Insurance from the Belgium government when the Fortis Bank was nationalized. A condition of the sale was that Fortis' systems needed to be migrated onto Amlin's core underwriting and reporting systems within 24 months. If this didn't occur, the Belgium government would fine Amlin daily until the systems were migrated. The premium, claims, and financial data went back over 35 years, and as the underwriting was mostly long-tail business, the data needed to be retained. What was required was a fast and efficient process of firstly taking the data out of the Fortis system, then processing it, including data validation and integrity checking, and lastly transferring onto Amlin's core systems. We needed it to be fast and reliable, so we used the Oracle database and Oracle HPC servers to process this information. We were absolutely astounded on the speed in which this information was processed and would use the same setup again. This project took the full two years but was delivered on time and to budget.

- We implemented a general ledger (GL) to track contractor costs and expenses throughout the group. Before, all of this was on spreadsheets, which made the tracking and reconciliation time consuming, with loads of errors. We decided that everything should be tracked electronically, which would make the month end, quarter close, and year end much faster to complete. We used Coda as a GL and were advised by them to place this on a server that was as powerful as possible following their support matrix.

- We have developed an underwriting engine that will estimate the life expectancy for individuals with medical conditions such as diabetes and cancers. This will then feed into our pricing engine to give an enhanced annuity rate. The application uses data gathered from different sources including our own and will process this to give a deduction on life expectancy. Although the calculations are simple when there is one medical condition, when there are multiple heath conditions the calculations around correlation matrices become complex. The look-up tables used are massive as the number of conditions and expected outcomes are vast. We needed a powerful calculations engine that would be able to process information within four seconds and as such needed powerful hardware.

- We had a project come in to aggregate all the funds we process and to include all the movements. Although we had a system to track fund amounts, we were are not able track and report on movements. We engaged with TCS in India to develop a system to track all movements so we could report on them. They developed a system using a high-performing server (HPC) to calculate the movements from the prior day to the current day.

- When we took over HSBC Insurance, we also took over their systems. Following a strategic review by the United States, they decided to consolidate the main broking applications into a single application. It was decided that we would design a system from scratch that would support the whole of EMEA. Due to the volumes of data, it was agreed that we would use high-end HPC servers to ensure that we had performance high enough that all of our users could go about their day-to-day work without any degradation. This project is still in flight; however, we have rolled it out to all users, with more enhancements still in development.

- The project still going on is to consolidate all data warehouses throughout the group to enable us to produce all reports from a single repository. We have over 28 different business and operational units throughout the world and have previously reported this on a disparate nature, and there was significant effort to bring this information together. Now we are able to report more quickly, and this reduces the cost and time to produce all statutory and legislative reporting data.

- Based on this project, there are many calculations that have to be done, which include impalements for patient-to-physician ratio, PMP payments, retro payments, etc. This is all part of the ACO (Accountable Care Movement). Diagnosis, DRGs, and other information will be collected to develop models in preventive care. We have set up complex data analytics/models that predict which patients are at high risk for hospital readmission.

- Allows large amounts of genomic data to be processed from raw to usable data in a time frame that is clinically acceptable, thus paving the way for individualized medicine pursuits.

- The project enabled us to track the upstream quality of resources being extracted through to the eventual release of all bio products including bitumen and finer oils. Through the process of the analysis, the results were constantly fed back into the system to enable us to track the quality of the raw resources and our refinement processes. This ensured we provided the best output and gave the best returns to the company. The vast volumes of data produced meant we needed an HPC environment to cope with the data.

- Correlate customer information from multiple contact channels to natural language and customer rep speech interactions (via voice-recording speech transcription) to identify major reasons for customer attrition, as well as retention saves to identify the best retention strategy for marketing.

- The HPC environment we used has made it enormously easier and cheaper to conduct large-scale research simulations of wireless network test beds. Our simulation time has been cut by a factor of four, and we can now do everything in-house.

- We developed a sorting algorithm that takes the basic idea of quick sorting. In the process of splitting the data, it takes three steps: an identifying step, position calculation using a prefix sum step, and a copying step. Every one of these three steps is highly parallelizable so we can achieve the maximum of concurrency. The end result is that our new algorithm is cost optimal, that is, we can prove that the cost of our parallel algorithm using p processors is the same as the cost of the best sequential sorting algorithm using one processor. (Note: "p" refers to any number of parallel processors.)

- We needed to upgrade our core systems as they were approaching end of life following years of low investment. A program was created to replace the system that processed charges on personal accounts for overseas customers. We undertook a review of all built systems on the market and could find nothing that suited our needs. As a result, we partnered with a third party to produce a system ourselves. We asked for consultancy from IBM, and they recommended we use their DX 360 HPC/M4 servers as a basis to host the applications.

- We delivered a HPC project to offer online investment quotations to IFAs (Independent Financial Advisors) for pension and investment products. The main success factor for the project was speed — from the moment the IFA requested a quote to the time it takes for a response to be produced. We use aggregator sites, which gives us the added constraint of if we don't return a quote with 20 seconds end to end, then we do not appear to quote on the Web site. We used high-end servers to give us the ability to process 1000s of quotes per minute if required and never be in a position where we have maxed out our processing capability.

- We have implemented a companywide sales and accounting system called Kerridge to streamline our accounts and purchase ledger and for ordering. We decided to use a single system for integration and costs savings; however, our major concern was how well the system would perform. We decided to use a thin-client rollout together with HPC for the database and applications parts. 18 months into the project, we had no performance issues reported with over 3,000 users. The average processor utilization is 35% for the database and 45% for the application servers. We have been able reduce the number of staff, which has lead to cost savings.

- Over the past five years, the financial services industry has been under close review following the banking crises in 2007. Following this, insurance companies are required to prove they have enough cash in reserves to pay claims on the business they write. The Lloyd's market has to prove to the regulatory authorities that as a whole we are solvent and have reserves to match our liabilities. We had to set up a database application to consolidate the entire market's data to report as a single entity to the authorities. This has meant we are receiving data from all the 120 underwriting companies within the market processing this and reporting as one. To put the volume of data in perspective, we now receive around 30GB of data per quarter. Following a review of hardware requirements, there was a need for high-grade servers to process the information and produce the required reports in good time.

- EuroCapital calculates GDP receiving information from many different sources, including government data, FTSE companies, and other financial feeds including Bloomberg and Reuters. There is a constant feed coming into the system that provides us valued information on the performance of the U.K. economy in real time. As the volumes of data coming into the system is constant and of high volumes, we required hardware and software to support this.

- Our database servers have been outgrowing our datacenter for a number of years, and we got to a point where we could not fit anymore servers into the racks we had. To overcome this, we decided to consolidate our application servers onto blade technology and the database servers onto a small number of high-performance servers. We carried out a review of what was on offer and decided on 4 Sun HPC servers attached to a Sun SAN. As a result, we were able to reduce from 23 physical servers down to 4. As a consequence, we dramatically improved performance across all applications, which made the business see tangible benefits.

- To remain competitive, we created a database to analyze all mobile and data usage to understand where and how our customers use their mobile devices. This gave us information that would influence our pricing strategy, and for both national and international usage. The data is reviewed to influence our pricing and tariff strategy and any offers we put out to our customers. We used multiple, clustered Sun servers and Oracle databases to process this information and give us the reports required.

- Following expansion and the acquisition of other travel companies, we had three different global hotel and airline booking systems. This made keeping track of availability of accommodations very difficult and made online bookings nearly impossible. We decided to integrate the two other systems into the main core system; however, that meant that the system would nearly double in size in data volumes and users. To ensure this system merge was successful, we upgraded the network and server hardware to make sure that the system was able to cope.

- The company needed new general ledger software to replace an old Sun accounts system. We needed a package that was high performing, was easily customizable, and would scale across other areas of the business. We picked Coda as it split the application server and database onto different hardware. We implemented both parts on high-performing servers, which enabled us to migrate and extend the software with integration into existing applications.

- We have over 30 customers that we service, and we run their applications to also include servers and databases. Our clients had been complaining of poor performance as the volumes of data had been growing steadily over the past 4 years. We decided to offer a new service where we would rent space on new HPC servers where performance would be significantly better. We surveyed clients to see what their appetite was for this service, and their response was positive, so we purchased 2 new HP ProLiant servers. Within 6 months, over

50% have migrated and are experiencing improved performance; thus we have created a new revenue stream, and within 12 months, these servers will have paid for themselves. At this point, we would be in profit mode and will look to increase this service even further.

- The project was to collate all financial data from financial services, manufacturing, and textiles companies and understand how the recession has affected profits, turnover, and efficiency. We gathered as much information from as far back from all stock exchange–listed companies and undertook detailed trend analysis on the performance of the companies and their sectors. We were looking for outperformers and whether this had any relation to their size, domicile, and customer base. In some cases, we went back 20 years and gathered vast volumes of financial data including share prices over the period.

- In the frame of the "Clean Sky" project, aimed at designing a new generation of environmentally friendly aircraft, Alenia Aermacchi is leading the Green Regional design team.

- While seeking the most promising configuration, two different wings have been shaped to fulfill separately the two goals of enhancing aerodynamic performances and reducing wing weight. In particular, two promising configurations were obtained, which respectively allow for an enhancement of 2.5% of cruise efficiency and a reduction of 4% of wing weight.

- Bankinter was founded in June 1965 as a Spanish industrial bank through a joint venture by Banco de Santander and Bank of America. It is currently listed among the top 10 banks in Spain. Bankinter has provided online banking services since 1996, when they pioneered the offering of real-time stock market operations. More than 60% of Bankinter transactions are performed through remote channels; 46% of those transactions are through the Internet. Today, Bankinter.com and Bankinter brokerage services continue to lead the European banking industry in online financial operations.

- Bankinter uses Amazon Web Services (AWS) as an integral part of their credit-risk simulation application, developing complex algorithms to simulate diverse scenarios in order to evaluate the financial health of Bankinter clients. "This requires high computational power," says Bankinter Director of Technological Innovation, Javier Roldán. "We need to perform at least 5,000,000 simulations to get realistic results."

- Through the use of AWS, Bankinter decreased the average time to solution from 23 hours to 20 minutes and dramatically reduced processing, with the ability to reduce even further when required. Amazon EC2 also allowed Bankinter to adapt from a big-batch process to a parallel paradigm, which was not previously possible. Costs were also dramatically reduced with this cloud-based approach.

- Bankinter used to spend anywhere from $100,000 to $200,000 on a traditional high performance computing–driven financial workflow. By migrating to the cloud, Bankinter saved more than 99% in infrastructure costs while reducing the average time to solution from 23 hours to 20 minutes.

- We did not have the internal firepower to do the type of processing we wanted to do for Elsevier's ScienceDirect data set, and it was cost prohibitive to purchase the needed resources. The high-performance computing and high memory instance types, AWS user documentation, and ease-of-use were primary drivers for using AWS. AWS provides a cost-effective, time-sensitive, large-scale data processing solution that is typically only available to enterprise-class businesses. If you have the technical know-how, AWS is a great equalizer.

- Most recently, we used AWS for calculating similarity scores and large-scale data modeling. Using multiple AWS HPC instances in parallel for a large-scale data mining algorithm, we achieved 95% runtime optimizations in both processes: The first job involved calculating similarity scores for a total of 8.6 trillion data pairs; running it on just a single core machine would take a whopping 9,856 days! Running on a single HPC instance, this would typically take our algorithm approximately 22 days to finish. With the use of multiple large HPC instances, we reduced the time for processing this data from 22 days to 33 hours, effecting a 95% optimization for the process.

- Most recently, we used AWS for calculating similarity scores and large-scale data modeling. Using multiple AWS HPC instances in parallel for a large-scale data mining algorithm, we achieved 95% runtime optimizations in both processes: The second job was optimizing a data modeling application to work on EC2's high-memory instance (m2.4xlarge), which contains a healthy 68.4GB of RAM and 8 cores. We optimized the data modeling application to best use the memory and cores available for the AWS high-memory instance and reduced our runtime processing for a job of 3.8 million ScienceDirect articles from 100 days on our infrastructure down to just 5 days of processing time on AWS.

- Experts in wind power units, specialists in CAD (computer-aided design), and CFD (computational fluid dynamics) professionals from different organizations were involved in a complex design scenario and needed to collaborate by effectively sharing different kind of information and resources. Engineers from the University of Trieste prepared the parametric CAD model and set up CFD simulations, while Airworks took care of aerodynamic performance calculations of the wind rotor blade and subsequently performed optimization analysis. The solution enabled the seamless execution of interorganizational simulation workflows. With Enterprise Suite, each partner of the simulation workflow was able to contribute to the project, providing its own high level of collaboration. The CAD expert generated the geometry of the blades, providing a parametric model and the CAD solution used to update it, while the CFD expert set up the aerodynamic simulation model, which used the geometry to evaluate the rotor aerodynamic performance. Both simulation workflows were made available to partners in a shared repository. Then the wind power unit expert was able to integrate his performance evaluation software into an automated multidisciplinary workflow. Ultimately, he used the workflow to evaluate the power efficiency of the system and to optimize it in a complete range of environmental conditions. "Through the entire process, simulation data and engineering knowledge were effectively managed and shared, allowing a faster process and a considerable resource saving," says Stefano Picinich, Airworks Engineering managing director. With the setup of the optimization workflow, Airworks professionals were able to explore and evaluate new parametric geometry, leading to innovative designs, analyzed by the decision maker via the Suite's postprocessing tools. The considerable result was of a wind turbine design with an outstanding power coefficient and an annual energy production increase of 1.26% and 0.47%, respectively.

- The purpose of this study was to determine the ROI for several of the HPCMP-supported projects, within the armor/antiarmor portfolio, that provide direct support to the Warfighter. The fundamental goals for using modeling and simulation in support of the armor/antiarmor development programs were to reduce time, resources, and risk while improving the quality of information available to designers, users, and decision makers. HPCMP capabilities enable the development and testing of new armor/antiarmor capabilities in the following areas: propulsion, interior ballistics, external trajectory determination, terminal ballistics, warhead analysis, and sensors, with a total upper-bound ROI of $935 million.

- This HPC innovation consists of a suite of cutting-edge computational efficiency enhancement methodologies applied to complex missile-related aero-propulsive problems with combustion. These methodologies are integrated into CRAFT CFD and CRUNCH CFD, two CFD codes in widespread use by DoD to support missile design and evaluation, and encompass gas-phase/multiphase combustion as well as laminar/turbulent chemistry, including tabulated/neural network approaches, reduced/adaptive chemistry, turbulent scalar fluctuation model (SFM), and GPU acceleration. By leveraging the state-of-the-art HPC resources provided by the DoD HPCMP, these innovative methodologies deliver improved higher-fidelity predictive capabilities to the analysis of missile systems/components, thereby enabling CFD to serve as a cost-effective design tool and virtual test bed for missile evaluation.

- Bottero has recently introduced on the market the E-MOC, a completely new family of mold opening and closing (MOC) mechanisms for the hollow glass industry. This wide project was very complex since it had to combine innovation in terms of molds movement (speed and precision), cooling systems, maintenance facility ("top mounted"), and "tradition," maintaining compatibility with previous standards. This aspect leads to hard space constraints, so a very complex three-dimensional kinematic schema was chosen, and it was impossible to tune in an analytical or experimental way. A high performance level is required for this mechanism, particularly regarding clamping forces in a closed position, the force available at the beginning of the molds stroke, reduced closing time, absence of vibrations during movement, and robustness with respect to irregularities in air supply.

- The central aspect of hollow glass production is the heat exchange, controlled by the MOC cooling system: It must be managed with high accuracy to ensure the correct formability of the molten glass. The E-MOC cooling system is based on a completely new concept since the air is coming from the arm and enters the mold at middle height: This ensures a wide range of advantages regarding the cooling system, which can be used on the complete machine cycle, and the mold air path, which can be customized to ensure the temperature profile most appropriate for a particular production. With respect to traditional concepts, it conjugates the "power heat extraction" of the axial cooling (usable only in the closed molds position) with the flexibility and precision of the radial one (usable only in the open molds position).

- In E-MOC, the air path shape changes with the arm position: In particular, the closed position duct is completely different from the opened one. To ensure the same air mass flow at the two or three molds that can be assembled on the arm, a deflector was introduced, designed trough optimization based on modeFRONTIER. This method is currently used in Bottero to ensure the correct air repartition of each designed mechanism version.

- Using the power of high-performance utility supercomputing, Victor Ruotti, a computational biologist at the Morgridge Institute for Research, ran over one million hours of compute — a compute century — in just one week. The unique HPC run was used to create an indexing system for stem cells and their derivatives, with the goal of building a database to speed development of potential therapies for diseases using stem cells.

- By using Cycle Computing's supercomputing software and infrastructure from Amazon Web Services, Ruotti was able to run 115 years of computation in just one week, for the price of $19,555. Without this technology, Ruotti would not have been able to complete this run. Theoretically, if Ruotti were to use an 80-core cluster on a sequencing instrument, this run would have provided a 100 times speed up (i.e., it would have taken him 100 weeks instead of 1).

- The project involved finding the optimum combination of weight distribution, brake distribution, and aero distribution for four typical Swedish circuits for the 2013 Swedish STCC/TTA season.

- The 2012 season of the Racing Elite League run by TTA introduced new rules, leaving little room for engineering design changes. Therefore, the Polestar Racing vehicle model had to be modified without considering the chassis design parameters, which were previously the core of the optimization analysis. Design simulation acquired an even greater importance as the best combination of the front-to-rear weight, aerodynamics, and brake distribution in less time (track testing was limited to three days) became the new target. The simulation tools used — Adams/Car by MSC Software and the lap time simulation (LTS) in-house code — were easily integrated into the modeFRONTIER optimization workflow to optimally tune the front-to-rear weight distribution parameters. "Considering the limited amount of time we have had for testing in this championship, this kind of interrelated parameters study would not have been possible without the capability of modeFRONTIER to extend the simulation analysis," says Per Blomberg, manager of Chassis Development. Improvements in weight, aerodynamics, and brake distribution led to a lap time reduction ranging from 0.19 to 0.50 seconds. In only a few hours of simulation time, modeFRONTIER piloted the evaluation of parameter combinations that might have never been tested, leading to enhanced configurations. The optimization framework initially set up for the TTA championship can be promptly adapted to the typical circuits of the STCC championship. Repeating the study considering different conditions is now "something that we can do even overnight on a weekend race, once we have completed the trial session and gathered some data about the track model." Similarly, it is rather easy to run a new optimization cycle.

- This project uses the coupled coastal circulation, storm surge, and wind wave model ADCIRC+SWAN to produce high-resolution forecasts of storm surge, nearshore waves, and water inundation when tropical or extratropical storms threaten the U.S. Atlantic or Gulf Coasts. The complete system, the ADCIRC Surge Guidance System (ASGS), includes an advanced Web-based display (NC-Coastal Emergency and Risk Assessment). During an active storm, ASGS is run 2–4 times each day on a 150-node Dell PowerEdge M610 cluster (2 x 2.8Ghz Intel Nehalem-EP 5560, quad core) at the Renaissance Computing Institute (RENCI). The output from these runs are incorporated into guidance and forecasting efforts by the National Weather Service, the National Hurricane Center, and agencies such as the U.S. Coast Guard, the U.S. Army Corps of Engineers, and FEMA as well as local and regional emergency management personnel. The resulting forecasts are used for evacuation decisions, to position supplies and response personnel, for search and rescue, and for other event-based decision support as needed.

- A comprehensive informatics framework enables medical decision support by systemizing genomic analysis and high-performance computing to mine genomic data for clinical and research use. The innovation focuses on improving patient care by optimizing a clinician's ability to identify treatment options that work. Using a data evidence–based approach, the identified evidence on similar patients is personalized and transformed into actionable data for quick use at the point of care while providing predictive insights on treatment options. The ability to do so offers the potential to reduce errors in treatment and prescription-related errors and drive down the cost of patient care.

- The Compression Molding Machine (CCM) produced by the packaging division of SACMI transforms plastic resin pellets into semifinished caps. During the injection phase, a pneumatic piston allows for the melted pellet to be injected inside the mold. The piston is subject to a high acceleration rate and continuous collisions, which call for a compromise between stress and weight. In order to limit the risk of failure, find a light but strong piston geometry to improve the Compression Molding Machine performance in terms of tensile and yield strength. The first optimization study was carried out manually for a 12mm piston, while the second study on a 16mm piston took advantage of the modeFRONTIER optimization platform by capturing the piston model, previously created in SolidWorks and analyzed in ANSYS Workbench, and improving the geometry and performance. Piston geometry modifications led to a 33% safety index increase in the case of manual analysis and 44% in the case of the model optimized with modeFRONTIER. modeFRONTIER enabled the CAE expert to exploit his original inspiration — which led to an improved piston geometry — by evaluating a higher number of configurations (+240%) in less time and by further enhancing the safety performances. The time for the optimization process dropped from 20 days to 4 days, 3 of which were required for manual setup and 1 for automated evaluation.

- The High Performance GeoComputing Laboratory, University of California at San Diego, has developed a highly scalable and efficient GPU-based finite difference code based on AWP-ODC, a community code developed and supported by the Southern California Earthquake Center for large-scale earthquake simulations. The AWP-ODC GPU achieved perfect scalability on Oak Ridge National Laboratory's (ORNL's) Titan and was used to simulate realistic 0–10Hz earthquake ground motions, the largest ever earthquake simulation performed. This code was restructured to enable maximized throughput, reduced time to solution, and scalability. The code achieves 100% parallel efficiency on 8,192 GPUs and sustained 2.33PFLOPS on Titan. Moreover, this GPU-powered code has been transformed to calculate Strain Green Tensors, resulting in a speedup in key strain tensor calculations critical to probabilistic seismic hazard analysis by a factor of 110. This achievement makes a California-statewide hazard model a goal reachable with existing supercomputers. The performance of the code is expected to take physics-based seismic hazard analysis to a new level using petascale heterogeneous computing resources, saving more than 500 millions of core hours as required by building engineering design. UCSD development team members are as follows: Yifeng Cui (lead), Jun Zhou, Efecan Poyraz, and Dongju Choi.

- Dario Alfè, principal investigator and a major user of CASINO, said, "The new shared memory facility is effectively speeding up the code by a factor equal to the number of cores per node for large jobs (i.e. a fourfold increase with the current quad cores, but clearly set to increase in the future)." [Note: HECToR was upgraded to 24 cores per node in 2010.] "The new second level of parallelism will allow an efficient use of at least 4 times as many cores as previously possible, therefore increasing the scalability of the code to well over 100,000 cores. Finally, the rewriting of the checkpointing routines (that were found to choke on more than 10,000 cores) allows us to cut restart times from over 1 hour to a few seconds. This work has helped CASINO to affirm itself as a modern code, well capable of exploiting current and future massively parallel machines."

- Dr. Graham Mann (University of Leeds), who was the principal investigator of the project and also a major user of GLOMAP/TOMCAT, estimated that users of GLOMAP-mode consumed around 3.2 million AUs (allocation units) of XT4 resources over a 15-month period on HECToR with a notional cost of £44,000. When extrapolated across future research using GLOMAP-mode TOMCAT code on HECToR and other supercomputers, the optimizations could deliver significant cost savings. The key result for researchers, however, is the reduction in CPU time per model time step, thus making it possible to do new science.

- Commenting on the dCSE success, Mann said, "The optimization work by NAG means the code runs 15–20% faster in general and, importantly for us, delivered much more economical scaling to 256 cores on the XT4. This improved scaling means we can achieve significantly reduced runtime per time step, thus enabling the investigation of scientific scenarios requiring many more time steps." Mann also noted that "NAG's addition of another level of parallelism to the code via OpenMP will also enable more efficient use of the XT6 (Phase 2b) supercomputer with its much greater number of cores per node."

- DL_POLY_3 is a general-purpose package for classical molecular dynamics (MD) simulations from STFC's Daresbury Laboratory. University of Warwick researchers Mark Rodger and David Quigley, in collaboration with colleagues at the University of Sheffield, used DL_POLY_3 and the HECToR supercomputers to study the role of a protein called ovocleidin-17 (OC-17) in chicken eggshell formation. Significant performance improvements were needed to make the modeling possible in feasible timescales using the HECToR supercomputers (especially in terms of parallel I/O).

- Commenting on the dCSE project success, David Quigley said, "Prior to these I/O improvements, DL_POLY_3 was unable to make effective use of the parallel file system on HECToR, severely crippling the performance of our simulations. The new code has reduced the time taken to write a single snapshot from 3 minutes to less than half a second, resulting in an overall factor of 20 improvement in our net performance."

- Christos Vassilicos, principal investigator, said, "In order to better understand the origins of the original properties of multiscale objects, it was necessary to undertake high-fidelity simulations of such unique complex flows. Because of the complexity of the flow configuration, these simulations require hundreds of millions of mesh points, and thus it was almost impossible two years ago to perform the numerical modeling. The new version of the code is a blessing as it will speed up our research and will allow us to do accurate comparisons with experiments."

- The CABARET (Compact Accurately Boundary-Adjusting high-REsolution Technique) code may be used to solve the compressible Navier–Stokes equations and, in the context of this project, for the investigation of aircraft noise. An important component of aircraft noise is due to airframe/engine installation effects; the reduction of this remains a very challenging problem. In particular, when deployed at a large angle of attack at approach conditions, the wing flaps become a very important noise source. For engine-under-a-wing configurations, flap interaction with the jet can even become a dominant noise component. A crucial element of any noise prediction scheme is the high-fidelity large eddy simulation (LES) model. For the airframe/engine noise problem, this model needs to accurately capture all important wing-flap, free jet, and wing-flap-jet interaction effects.

- Commenting on the dCSE project success, Sergey Karabasov, principal investigator, said, "The dCSE project has been very useful in several respects. The acceleration provided by empowering the general-purpose LES CABARET code with HPC capabilities on HECToR allows us to use a significantly higher resolution than was previously possible on local small-size clusters."

- The project was focused on three enhancements: Improving the scalability of DSTAR by incorporating a 2D domain decomposition scheme, optimizing the I/O operations for monitoring the evolutionary physical quantities in the simulation, and refactoring the legacy FORTRAN 77 code. Commenting on the dCSE project, Prof. Kai Luo of the School of Engineering and the Environment at the University of Southampton said, "This dCSE project has transformed DSTAR by incorporating state-of-the-art parallel programming techniques for multicore architecture. It prepares DSTAR well for the next generation of massively parallel computers in the United Kingdom running at petaflops. The new DSTAR code is able to exploit multicore HPC architectures and run jobs efficiently on a vastly larger number of processes. This allows the code to make use of much more computing power than before and solve more complex problems. For example, more realistic chemical mechanisms can now be simulated. As a result, combustion simulation will be more accurate or combustion of more types of fuels can be explored."

- CASTEP is a commercial (and academic) software package that uses density functional theory with a plane wave basis set to calculate electronic properties of solids from first principles. CASTEP is a fully featured first principles code, and as such, its capabilities are numerous. Aiming to calculate any physical property of the system from first principles, the basic quantity is the total energy from which many other quantities are derived. (For details, go to **www.castep.org**.)

- Dr. Andrew Coward, who is the manager of the Global Ocean Modelling Consortium on HECToR, estimated that their group used around 6 million allocation units running NEMO last year. Reducing the wall clock time of NEMO by up to 25%, as enabled by this dCSE project, could result in a saving in AU cost of as much as £95,000 per year (up to £400,000 for the remainder of the service) for only six months of person effort. Other consortia using NEMO on HECToR have used around 40 million AUs over the same period, potentially meaning multimillion-pound savings overall.

- Dr. Ben Slater of University College London (UCL), principal investigator, and his colleagues have been able to conduct a detailed study of defect segregation in ice as a result of this HECToR CSE work. Commenting on the success of the project, Slater said, "For the published study, it was imperative to use relatively large models of ice slabs in order to capture the detail of the ice surface structure and probe the surface segregation energies of molecular defects. Using the optimized version of CP2K, we were able to routinely model cells containing over 1,000 atoms. The CSE investment into CP2K was crucial to making the study tractable."

- Dr. van Hunen, from the University of Durham, who is both the dCSE project principal investigator and a major user of CITCOM on HECToR, has used around 1.8 million AUs at a nominal cost of around £25,000 so far on HECToR. When extrapolated across future research on HECToR and other supercomputers used to run the CITCOM code, the optimizations achieved by this dCSE project could deliver significant cost savings. The key result for researchers, however, is the scalability and faster convergence, thus making it possible to do new science. In particular, the significant development of local mesh refinement in combination with multigrid, and future work based on this important pioneering study, will enable CITCOM users to address much larger numerical problems.

- This dCSE project breathed new life into a code that was facing extinction with regard to its use on current and future HPC resources. As a result of Dr. David Scott's work, the EBL code is now — and is expected to continue to be for the foreseeable future — able to again perform the numerical experiments on canonical turbulent flows for which it was designed. Describing the new science enabled by this project, Dr. Gary Coleman of the University of Southampton said, "The current focus is upon understanding the interaction of arrays of wind turbines and the atmospheric boundary layer within which they are embedded. This should yield important new insight into how best the turbines should be arranged to optimize their efficiency. EBL will also be applied to more fundamental studies, to address questions such as the degree to which wall-bounded turbulence exhibits universality at Reynolds numbers much higher than previously considered, and how best this type of turbulence can be modeled by engineers and meteorologists."

- Added an extra layer of parallelism to the ChemShell package, enabling simulations to scale to thousands of cores and complete up to eight times faster than previously.

- Describing the science impact, Principal Investigator Prof. Catlow (UCL) said, "In a number of projects, particularly within the themes of chemical reactivity, and nanomaterials and defects, the work of the Materials Chemistry Consortium depends crucially on the advanced computational chemistry tools provided by the code ChemShell. One important functionality of this software is the development and utilization of complex computational models including hybrid quantum-mechanical/molecular-mechanical embedding. Members of the consortium apply these models to study the structure and reactivity of heterogeneous catalysts, physical and chemical processes on the defects, and active sites in the bulk and at the surface, and interfaces of advanced functional materials. One of the ubiquitous and nontrivial tasks in this work is the CPU-intensive investigation of the potential energy hypersurfaces, or landscapes. The latest software development within the dCSE project allows us to run efficiently across in excess of 1,000 cores, and there is also a reduction of this most time-consuming part of approximately 50–90% compared to the older implementations."

- Fluidity-ICOM is built on top of Fluidity, an adaptive unstructured finite element code for computational fluid dynamics. It consists of a three-dimensional nonhydrostatic parallel multiscale ocean model, which implements various finite element and finite volume discretization methods on unstructured anisotropic adaptive meshes so that a very wide range of coupled solution structures may be accurately and efficiently represented in a single numerical simulation without the need for nested grids.

- Commenting on the dCSE project success, coinvestigator Dr. Gerard Gorman (Imperial College London [ICL]) said, "Achieving excellent, strong scaling up to 4,096 cores broke new territory for Fluidity-ICOM. Moving from computational platforms of the scale available at universities to HECToR allowed scientists to 'think big.' For many complex physical problems, resolution is a key limiting factor to what can be simulated. Indeed, a lot of physics has to be parameterized in crude ways in order to allow simulations to execute within a given computational platform. Using HECToR with this dCSE-improved code greatly increases the accuracy to which the physics can be captured."

- CASTEP is an electronic structure code based on the density function theory used in areas such as inorganic and organic photovoltaic materials, catalytic reactions at surfaces, light-emitting polymer materials for optical displays, and femtosecond laser chemistry. CASTEP is one of the most heavily used applications on HECToR. Commenting on the dCSE project success, Keith Refson, principal investigator, said, "This project has developed a major new functionality in CASTEP, allowing the computation of excitation spectra using TDDFT. We anticipate that this will have twofold benefits in the next year — first, enabling new science using the new functionality as developed and, secondly, using this as a basis for further methods development for the treatment of excited states. A poster on this work presented by Dominik Jochym at the recent 'Trieste Total Energy' electronic structure conference was well received by peers and potential users."

- Ported the AVS/Express Distributed Data Renderer (DDR) visualization application to HECToR's Cray XT4. This enables researchers to perform visualization of data sets that are too large for local GPU hardware. AVS/Express DDR provides parallel module processing, where various visualization techniques are applied to domain-decomposed data. Parallel rendering and image compositing, utilizing the distributed memory on HECToR, allow significantly larger data sets to be visualized. The initial users are materials scientists that wish to examine (large) volume data sets acquired through CT x-ray scanning techniques.

- Commenting on the dCSE project success, Martin Turner, principal investigator, said, "The MRBV project has allowed us to run our visualization code on much larger problem sizes than have been possible on any of our existing visualization hardware. In particular, we are able to visualize a 350Gb volume. It has also allowed us to verify the approach taken in developing the parallel AVS code. Prior to running on HECToR, the product had only been tested on small-scale clusters (32 cores). While certain architectural changes within the software were needed to run the code on HECToR, it is useful to see the code scaling to 1,024 cores." CARP is a widely used software package designed for large-scale simulation studies of hearts. The aims of these studies are to provide detailed personalized therapies for treatment of medical conditions such as cardiac arrhythmias or heart failure. The areas where incorporation of in silico modeling into the clinical workflow are feasible include cardiac resynchronization therapy (for treatment of heart failure), drug trials (looking for possible side effects), and development of better defibrillators.

- Commenting on the dCSE project success, Gernot Plank, principal investigator, said, "The progress achieved during the 8-month dCSE project led to a step change in code performance. For the large problem we are interested in, that is the simulation of a human heart beat, the code performance improved by a factor of 20, which enables us to simulate one cardiac cycle in less than 5 minutes. This opens entirely new perspectives for cardiac modeling in the context of clinical applications where simulations are expected to play an important role in the future to optimize planning of therapies."

- CONQUEST is a linear scaling, or O(N), DFT electronic structure code developed jointly by the National Institute for Materials Science (NIMS), Japan, and UCL. The code is designed to perform DFT calculations on very large systems (containing tens of thousands, hundreds of thousands, or even millions of atoms). It can be run at different levels of precision, ranging from ab initio tight binding up to full DFT with plane wave accuracy. CONQUEST is currently used by academics at UCL and Imperial College London. It is also in use on the Jaguar petaflops supercomputer at Oak Ridge National Laboratory and is one of the main codes selected for optimization on the Japanese 10PF "K" supercomputer. As well as linear scaling functionality, CONQUEST can be run with exact diagonalization using ScaLAPACK, and optimizing this part of the code for calculations on thousands of atoms for metallic systems was the target of the project. The code performance has been improved by optimizing ScaLAPACK use and implementing k-point parallelism. Implementation of Kerker and wave-dependent metric preconditioners has enabled calculations to achieve self-consistency that would be otherwise not possible with linear mixing. With the improvements from this work and the already excellent scaling properties, CONQUEST will become a valuable tool for large-scale ab initio electronic structure simulation on metallic materials on HECToR.

- Dr. Matt Probert of the Department of Physics at the University of York said, "[The development of the new geometry optimizer for CASTEP and ONETEP] has been a great success. The 'time to science' — that is, to do a full structural optimization in either code — has been maintained or even reduced in certain circumstances, whilst the amount of memory required has been dramatically reduced. This latter point is key, as with the move to multicore, the amount of memory per core is going down, and the old geometry optimizer consumed a large amount of nondistributed memory and was beginning to limit the size of problem that could be tackled. As a consequence, users of both CASTEP and ONETEP will now be able to tackle significantly larger systems: With the old code, there was an effective maximum size of around 3,000 atoms on HECToR Phase 3, which has now grown to 100,000 atoms. This will enable users of CASTEP to study larger structures, such as grain boundaries, and for ONETEP to tackle larger molecules, such as proteins and DNA segments. Obviously, not all CASTEP and ONETEP users on HECToR will be studying systems of this size, but there is no downside to the new minimizer: All calculations will take the same or less memory [with the exception of very small systems, which cannot be run usefully on HECToR] and converge in the same or less time than before. The new minimizer will be the default algorithm in the next release versions of both CASTEP and ONETEP."

- Dr. Ricardo Grau-Crespo of the Department of Chemistry at UCL said, "Recent progress in the development of the VASP [Vienna Ab-initio Simulation Package] code has made possible a very accurate description of the electronic structure of some types of solids (e.g., transition metal and rare earth compounds) for which traditional methods fail. This opens the opportunity for new research in fields like catalysis and electronics. The remaining limitation is in the computational cost of these calculations, in particular given the need to comply with short job completion times on HECToR. Achieving good scalability in the parallelization of the code is therefore essential in order to take advantage of these new developments. This dCSE work has significantly expanded the scalability of the VASP code in HECToR by exploiting k-point parallelism, and exciting new research in computational materials science will now be possible thanks to this work." In addition, Dr. Aaron Walsh of the Department of Chemistry at the University of Bath commented, "K-point. This will remain the case for HECToR Phase 3 and any other HPC system."

- Dr. Barbara Montanari of the Computational Science and Engineering Department at the Rutherford Appleton Laboratory, STFC, said, "This dCSE project allows CRYSTAL to exploit fully the resources of HECToR and to apply advanced density functionals — such as hybrid exchange — to systems that have previously been impossible to study. The first work to benefit has been in renewable energy where extensive surveys of bulk and surface structures are now under way: Applications include the photovoltaic production of hydrogen, solid oxide fuel cells, and lithium ion batteries."

- The dCSE software development program has been very effective in producing a novel development of CRYSTAL performed in close collaboration with the high-end computing support [...] at the Rutherford Appleton Laboratory, who provide long-term and in-depth support for the CRYSTAL suite of codes.

- Prof. Wyn Williams of the School of Geosciences at the University of Edinburgh said, "Over the last 20 years, micromagnetic modeling has yielded a wealth of information both on the fundamental nature of the physics of magnetic recording processes and as a design tool for developing new forms of magnetic recording technology. Modeling codes are being developed by just a handful of researchers worldwide. It is extremely difficult to get U.K. funding for such initiatives through normal funding, where the focus is on the scientific outputs rather than technical developments. It is also difficult to recruit the skilled staff needed to get the code development done effectively. The dCSE awards have become a vital resource, providing scientific researchers the specialist knowledge and expertise that allow rapid development and coding of the numerical models where the researcher can focus on the scientific goals rather than the details [...] of the parallelization."

- Prof. Williams further commented that there was a need to address two scientific issues using MicroMag and HECToR. The first is the magnetic stability of rock particles, for which "to match experiment, computational models require resolutions in the order of 200 times the current practical limit of MicroMag." The second is the thermal behavior of nonuniformly magnetized particles, where "the availability of the new MicroMag code and the computing resource of HECToR mean that there is now real hope of significant progress on what has eluded analytic theory for over 50 years."

- Professor Kenneth Taylor of the Department of Applied Mathematics and Theoretical Physics at Queen's University Belfast (QUB) said, "dCSE support was essential to the successful development of several very large and complex HPC parallel programs. Each of these projects required months to years of full-time attention by someone who was both a specialist in the science and had experience developing software for parallel processors. In each case, the software packages enabled calculations that cannot to our knowledge be performed elsewhere. A generalized version of HELIUM for crossed and circularly polarized laser fields was developed and tested with dCSE support. We now have the ability to integrate the helium-laser Schrödinger equation in its full generality (essentially six spatial dimensions plus time integration). This ability is unique in the world. The code has been used to search for failures in reduced dimensionality models of atom-laser physics. Full exploitation of this high-dimensional code awaits HECToR's successor for reasonable performance. A particularly successful outcome of the HELIUM dCSE project was the development of postprocessing code that calculates the energy spectrum of double ionization. It has recently been used to provide a high-integrity analysis of an experiment that has been controversial for several years. dCSE support and the full performance of HECToR were both vital to the success of this research."

- The project was focused on the incorporation of two enhancements: A load-balancing scheme, which would improve the efficiency of the code when modeling systems exhibited spatially nonuniform atomic density, and rigid body dynamics, which would be a better treatment of the forces and equations of motion. Commenting on the dCSE project, Dr. Ilian Todorov of STFC's Daresbury Laboratory said, "[This] was a useful project in terms of improving our understanding of the particular load balancer and appreciating its potential and power. However, the level of intricacy coupled with the limited effort available make this development bear more academic value than actual CSE value. It is just too expensive to develop and incorporate into already big code without any recurrent funding. The effort this dCSE project awarded to the [rigid body dynamics] algorithms development was the invaluable time for testing them extensively and learn more about coupling topological objects dynamically and handling their various degrees of freedom with the domain decomposition framework. Overall, this project taught us how grossly we had underestimated the effort needed for these developments."

- General R-matrix methods successfully model multielectron atom-laser and molecule-laser interactions, but only in the time-independent limit. HELIUM, which has been in heavy use on massively parallel machines for over 15 years, successfully models time-dependent atom-laser interactions but is limited to two-electron atoms. RMT removes both of these limitations. Significant performance and load-balancing optimizations to the RMT code make it possible to achieve a 30% speedup using 8,192 cores.

- Commenting on the dCSE project success, Prof. Kenneth Taylor of the Department of Applied Mathematics and Theoretical Physics at QUB said, "dCSE support for software development has underpinned the success of the RMT code (R-matrix incorporating time). RMT is a large complex parallel code that exploits the high accuracy and high efficiency of the finite-difference and time-propagation methods of HELIUM. RMT is capable of handling laser interactions with many-electron systems at a level quite impossible before. HELIUM is a well-established HPC code for calculating such laser interactions with two-electron systems and has been developed in Belfast over the past 20 years."

- "The time-independent R-matrix code (RMATRX) coupled in has again been (largely) developed in Belfast over several decades, and combining the two codes (HELIUM and RMATRX) successfully was a project demanding several years of work by researchers intimately familiar with the scientific domain, the specialized numerical methods, and parallel programming methods. RMT has now been applied with success to ab initio calculation and analysis complementary to recent experimental measurements of attosecond time delays in the photoemission from the outer subshells of neon and to calculations of high harmonic generation in helium. Additional application areas of RMT include the study of inner shell excitations and decays in complex atoms and intense-field atom-laser interactions at XUV frequencies generated by the new free-electron x-ray lasers. Preliminary research in several of these domains is now being moved from small computers to HECToR for large-scale production runs."

- Before this work, both SWT and SS3F used a 1D parallel data decomposition, which becomes restrictive at higher numbers of processors since the decomposition is limited by the number of collocation points in one of the three spatial dimensions. By implementing better data decompositions in both codes, their strong scalability has been substantially improved: SS3F now scales to over 12,000 cores and SWT to 8,192 cores for a representative fixed problem size.

- Commenting on the dCSE project success, Dr. Gary Coleman of the University of Southampton said, "This work allows us to make efficient use of current and future HEC resources and consider problems and regimes that otherwise would have gone unaddressed. CONQUEST is a linear scaling, or O(N), DFT electronic structure code developed jointly by NIMS and UCL. The code is designed to perform DFT calculations on very large systems (a system containing two millions atoms of silicon has been demonstrated). It can be run at different levels of precision, ranging from ab initio tight binding up to full DFT with plane wave accuracy. CONQUEST is currently used by academics at UCL and Imperial. It is also in use on the Jaguar supercomputer at Oak Ridge National Laboratory (ORNL) and is one of the main codes selected for optimization on the Japanese 10PF 'K' supercomputer. Prior to this work, CONQUEST use was limited to systems with relatively simple bonding, where spin and van der Waals interactions were not important. As a result of this project, the code can now be used for weak bonding in biomolecular systems and magnetic systems. Furthermore, as these new application areas require the use of a scalable code and large HPC resource, scalability was also improved. This was achieved by developing hybrid parallelization within CONQUEST via the implementation of OpenMP with the existing MPI."

- Improved simulations in ocean modeling, renewable energy, and geophysical fluid dynamics. Furthermore, Fluidity will be able to be coupled with applications that require nonstandard discretizations such as atmospheric and ice-sheet models.

- Fluidity is a nonhydrostatic finite element/control volume CFD numerical forward model that is used in a number of scientific areas: geodynamics, ocean modeling, renewable energy, and geophysical fluid dynamics. The applications cover a range of scales from laboratory-scale problems through to whole earth mantle simulations. Fluidity is parallelized using MPI; prior to this, work scalability was limited to around 256 cores on HECToR Phase 2b. Fluidity is now capable of scaling to many thousands of processors on HECToR Phase 3. One of the unique features of Fluidity is its ability to adapt the computational mesh to the current simulated state: dynamic adaptive remeshing. On many processor simulations, the adapted mesh must be load balanced via dynamic load balancing. Zoltan is capable of performing very efficient dynamic load balancing, and as a result, its use will extend the original functionality of Fluidity on large HPC systems and thereby increase the range of scientific areas that it can be applied to.

- Commenting on the dCSE project success, Dr. Jon Hill of the Applied Modelling and Computation Group at ICL said, "Previous to this project, Fluidity could only handle a single discretization type in parallel when using adaptivity, limiting the problems that could be solved on large systems such as HECToR. Coupled with research into new discretizations that would need to also be integrated into existing software, the capabilities of Zoltan will allow any current and future discretization to be used. Without this project, significant effort would have to be found for each current and new discretization. This would have been a major software maintainability issue."

- "The dCSE program enabled funding of work with relatively little overhead, which would be almost impossible under other funding streams. The expert software engineer from NAG worked well with our group and quickly got to grips with our testing infrastructure and software development methodologies. The Zoltan code was integrated into the main code branch and is now being actively used by all users of the code. In addition, the functionality of Zoltan is being used in a number of other areas, such as OpenMP-MPI hybrid optimization, as it contains useful functions, such as graph coloring."

- Parallel simulations to study low-frequency turbulence in magnetized plasma generally involve FFT calculations along with associated data copying and MPI communications. This work has optimized the transformation of data between the linear and nonlinear parts of the code. By implementing a more efficient approach for the indirect addressing used for the data copying functionality, performance has been improved. Furthermore, an improved "unbalanced" data decomposition has been developed that provides more flexibility to users by allowing them to select the process count that matches the exact simulation, resources, and system they are using. GS2 can now be used efficiently on HECToR with 2,048 cores, with a representative reduction in overall runtime time of 17–20%.

- Commenting on the dCSE project success, Dr. Colin Roach of the Culham Centre for Fusion Energy said, "The need to redistribute very large 5D data arrays is one of the main challenges limiting the maximum practical number of cores that can be used for gyrokinetic simulations. This collaborative dCSE project has allowed us to drill deeply into the innards of the open source GS2 code and to make impressive improvements in code performance at both low and high core counts. GS2 has been around for some years, and these improvements exceeded my expectations. Furthermore, the project has also exposed ideas of very useful improvements that we could make in the future. The outcomes from this project will be of direct benefit to the wide user base of GS2.

We are now exploiting this enhanced version of the code to make more efficient use of HECToR in scientific projects, where we are exploring the losses of heat and particles due to electron- and ion-driven turbulence in tokamak plasmas."

- Prof. Michael Fagan of the School of Engineering at the University of Hull said, "Bones have extremely complex geometries with multiscale features that can vary by several orders of magnitude. Capturing that detail is essential to mimic and understand fully the biomechanics of bone, but results in huge FE model files. This dCSE project has been invaluable in providing a way to manage those very large files. It opens up the possibility of modelling whole bones at unprecedented resolution but also now provides the option of repeated adaptive remodeling simulations to allow the modelling of bone growth and optimization of the next generation of orthopedic and dental implants."

- Dr. Sylvain Laizet of ICL said, "I am currently undertaking direct numerical simulations with incompact3d. The work is about turbulent plumes in a channel flow configuration with a tilted ramp. In this numerical work, the mixing of fresh water with salty water in a channel flow configuration along with particle settling processes is investigated. In particular, the influence of a tilted ramp inside the computational domain that is modeled using an immersed boundary method is studied in order to reproduce real-life configurations. 3D interpolations are needed to prescribe the correct boundary conditions on the ramp, and it can only be done with the halo-cell communication code newly introduced. The idea is to use an explicit stencil-based method for 3D interpolations to correctly impose boundary conditions for the velocity, for the salinity, and for the particles (while the rest of the code still uses an implicit compact finite difference method). Concerning compact3d, some simulations of a 3D temporal mixing layer at relatively high Reynolds numbers are currently running in the framework of a collaboration between the University of Poitiers in France and Imperial College (PRACE project for acoustic predictions). So far, compact3d has only been run on a single node. However, the new version of the code will allow groundbreaking progress in compressible flows, with configurations very close to experiments. We are also planning to combine an immersed boundary method with compact3d, in the same way as incompact3d, in order to put a solid body inside the computational domain. The aim is to investigate grid-generated turbulence for compressible flows."

- "The work completed in the framework of this project has been highly successful and has had a very significant impact on the widespread usage of Code_Saturne within our research group at the University of Manchester. Substantial improvements have been introduced across a range of items: the overall scalability, I/O performance, parallelization of pre/postprocessing stages, and domain-decomposition tools. All of these features combined have enabled a step change in the scale and depth of our computational turbulence research via more efficient usage of HECToR, which has directly benefitted the work undertaken in a number of other research projects." "Incompact3D is an in-house computational fluid dynamics code used by the Turbulence, Mixing and Flow Control group at Imperial College and its academic collaborators to conduct state-of-the-art turbulence studies. The work follows on from two previously successful dCSE incompact3D projects that significantly improved the scalability of the code, enabling it to use more than 10,000 cores on HECToR for production runs. However, as more cores are used, time spent communicating between nodes becomes more prominent, indicating that there is room for further development. The objective of this project was to reduce the time taken in all-to-all communications, which is particularly important when utilizing large core counts on HECToR. This was accomplished by further developing the 2D decomposition and Fast Fourier Transform library, 2DECOMP&FFT, to perform computations that overlap across-node communications."

- Commenting on the dCSE project success, Dr. Sylvain Laizet of ICL said, "When the recent update of HECToR increased the number of cores per processor to 32, we observed that the communication within a processor could be a limiting factor for running large simulations with incompact3D. With dCSE software support, further development of the code reduced the time taken by the all-to-all communications in incompact3D by about 15%. By using a library solution, 2DECOMP&FFT, there was little impact on end users. Our code incompact3D is now available as an open source code through a Google project with about 75 new users worldwide since November 2012, opening many possible collaborations within the United Kingdom. These three successful dCSE projects related to incompact3D now enable us to perform highly accurate large-scale simulations in conjunction with our experimental results that form a central component of our research effort. As such, the added value is critical."

- CABARET a general-purpose numerical scheme that may be used to solve the complex flow of liquids and gases, is well suited for computational aeronautics and geophysics problems. One application of CABARET is in the investigation of aircraft noise, which is currently a very important environmental concern. A significant component of all aircraft noise results from air flowing around the airframe, engines, and wing flaps. Reducing this noise remains a very challenging design problem that requires this model to accurately simulate the wing-flap, free jet, and wing-flap-jet interaction effects that can occur so they can be minimized. Another CABARET application simulates medium- to large-scale circulation effects of eddys in the ocean, which is a major concern in the study of climate change.

- Commenting on the dCSE project success, Dr. Sergey Karabasov of the School of Engineering and Materials Science at Queen Mary University of London (QMUL) said, "The dCSE project has been mainly (75%) focused on extending the general CABARET method to wider scientific applications. Specifically, thanks to this project, a new computational geophysical fluid dynamics (GFD) code PEQUOD (the Parallel Quasi-Geostrophic Model) has been introduced to HPC for high-resolution ocean modeling simulation at eddy-resolving regimes. This work has already led to secure NERC funding for a new three-year project 'Turbulent Oscillator' (Dr. Pavel Berloff, principal investigator, and Dr. Karabasov, coinvestigator) where the HPC PEQUOD code will be used to study long-scale variability in oceanic flows."

- "The remaining 25% of the current dCSE project has been used for improving I/O of the unstructured CABARET code used for aeroacoustics applications. This work has been essential since the original way of postprocessing results of the

CABARET calculations was largely serial, which was beginning to cause problems on HPC file systems. The dCSE software support has been very useful for code development."

- Dr. Jimena Gorfinkiel of the Department of Physical Sciences at The Open University said, "The dCSE project allowed us to carry out development work that would have been extremely difficult to perform otherwise. By providing the funds and mechanism to hire an experienced developer to work intensively on this specific task (the parallelization of the Hamiltonian construction and diagonalization), it enabled us to carry out this work effectively and to a high standard. The dCSE work will allow several groups in the United Kingdom (at UCL, Open University, and QUB) to perform new science in the very near future. In the short term: Electron collisions with DNA constituents — we will run significantly improved calculations for electron scattering from biologically relevant molecules containing 40+ electrons and 10+ nuclei. This work is essential to understand how low-energy electrons damage DNA; electron collisions with molecular clusters — in order to model the effect of the medium in electron scattering from biological material in the cell, experimental and theoretical studies of collisions with small molecular clusters need to be carried out. Again, the DCSE work will allow us to do this for the cluster of, for example, a DNA base and several water molecules; and positron-molecule collisions — it is essential for these calculations to be accurate that a good description of the polarization of the target is achieved; we do this by including enough pseudostates in the calculations. In the medium term, and together with software development being carried out under the EPSRC UK-RAMP grant, the developments will allow us to describe the multielectron response of polyatomic molecules to intense short laser pulses. The software in development will be world leading in this very active area of research."

- The Vienna Ab-initio Simulation Package is a quantum mechanical software package used to simulate the electronic structure of condensed phase materials. VASP is one of the most important and utilized materials science and chemistry codes in the world. VASP is currently the most heavily used single application on HECToR, consuming 20–25% of the total of the machine resource. During 2011, VASP used 267,588 kAUs on HECToR (12.65% of the overall machine and 17.61% of all jobs), and during 2012, this figure was about 651,061 kAUs (20.4% of the overall machine and 21.1% of all jobs) — equivalent to a notional cost of over £1.5 million. At that time, there were at least 63 registered HECToR users using the package from 14 different HECToR projects.

- We now highlight one specific scientific project that is only possible through improved scaling in hybrid-density functionals. A new class of quaternary Cu-based semiconductors based on the kesterite mineral structure have become the subject of intense interest because they are considered as ideal candidate absorber materials for low-cost thin-film solar cells.

- To further improve the solar conversion efficiencies, it is crucial to understand the basic material and device properties of these kesterite materials. Unfortunately, standard-density functionals applied to CZTS result in an electronic band gap of just 0.09 eV, an error of 94% with respect to experiment. We have recently shown for the bulk system that a hybrid-density functional (HSE06) can reproduce experiment exactly. The performance enhancements of hybrid-density functionals will be used to model the formation of a heterostructure of this material, with the CdS window layer used in real solar cell devices. This will involve the construction of a periodic superlattice of the order of a few nm, containing hundreds of atoms, and simulating the charge distribution occurring at the interface, which determines the separation of electrons and holes when light is absorbed. Having an accurate description of the electronic states is essential to understanding the interfacial electronic structure and for extracting the key physical parameters associated with the heterojunction, which can be used to model and improve photovoltaic device performance. Ultimately, these simulations, which explicitly rely on the use of hybrid functionals, could lead to improved light to electricity conversion efficiencies in next-generation devices.

- Nucleus for European Modelling of the Ocean (NEMO) is an ocean modeling code of great strategic importance for the United Kingdom and European oceanographic communities. Although NEMO has been used successfully for a number of years in global and ocean basin applications, its use as a shelf-sea model is less well developed. Also, NEMO was originally designed for vector architectures, leaving room for improvement in its performance on massively parallel architectures such as HECToR. This work focused on shallow-sea problems and the performance and scalability of the code over many thousands of cores on modern architectures by introducing more flexible domain-decomposition techniques and permuting array indices to improve halo exchange performance and enable the elimination of redundant computations on land-only cells.

- Commenting on the dCSE project success, Dr. Stephen Pickles of the Advanced Research Computing Group at STFC's Daresbury Laboratory said, "The dCSE NEMO project has improved the usability of the NEMO ocean modeling code by introducing dynamic memory allocation so that the code no longer needs to be rebuilt to run on a different number of processors. It has also provided essential data to inform NERC's long-term strategy for ocean modeling within the United Kingdom through the NERC Ocean Roadmap Project. The dCSE program as a whole is extremely valuable to the U.K. computational science community. The continued competitiveness of scientific modelling codes depends on sustained development effort. The dCSE program fills an important gap, helping to maintain capability and expertise in high-performance computing and software engineering and, through this, to sustain the software assets that U.K. scientists need in order to remain competitive on the world stage."

- Block 32 is one of the world's most difficult petroleum objects to image because two-thirds of the block surface is covered by salt, and salt bodies considerably disrupt wave propagation. So there was no choice but innovation! For that, the best industrial algorithm has been developed in FWI [full waveform inversion], with impressive increased resolution of the velocity model thanks to combining methods, mathematics, models, and HPC. The result was fast processing and accurate imaging, resulting in 12 significant discoveries out of 16 exploration wells. This was performed for a study done for Exxon and resulted in $20 million just for the study. In 2003, a seismic analysis of a region in the Gulf of Mexico on a 64-processor cluster showed a structure shaped like a bowler hat which is typical for a petroleum zone. Based on this image, TOTAL was ready to install boring equipment on this side. Luckily, sometime later, a fresh data analysis on a 10,000-core supercomputer revealed that the structure was an artifact. Thanks to better HPC, TOTAL saved $80 million. On average, only 30% of an oil field is deployed, the other 70% of oil remain in the oil field because it's simply too difficult for exploration. Additional EOR (enhanced oil recovery) mechanisms requiring thermohydraulic modeling and chemical reaction simulation are essential for new fractured reservoirs, shale oil, and shale gas. Only a 10% increase of oil deployment due to these methods would result in earning of billions of dollars. This is achieved by both improved reservoir modeling through enhanced HPC and EOR.

- Multiscale CFD, from molecule to cell to component to the environment. Fluid catalytic cracking (FCC) is one of the most important conversion processes used in petroleum refineries. It is widely used to convert the high-boiling, high-molecular-weight hydrocarbon fractions of petroleum crude oils to more valuable gasoline, olefinic gases, and other products. Accomplishment is running on 10,000s of cores to simulate 3D industrial Riser for up to 3 million cells on bubble/laminar/turbulent regimes. When explosion is included in the simulation, one is able to understand the mechanisms controlling the explosion, typically in offshore oil and gas producing platforms, to cope with the societal, economical, and physical challenges of prime importance.

- Recent 22nm microprocessor technology brought innovation and large economic returns. In 2011, total investments were $100 million, with an HPC investment of just 70TFLOPS for molecular modeling simulations.

- Beijing Computing Center (BCC) has been providing cloud computing services to small and medium-sized organizations to run their simulation applications since 2009. At present, a high-performance computing platform has been established, with 600 processor cores (50 nodes). A variety of simulation software, including Abaqus, LS-DYNA, ANSYS (Fluent), and HyperWorks, are running on the platform to provide CAE consulting service for small and medium-sized enterprises. The industries benefited from the BCC cloud service range from automobile, mechanics, and civil engineering to construction. BCC's HPC cloud service has helped those small and medium-sized firms accelerate their CAE simulation application to develop better products and bring out more innovations. The total investment for this project is $1.36 million, made up of $560,000 on software and hardware, $320,000 on power, and $480,000 on personnel costs over the past four years. BCC is the first and only computing center in China dedicated to helping small and medium-sized organizations to gain access to supercomputer systems and better compete in their market space.

- Shanghai Jiao Tong University (SJTU) is one of few top universities in China with a strong academic reputation in science, engineering, and medicine. To meet increasingly high demand for computing resources from scientists and researchers in recent years, we built the fastest supercomputer among all the universities under China's Ministry of Education (MOE) in April 2013. It is a heterogeneous supercomputer equipped with both CPUs and accelerators, with Rpeak at 260TFLOPS.

- Working closely with the CFD team in the Commercial Aircraft Corporation of China (COMAC), SJTU has developed an in-house code called SJTU-NS3D to solve 3D RANS (Reynolds-averaged Navier–Stokes) equations on structured grids by FVM (finite volume method), which has been widely used in designing wing models. To meet the rising demand for computation resources for higher precision, we have designed the CUDA version of SJTU-NS3D. As the hot spot, the Runge–Kutta iteration has been offloaded to the GPU as kernel functions. Then we have further optimized the algorithm by using four different approaches: for better locality, transformed 1D matrices stored intermediate variables to 3D matrices; to update data in neighbor vertexes, used multiple kernels to synchronize globally; to eliminate data transfer between host and device memory, paralleled one direction each time in implicit residual smoothing algorithm; for better access to global memory after matrices transformation, used intrinsic function cudaMalloc() to coalesce access. The optimized CUDA version has achieved 20-fold speedup for the ORENA M6 wing model compared with the single thread on Intel i7 920@2.67GHz, without sacrificing any accuracy, and 37-fold speedup for a real wing model candidate from COMAC on a single Fermi C2050. After testing with other real cases, the CFD team in COMAC has adopted the code in daily work. A postdoctoral has been employed recently to further optimize the code to adopt higher-accuracy schemes in convective terms adaptation besides Jameson Central Difference and expand to multiple Fermi cards to meet larger grid numbers. Interactions of intense laser pulses with plasmas have very broad applications such as fusion energy generation, high-energy electron and ion acceleration, and powerful terahertz and x-ray radiation generation. These interactions can be studied significantly by the particle-in-cell (PIC) code, which is a well-established first-principle model: electromagnetic fields by solving Maxwell's equations on a grid using finite difference, and particle motion by solving the relativistic Newton's equation. Due to the first-principle nature, PIC simulations require intensive computation, and PIC codes have long been at the frontier of high-performance computing. In particular, a huge number of collision times are also involved in fusion energy generation, which needs much more intensive computation and cannot be achieved easily by usual parallel CPU computer clusters. A preliminary result shows a speedup of near 30 times when GPUs are employed.

- Application in medicine is in ion transport and selectivity with applications in drug design and disease study, with the Institute of Natural Sciences, SJTU. Ion transport through ion channels is an application developed for understanding many diseases and providing insights for new drug designs. The physics of the transport process involves electrostatics interaction and ion-ion correlation and beyond — Poisson-Boltzmann electrolyte theory, protein dynamics and interactions, and phase transition theory. This process again involves physical phenomena of many time and spatial scales that require multiscale computing. The large disparities in time and spatial scales can be seen from the ion flow scale in 0.1ms — 1-minute vis molecular bond vibrations fs and the spatial ion channel 0.7nm vis membrane scale in microns. "Novel computing technologies are emerging. The best example is supercomputing infrastructure, which provides massive computing power, significant storage capacity, and interconnection bandwidth. These fundamental advantages take genomics research to a new attitude in terms of improving analysis throughput by orders of magnitude and enabling challenging computing-intensive research that was impossible in the past. These new capabilities bring cost down in application scenarios like clinical diagnostics and treatment as well as provide new insight on disease, even a fundamental understanding of living.

- BGI has been collaborating with the National Supercomputing Center in Tianjin (Tianhe-1A), China, for quite a time to use the most cutting-edge supercomputer systems to enable basic research and boost business development. Case 1: Integrating GPU computing power and the Hadoop distributed processing framework, the team is able to conduct metagenomics research with just one-fortieth of the original hours. As we know, gut bacteria could play key role in the development of various diseases such as type 2 diabetes. To investigate this phenomenon, the various species in the gut need to be clustered and analyzed according to sample characteristics; thus each of them can be identified and modeled. The development of this package enables research over thousands of samples, which is promising to obtain a clear picture of the gut bacteria–to-disease relationship. Case 2: Running on Tianhe-1A supercomputing systems, the *de novo* assembly of human genome from second-generation short reads takes only two hours. This was made possible by the state-of-the-art parallel algorithms and leading systems architecture of Tianhe-1A. With fuller genome assembly, we can more exactly inspect the genetic variation landscape. This is of great importance for understanding and fighting disease as well as for better healthcare. Case 3: The computing capacity of Tianhe-1A helps the healthcare business, which is time critical. The best example is prenatal diagnostics, which can help prevent birth defects. Compared with the traditional approach, sequencing-based diagnosis is almost 100% accurate and risk free. The software infrastructure ensures data transfer, analysis execution, result generation, and report delivery. The analysis throughput had never been an issue on Tianhe-1A, and the elastic design of the framework reduces time to solution from days to just hours, which makes the business very competitive.

- In recent years, the Shanghai Supercomputer Center (SSC) got more and more industrial computing simulation requirements from industries such as aviation, marine, automobile, steel, nuclear power, semiconductor, and rendering. SSC's old system is very limited and not enough to fulfill these requirements. So SSC built a new industrial computing environment — the Hummingbird Cluster System — in 2012 to provide more computing resource for industrial or business computing. This system is made up of the IBM blade cluster system and DDN storage system and only provided for industrial users. In 2009, the Shanghai Supercomputer Center launched its third-generation supercomputer, Magic Cube Cluster system, which was produced by Dawning. This machine has 200TFLOPS computing power and was ranked number 10 in the world and number 1 in Asia in ISC09. Until now, over 500 scientific research groups have used or are using this system and have produced over 1,000 research papers or reports and finished over 1,000 research projects.

- The parallel file system Cappella, a distributed file system, is specially designed for the needs of the China Center for Resources Satellite Data and Applications (CRESDA). The system of CRESDA receives, processes, distributes, and archives satellite data, and it is expected to produce 12TB of new data received each day in less than 60 minutes (normal mode) or 10 minutes (express mode). Typically, the final production data are small files, about 100KB. That means the system requires 10,000 to 100,000 metadata operations per second (ops/s), which will increase sharply when the system scales up to process data with higher resolution. For traditional parallel file systems with a single metadata server (MDS), this will be a challenge for the MDS and is prone to be a bottleneck. Besides, it requires at least 50GBps aggregated I/O throughput at peak times to support all the possible operations. Cappella holds several unique characteristics, the design of which is motivated by the observations of specific CRESDA application workloads and technological requirements (e.g., the volume of ultralarge file systems, fast response to huge metadata requests, and high aggregated I/O throughput). Cappella has a scalable and adaptive metadata management structure, called G-HBA, and holds a novel, decentralized, semantic-aware metadata organization to store metadata and support fast metadata lookups in an ultralarge-scale file system with multiple metadata servers. Experiments show that it has similar aggregated I/O throughput as Lustre but higher metadata throughput to meet the demands of CRESDA. When using Cappella, the application demonstrated 2 times performance improvement. The key ideas in the metadata design were cited by papers in HPCC, FAST, SC and so on.

- In the past 10 years, the new generation of high energy physics (HEP) experiments, such as the experiments on LHC (Large Hadron Collider at CERN) and the BESIII experiment on BEPC (Beijing Electron–Positron Collider), has been creating a huge amount of data. Handling the gigantic amount data generated from the large-scale experiments is a great challenge in scientific computing. At the Institute of High Energy Physics (IHEP), a self-made data-intensive computing environment was designed and established to support international and domestic high energy physics projects. The computing environment includes an HPC cluster, high-throughput storage, and a high-bandwidth network that links to the United States and Europe. Our system is integrated with the U.S. and European grid computing systems, together forming a globally distributed computing grid. The computing environment provides a new infrastructure for data process and physics analysis. The speed of data processing increased three orders (10^3) compared with that on the previous system. This makes scientific computing possible for new HEP experiments. More than 20 million CPU hours of computing service were provided and many petabytes of data were processed every year. Our computing center has been making critical contributions to physics researches, such as the Higgs discovery and new neutrino phenomena.

- With seismic field high-density geometry acquisition technique development and application, 3D seismic data volumes are getting bigger (a single file may exceed 50TB–100TB; some are over 200TB), driving the demand for more optimized HPC configurations. We have installed a lot of systems in the past three years. The following are three examples: 148 nodes (Sandy Bridge E5-2670), 296

CPUs, 2,368 cores, and 3000TB storage, and finished 4 Big Data 3D survey projects (2004KM2, 144TB, $3.22 million) from January 2013 to May 2013 (innovation points include big data processing, seismic application software I/O optimizing, new processing technology, 10GbE networking usage, and parallel storage cluster); 558 CPUs, 504 GPUs (Sandy Bridge E5-2660, M2090), a parallel processing depth image system, and 2,000TB storage, and finished 32 3D survey processing projects (28656KM2, 3.871TB, $25.75 million) from June 2012 to December 2012 (innovation points include big data processing, seismic application software GPU parallel optimizing, IB networking usage, and parallel storage cluster); and 256 CPUs (Westmere, 2.80Ghz), PC cluster system upgrading (initial installation was in 2011), and 3,000TB storage in January 2013, and finished 4 Big Data 3D survey processing projects (39TB, 1160KM2, $3 million) from January 2013 to May 2013 (the upgrading project has saved $960,000 in cost) (innovation points include big data processing, seismic application software I/O optimizing, 10GbE networking usage, parallel storage cluster, and upgraded memory, disk, and networking).

- Macromolecular structures and functions are the current challenges in molecular biology. Without HPC resources, it is difficult to solve the 3D structures of macromolecular complexes and simulate their dynamics. We utilized our HPC facilities supported by Inspur to perform structural studies of many important biological macromolecular complexes (including virus, chaperonin, and filaments) and got great achievements. Our research achievements are based on massive computations and have improved our understanding of life in molecular details.

- By employing HPC simulation, the company was able to simulate and analyze their pallets in a highly predictive and time-efficient manner. Without these HPC resources, they would not have earned a multiyear contract from a large German automotive OEM.

- Improvements to Jeco's pallet product have impacted their bottom line as sales revenue is expected to double, payroll will increase by 35% at their plant, and they will be in contention for additional high-margin domestic and export business projects. From an MS&A perspective, NDEMC facilitated Jeco's access to SIMULIA Abaqus Unified FEA, which ordinarily would have been beyond the realm of possibility due to budgetary constraints. By gaining access to MS&A and technical expertise, Jeco had the ability to develop creative technological solutions in the final, time-critical phase of the product innovation process.

- As recently as eight years ago, the application of high-performance computing techniques to drug discovery efforts was problematic at best. Using the best artificial intelligence platforms available at the time, even clusters composed of 40 or 50 processors could take up to 12 months to run through the DNA sequence data and the corresponding gene expression and clinical response data needed to identify the important genes in a tumor when compared with normal tissue. Today, due to advances in supercomputing and software platforms from companies like GNS with its REFSTM computational environment, Wolfram Research with Mathematica, and MathWorks with MATLAB, results of this type can now be achieved in weeks.

- 10 years ago, using the high-powered workstations available at the time, Dana engineers were only able to model and simulate the performance of one engine cylinder operating at room temperature. Also, cylinder head gaskets are composed of multiple layers but, in early simulations, could be represented by just a combined single layer, using so-called gasket elements. And even these crude simulations would take several days to run. Much of a product's testing and analysis at this time relied on complex, time-consuming, and costly physical prototyping. A typical example is the design of cylinder head gaskets — 90% of all new releases are developed using CAE to create the final design and obtain customer sign-off. The old trial-and-error technique of building and testing multiple physical prototypes is a thing of the past. Just a few examples of key initiatives that rely on the company's extensive HPC computational resources include fuel economy and Ford's EcoBoost engine technology; safety, always a prime attribute at Ford; and internal cabin noise, a major factor in consumer satisfaction. HPC and CAE played a pivotal role in the development of Ford's EcoBoost engine technology, which will be available in more than 80% of Ford vehicles by 2013. Today, 3.5L V6 EcoBoost engines are available in the Lincoln MKS, Lincoln MKT, Ford Taurus SHO, and Ford Flex. The new 2011 MY Ford Edge is available with the new 2.0L 4-cylinder EcoBoost engine. That same engine will also go into the next-generation Ford Explorer. The EcoBoost family of 4-cylinder and 6-cylinder engines uses turbocharging and direct injection technology to achieve a 15–20% increase in fuel economy when compared with other midsize utility vehicles. Ford's Powertrain team used HPC technology along with computational fluid dynamics and CAE applications to optimize the design of the EcoBoost. In particular, the engineers worked on optimizing combustion and structural aspects of the EcoBoost Powertrain technologies. HPC resources are also used to develop both passive and active safety attributes. Passive safety focuses on improving structural performance and airbag deployment to reduce intrusion into the vehicle and help protect the occupants. Ford's active safety initiatives include Adaptive Cruise Control and Collision Warning with Brake Support, which uses radar to detect moving vehicles directly ahead. When the danger of a collision is detected, the system warns the driver, automatically precharges brakes, and engages a brake-assist feature that helps drivers quickly reach maximum braking once the brakes are engaged. The technology was introduced in the summer of 2009 on the 2010 Ford Taurus, Lincoln MKS sedan, and Lincoln MKT crossover and will be made available on other Ford vehicles. The entire vehicle is modeled to assess both the active and the passive safety designs. Engineers simulate the results of crashes based on a wide variety of design and environmental factors without actually building a physical prototype for testing. HPC is also used to model what is known internally as "NVH" — noise, vibration and harshness. Controlling interior noise in the automobile is a major factor in customer satisfaction.

- How surfactants — the dominant ingredient in soaps, detergents, lotions, and shampoos — behave at the molecular level can have a profound impact on a product's properties and, ultimately, the product's acceptance by consumers. But new imperatives centered on environmental sustainability have researchers pushing their investigatory limits. With unprecedented oil prices, limited water in third-world countries, and a global push for products that are more environmentally friendly, Procter & Gamble (P&G) must rethink many of the products that consumers have come to depend on around the world. The scale of the research challenge is experimentally overwhelming. To address this, P&G has invested in high-performance computing, but this too presents sophisticated challenges in both theoretical and computational capabilities. They produce many of the popular brands of cleaning products that are in homes around the world, including Tide, Cascade, Dawn, Cheer, Pantene, Downy, Head and Shoulders, and Oil of Olay. The research team has already been able to produce virtual (i.e., computer-generated) "phase diagrams" that examine the behavior of materials at various concentrations. They can now predict at what concentration these materials will make a phase change into a different structure.

- This information is then passed on to the P&G formulators who mix and match the ingredients to create a new or improved product. Which ingredients they choose determines how well the product will perform (e.g., amount of sudsing, shelf life, color, the ability to remove stains, and a myriad of other characteristics). Formulators also must balance the cost and availability of ingredients with the wants and desires of the consumer. Dr. Kelly Anderson's access to HPC at Argonne allows him to simulate and capture information that the formulators cannot access using physical experiments. This data helps make their work that much easier and more productive.

- One of the areas Boeing researchers have modeled using ORNL's supercomputers is aeroelasticity — that is, the effect of aerodynamic loads on airplane structures. Doug Ball explains that when a wing is subjected to lift, loading causes the wing to deform — it tends to bend upward and twist. To get the best possible performance from the airplane, engineers use the wings' aeroelasticity to obtain maximum wing loading — this provides optimal lift and reduces drag that, in turn, lowers fuel consumption. This is static aeroelasticity. The other physical condition under investigation is dynamic aeroelasticity, which occurs when the wing is subjected to sudden forces such as a wind gust. Designing the wing to handle oscillations induced by these forces makes for a smoother and safer ride. "Of course, one way to solve the problem is to build a big, beefy, superstrong wing," Ball says. "But a heavy wing is bad design — what we're really trying to do is design a wing with minimal weight that meets our design criteria. For example, one approach is to add more structure only in those places that are taking the brunt of the static and dynamic aeroelasticity forces. In addition, we are advancing the state of the art in the use of composites. The Dreamliner wing, the first Boeing product to make significant use of composites, is lighter and safer than its metal predecessors, and its greater efficiency results in reduced fuel consumption and emissions. Understanding how these new materials can be used in future wing design is a major goal of Ball's group. Here's how Ball summarizes the major benefits of HPC: "It lets engineers design better airplanes with fewer resources, in less time, and with far less physical simulation based on wind tunnel testing. For example, when we were designing the 767 back in the 1980s, we built and tested about 77 wings. By using supercomputers to simulate the properties of the wings on recent models such as the 787 and the 747-8, we only had to design seven wings, a tremendous savings in time and cost, especially since the price tag for wind tunnel testing has skyrocketed over the past 25 years. This is all the more impressive when you consider the fact that the new wings have been completely redesigned to take advantage of our latest research into aeroelasticity and advanced composite materials. So, while the amount of wind tunnel testing has decreased by about 50%, the characteristics of our current computer-based testing is dramatically different." They are now able to look more closely at various design conditions, failure modes, and loading models and are able build a much larger database. "For example, back in the days of the Boeing 707, we could handle six or seven load cases. Today, we have designed thousands of load cases to simulate on the supercomputer."

- Movie rendering is getting more sophisticated than before and needs tremendous computing power to complete in the timely fashion. At the China Film Corporation postproduction branch, we are doing rendering on 200 compute nodes 24 x 7. With the latest HPC gears, not only can we run more sophisticated models, but they can be finished within the same time window as we did with the previous systems. Overall, we see a 3 times' efficiency gain compared with the previous operation when fewer HPC systems were available. The more sophisticated models deliver higher picture fidelity, which in turn leads to better audience experience.

- To support the increased complexity of our PAM-CRASH model, we have increased our compute capability from 400 CPU cores to 660 CPU cores within a year. Our center can now run more sophisticated models on the clusters within the same time envelope and many more scenarios than we did before.

- Supporting 19 different colleges and research departments in Tsinghua University with a 700-node cluster enabled researchers and students to do more and better science across various disciplines. The areas of research include material science, cosmic modeling, weather prediction, and chemistry research. Over 80 papers were published based on the results from the HPC cluster; some were chosen by the top publications in their field such as *Nature* magazine. The cluster is now running at full capacity after a year, and we see continued demand for more computing resources.

- 48 frames or images have to be computationally created for every second of stereo animation, a feat that demands HPC capabilities. Every frame requires that the camera's moves be designed and every character, effect, or prop on screen to be animated. An animator painstakingly crafts the animation of every character, creating only seconds of a single character's animation every week.

- While this process requires significant amounts of computation, it is the next step in the filmmaking process called "rendering" that requires the enormous amounts of computing power. Rendering is the process of adding light and color and involves sophisticated algorithms to create the various lighting effects. The rendering artists, called lighters, use HPC to set and control the lighting effects in a scene. These lighting effects can include both direct and indirect lighting, bounce lighting, reflections, and shadowing. DreamWorks Animation's HPC clusters are production systems dedicated to creating those two or three CG feature films every year. Exploring a new, highly scalable rendering solution in-house would take a substantial amount of the cluster's processing time. And even if time were available on the DreamWorks clusters, they lacked the necessary configuration and raw horsepower to test their ideas. The DreamWorks Animation team applied for and was awarded time on a leadership-class HPC system at the Oak Ridge National Laboratory. The award was made through the Department of Energy's INCITE Program. INCITE provides access to computing time and expertise at some of the world's most powerful supercomputer centers at ORNL, Argonne National Laboratory, Lawrence Berkeley National Laboratory, and Pacific Northwest National Laboratory. The program supports innovative, large-scale projects that enable scientific advancements. The exploration of DreamWorks' vision to create dramatically faster, more powerful rendering software met the INCITE criteria, and precious time on the ORNL supercomputer was made available.

- In 2003 and 2004, the Goodyear Tire and Rubber Company found itself in a definite slump, suffering declining revenue and losing out to its two main competitors, Michelin and Bridgestone. In response, Goodyear leveraged its high-performance computer clusters and its ongoing collaborative relationship with the Sandia National Laboratories to change the way it developed tires. Rather than designing, building, and testing physical prototypes, Goodyear engineers used modeling and simulation to test virtual models and significantly cut time to market. The result was the Assurance all-weather tire featuring TripleTred Technology, a huge hit that helped Goodyear not only climb out of the hole it was in but continue on to launch a flurry of new tires that resulted in record profits. Using the software developed in collaboration with Sandia, the Goodyear engineers working on the TripleTred Technology were able to create more complex simulations. This allowed them to realize maximum performance from their HPC clusters and significantly speed up their initial modeling process. One of the major factors that allowed the company to turn its revenue sheet around was that the Sandia-Goodyear association allowed Goodyear to innovate more rapidly than its competition. The introduction of the Assurance tire is a prime example of that innovation, as is the subsequent rapid-fire introductions of new additions to its Fortera and Wrangler light truck and SUV tires and the new Eagle with its ResponsEdge and carbon fiber technology.

- New club designs may be developed on a computer screen, but the initial inspiration may be a sketch on paper or the back of an envelope. Using the supercomputer, the engineers translate these initial ideas into a three-dimensional matrix that simulates how the proposed design will perform in real life. The system can simulate what happens to the club and the golf ball when the two collide, how the components shift during the stroke, and what happens if different materials are used in the club head and shaft. The HPC system provides a 3D animated simulation of the design that allows the engineers to see what is happening inside the club as well on the outside. By examining cross-sections of the club during the various phases of its collision with the golf ball, the PING engineers can spot potential problems and correct them in virtual space rather than producing multiple expensive physical prototypes. Simulations that formerly took 10 hours can now be accomplished in 15–20 minutes. The series of tests that previously took months are completed in under a week. PING has accelerated its time to market with new products by an order of magnitude — an important consideration for a company that derives 85% of its income from new offerings. And because they no longer have to go through the expensive and time-consuming cycle of building multiple physical prototypes, the engineers can incorporate top-quality design into the new clubs from the very first stages of the production cycle.

- High-performance computing is enabling researchers at Pioneer Hi-Bred, a DuPont business, to conduct leading-edge research into plant genetics to create improved varieties of seeds. Pioneer researchers use HPC to manage and analyze massive amounts of molecular, plant, environmental, and farm management data, allowing them to make product development decisions much faster than by using traditional experiments and testing alone. HPC also gives Pioneer a window into the future, allowing them to make more informed decisions about their applied breeding programs and frontload experiments with predicted potential winners.

- For Pioneer, the result has been faster improvement in new seed products, staying ahead of the competition, a major jump in innovation and productivity, and the ability to help meet some of the world's most pressing demands regarding the availability of food, feed, fuel, and materials. Pioneer performs exhaustive testing of a large number of genetic combinations in thousands of experiments each year. When they decide to commercialize a new corn hybrid, they want to know if the behavior they have seen in their experiments is representative of the environmental and land management conditions the product will encounter in the future. If the corn is going to be used for biofuels, the combination of the big three — genetics, environment, and management — should yield corn that has a higher extractable starch content, the ingredient used to make ethanol. Pioneer uses its HPC capabilities in two primary ways. On the one hand, the clusters serve as a high-powered batch-processing system that crunches huge amounts of data and then applies statistical analyses to this data. On the other hand, the same HPC cluster configuration is used for modeling and simulation of the breeding process. This allows Pioneer to determine which direction is likely to bring the most positive results in terms of genetic improvement. Software visualization capabilities allow the researchers to work more easily with massive amounts of data. The results of both approaches — number crunching and modeling/simulation — have allowed the researchers to engage in deep theoretical work to study the genetic architecture of many traits of corn that is now being put into practice.

- Virtual product development using in-house supercomputing capabilities has allowed Whirlpool Corp. to take a systems approach to designing not only its broad and varied line of appliances but also the packaging that protects the products during shipping. High-performance computing at the company is operating in a full-fledged production environment, delivering significant payback. Top management regards Whirlpool's supercomputing capabilities as key strategic assets: The packed products may undergo a number of indignities as they make their way from the manufacturer to the distribution center, then on to the retailer, and finally into the customer's home. They can be dropped, squeezed, stacked, crushed, vibrated, shocked, and/or subjected to extremes of heat, cold, and humidity. For example, a forklift — known as a clamp truck — can close its metal jaws around a boxed appliance with 2,000lb of pressure. An appliance can be clamped and moved up to 29 times as it wends its way through the supply chain to its final destination. These products can cost thousands of dollars. So if an appliance or even its packaging is damaged in transit, the retailer or customer may refuse to accept delivery. So in 2004, Whirlpool first dipped its toe into the supercomputer waters with the purchase of a small HPC "cluster" system composed of 48 processors. Within a year and a half, this cluster was overwhelmed. The company then invested in a more powerful high-performance system built around 168 processors. Today, this machine is logging 90% utilization and running 24 x 7 to meet the company's modeling and simulation needs.

- Using HPC, Whirlpool models product designs computationally in "virtual space," exploring a wide variety of permutations and conducting tests using simulation. When designers finally do build the physical prototypes, they use them for validation, not development. HPC allows Whirlpool engineers to run more complex simulations and do them faster. In addition, engineers are now beginning to run stochastic models in order to better predict how a product or material will function in the real world and optimize the design accordingly. Stochastic modeling, also known as the Monte Carlo method, tests sensitivity to multiple variations of a single characteristic, such as temperature or pressure, and creates a probability distribution of that sensitivity. For example, a stochastic model may test the durability of packaging materials to temperatures ranging from subzero to 100F, in 1-degree increments, and forklift pressure from 1,000–2,000lb in 1lb increments. Such detailed modeling allows engineers to identify critical characteristics that can lead to more robust and salable products.

- Medrad, a provider of drug delivery systems, MRI imaging accessories, and catheters, had purchased patents for a promising interventional catheter device to mechanically remove blood clots associated with a stroke. Breaking with a long tradition of building numerous physical prototypes to research the potential of a new technology, Medrad turned to the NSF-funded Pittsburgh Supercomputing Center (PSC), experts at Carnegie Mellon University, and the use of complex numerical simulations running on high-performance computers to determine if the catheter technology was worth pursuing. It was. The classic approach to researching a potential biomedical product involves making benchtop models, subjecting each one to a variety of trial conditions and then moving into animal and human testing. But this approach could not efficiently capture the complicated interactions between the blood cells, the vessel walls, the clot, and the device itself. Dr. John Kalafut felt that the catheter project was an excellent opportunity to take a new approach: replace the traditional build-and-test process with numerical simulations, both to establish if the catheter would really work and then to examine a variety of parameters to determine the optimal design solution. "Not only did we need to understand the physics of how the device worked, we also wanted to explore different design and manufacturing approaches," says Kalafut. "We felt that doing this computationally would be more efficient and faster than building lots of different physical prototypes."

- The physics involved with a blocked circulatory system and the interventional catheter device includes complex fluid dynamics that can be represented mathematically — so-called computational fluid dynamics. In particular, the group wanted to use the powerful algorithms associated with CFD to establish if the catheter would break up the blood clots as claimed. However, the R&D group's high-end workstations lacked the horsepower to conduct the complex simulations. They also did not have the in-house expertise required to develop the detailed CFD codes needed for this kind of investigation. They needed access to HPC and software and the expertise to help them harness their full potential. Medrad used HPC at two critical points in their research and development. Initially, it was employed to investigate the physics of the catheter intervention device in relation to the blood vessels (vasculature) where a clot had to be broken up or removed. In other words, Medrad used HPC to simulate the process of the catheter destroying the clots, adjusting the parameters again and again to ensure that the phenomenon was repeatable. This validated that the science behind the patent's theory was solid and that the device would do what its inventors claimed. Then HPC was used to mathematically refine the prototypes by simulating many different combinations of changes — more than could be done physically in the time frame or budget available — to arrive at the best design. "Using the PSC supercomputer," John Kalafut says, "we have been able to look at multiple iterations of different design parameters without building numerous expensive prototypes."

- Building on the success of the supercomputer CFD studies, a team from the research group and one of the business units now has demonstrated the viability of the device in repeated animal studies. As a result, Medrad is confident that they can manufacture the device.

- To improve the freshness of its Folgers coffee and reduce the costs of packaging, researchers at Procter & Gamble decided to switch from a metal can to a plastic container. However, they ran into a number of problems related to gas buildup inside the containers, fluctuations caused by changes in atmospheric pressure during shipment, and a problem associated with the coffee cans imploding while being trucked to their destination. This latter problem almost caused the entire research project to fail. The group used high-performance computing to both solve their problems in a timely fashion and provide the company with a competitive advantage. "For coffee to be fresh when a consumer opens a coffee can, you have to roast it, grind it, and package it very quickly," Tom Lange explains. "But even after we seal the container at the factory, the coffee continues to give off gas. This gas buildup exerts pressure against the coffee container and potentially could cause it to explode. A metal container can handle the pressure, but there's a problem. After the can is opened, oxygen, the oil in the coffee, and the metal all interact, and the grind quickly becomes stale — even if you put a plastic lid on it." The packaging designers found they could solve the freshness problem by packaging the coffee in a plastic container. This had several advantages — not only were the plastic cans less expensive, but they kept the coffee fresher. Some of the oil is absorbed into the plastic and creates what is called an aroma headspace over the top of the coffee. This prevents loss of flavor after the can is opened, and there is no metal to act as a catalyst and cause staling. Supercomputers were used throughout the process to perform modeling and simulation of the coffee container — primarily finite element analysis, a computer simulation technique used in engineering.

- HPC helped the researchers devise three inventions that met the challenges associated with pressure and staleness. Two of the inventions addressed anticipated issues.

- To advance the state of the art of its proprietary photochromic technology used in the Transitions eyewear and speed its time to market, PPG Industries enlisted the help of high-performance computing. Advancing photochromic technology, which allows lenses to change from clear to dark and block harmful ultraviolet rays, involves complex modeling and simulation of molecules at the atomic and quantum levels. By enlisting the help of the NSF-funded Pittsburgh Supercomputing Center and its powerful high-performance computing capabilities, PPG's R&D organization was able to rapidly create the next generation of photochromic dyes and move out in front of its competition. PPG became a member of PSC's industrial affiliates program and began to access the center's high-end supercomputer and the expertise of the center's personnel.

- The PSC system, known as "Big Ben," with its thousands of high-speed processors, has an order of magnitude more capacity compared with PPG's in-house systems. Using parallel processing, PSC's HPC machine can run massive jobs that are simply beyond the reach of resources available within PPG.

- Dr. Michael Makowski estimates that a complex calculation that might have taken a week on their in-house machines — if it could be run at all — now takes only four to eight hours on the center's machine. This permits PPG to predict many of the performance characteristics of molecular structures and how they will behave under a variety of conditions, without having to construct a physical prototype. "Using the PSC supercomputer, we can computationally screen a whole series or family of new molecular structures proposed by our organic chemists and weed out the 80% that will ultimately fail when they are tested experimentally. This allows us to focus our physical testing on the candidates that have the highest probability of success and those whose behavior we can predictably understand," Makowski says. Makowski says that the business case for using HPC has been made without a doubt. PPG's work with PSC has allowed the company to accelerate the R&D process, which translates into faster product development and more new products. The HPC work on the PSC supercomputer allowed PPG to bring the Gen 5 technology to market in advance of its competitors and has allowed PPG to continue to gain market share and increase its sales and earnings. Ongoing work now continues, and PPG's sixth-generation product is slated for launch within the next year. By using the PSC system to perform sophisticated modeling, the PPG research team is able to gather information that they would not be able to obtain experimentally. The research also adds to PPG's fundamental understanding of the science and physics of its photochromic materials and allows the researchers to continually fine-tune the models used to investigate these materials at the molecular and quantum levels. This provides insights into the material's properties that would not be possible with physical observation and testing. In addition to speeding up time to market with new photochromic products, PPG knows it is reducing costs, although exactly how much is hard to quantify.

- Researchers at the Salk Institute are using supercomputers at the nearby NSF-funded San Diego Supercomputer Center (SDSC) to investigate how the synapses of the brain work. Their research has the potential to help people suffering from mental disorders such as Alzheimer's, schizophrenia, and manic depressive disorders. In addition, the use of supercomputers is helping to change the very nature of biology — from a science that has relied primarily on observation to a science that relies on high-performance computing to achieve previously impossible in-depth quantitative results. The Salk Institute researchers have been using the San Diego Supercomputer Center, one of several university supercomputing centers supported through the National Science Foundation to provide HPC capabilities to researchers across the country. Given the complexity of the current research, the Salk team simply could not run the simulations without access to the SDSC system.

- The parameter sweeps and simulations executed on the SDSC high-performance computer had some surprises in store for the Salk investigators. The classic view of how synapses work, derived from laboratory investigation, is that neural transmissions occur primarily in dense, protein-rich areas called active zones. But when the Salk team ran their models on the SDSC system, the results indicated that neural communication was not confined to just the synaptic active zones but took place in peripheral areas as well, outside of the synapse. This was highly unexpected and exploded the traditional thinking of how synapses work.

- It took high-performance computing to reveal how limited our concept was of how synapses functioned," Dr. Terrence Sejnowski explains. "It appears there is a second channel outside the active zone of synapses that may be used to process different types of information. While we had been aware of these 'extrasynaptic' receptors, we really had no idea until now what they were for. But now we are beginning to understand that there is a completely separate mode of neural communication, perhaps going on in parallel with the synapse, which was not apparent until we ran the supercomputer simulations." New experiments conducted by other researchers seem to support the Salk team's unconventional findings. By studying the ciliary ganglion, researchers at the Salk Institute are accumulating valuable information that can be applied to the way other synapses in the brain and neuromuscular connections work, not just in chickens, but in human beings as well. Sejnowski and Dr. Thomas Bartol indicate that findings from the pure computational research they are conducting will impact the treatment of many neurological diseases. Most of today's pharmacological agents affect synaptic transmissions. But now it appears that there is significant chemical communications between neurons at other places in the brain and body as well, and these may be implicated in neurological and mental disorders. This revelation may open up pathways for new drug treatments. But the benefits derived from the combination of the Salk Institute research and HPC have even wider implications. For example, HPC allows the researchers to bring together anatomical, physiological, and biochemical data and draw conclusions that are not readily apparent when these and other related disciplines are studied on their own."

- To conduct leading-edge astrophysical research, educate the public, and create its spectacular and highly popular shows based on real science, the American Museum of Natural History's (AMNH's) Hayden Planetarium augments its own computing with the extensive supercomputer capabilities available through the National Science Foundation's high-performance computing centers, such as the San Diego Supercomputer Center. The process of computationally transforming mathematical data into realistic 3D images is called visualization. This is the process used to create the hugely popular animated movies enjoyed around the world. At the Hayden Planetarium, the same process is used to transform enormous volumes of actual scientific data from observations and physics-based computer models (also called simulations) into visualizations of the real universe. Processing this amount of data in a realistic time frame is impossible without high-performance computing. To create its three planetarium shows, the AMNH team used its own HPC resources, currently a powerful cluster of 100 off-the-shelf microprocessors. But for the really huge number crunching needed to render the most difficult and complex scenes, they turned to the HPC capabilities available through the National Science Foundation's supercomputer centers, a partnership between NSF and some of the nation's top universities. Harnessing the thousands of processors available at these HPC facilities has enabled a level of scientific simulation and visualization that was not otherwise possible. For example, sequences in Passport made extensive use of the HPC capabilities and expertise available at the San Diego Supercomputer Center, one of the NSF partners. This included rendering the Orion Nebula sequence. Rendering is the process of following or tracing the light rays to create an accurate image for the visualization. This process is similar in concept to a camera recording a picture when light bounces off an image and strikes the camera lens. Rendering the Orion Nebula was an HPC-sized challenge because millions of light rays had to be traced from the stars through the gas to create an accurate and authentic visualization.

- Microsoft is working with the NSF-funded National Center for Supercomputing Applications at the University of Illinois to develop high-performance computing software products that can be used by hundreds of thousands of users. As high-performance computing has moved out of the government laboratories and universities and is being adopted more broadly throughout the public and private sectors, the need for out-of-the box supercomputing software is growing rapidly. The supercomputer center and Microsoft engineers work as a team — engaging in a two-way knowledge transfer — that allows the software company to develop new software and tools to "deliver HPC to the masses." For Microsoft, lowering the barrier of HPC adoption means providing HPC software that can be used by hundreds of thousands of people to more effectively do their work. This could be a daunting task for a company that is not itself a customer of HPC systems in the traditional sense — Microsoft does not use supercomputers to design new drugs or test the safety of automobiles.

- Researchers at the Science Applications International Corporation in San Diego have been studying the Sun's corona and its impact on space weather for well over a decade. Advances in their understanding of the physics of solar activity, made possible by the use of supercomputers at the nearby NSF-funded San Diego Supercomputer Center, have allowed them to move from one-dimensional models of solar flares and coronal mass ejections to 3D models that can more accurately predict the impact of these phenomena on weather in space. Space weather can have a major impact on Earth's power grids, communications, satellites, and other essential systems.

- Since 1986, Motorola has been using high-performance computing to model cellular networks and telephones, an approach that has given them a clear competitive advantage and made the company a leader in wireless communications. The firm accesses the supercomputers at the NSF-funded National Center for Supercomputing Applications at the University of Illinois to create highly complex models of their wireless devices and infrastructure as different communications technologies emerge. They are now creating models that reflect the newest generation of cellular systems. Their work continues to help Motorola retain its leadership position in the global wireless marketplace.

- Alcoa, the world's leading producer of aluminum products, was one of the first companies to make major use of high-performance computing when it became the NSF-funded Pittsburgh Supercomputing Center's first industrial affiliate in 1987. When aluminum faced growing competition from plastic and other composite materials, the company used the PSC supercomputer to handle the complex modeling and simulations needed to get the jump on their competitors by successfully redesigning cans for the beverage industry and a variety of components for the automotive, aerospace, and building and construction industries.

LEARN MORE

Related Research

Additional research from IDC in the Technical Computing hardware program includes the following documents:

- *Worldwide High-Performance Data Analysis 2013–2017 Forecast,* Steve Conway, Chirag Dekate, Ph.D., and Earl C. Joseph, Ph.D. (IDC #241315, June 2013)

- *Experiences in HPC: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona,* Chirag Dekate, Ph.D., Steve Conway, and Earl C. Joseph, Ph.D. (IDC #241455, June 2013)

- *HPC in Aerospace, Astrophysics, and Astronomy: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona,* Steve Conway, Earl C. Joseph, Ph.D., and Chirag Dekate, Ph.D. (IDC #241451, June 2013)

- *Top Issues for HPC Sites: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona,* Chirag Dekate, Ph.D., Steve Conway, and Earl C. Joseph, Ph.D. (IDC #241463, June 2013)

- *Potential Disruptive Technologies Panel: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona,* Chirag Dekate, Ph.D., Steve Conway, and Earl C. Joseph, Ph.D. (IDC #241452, June 2013)

- *Advanced Visualization: HPC User Forum, April 29–May 1, 2013, Tucson, Arizona,* Steve Conway, Earl C. Joseph, Ph.D., and Chirag Dekate, Ph.D. (IDC #241446, June 2013)

- *Worldwide Technical Computing Server 2013–2017 Forecast,* Chirag Dekate, Ph.D., Earl C. Joseph, Ph.D., and Steve Conway (IDC #241154, May 2013)

- *Supercomputers Exceed 50% of the HPC Server Market in 2012,* Chirag Dekate, Ph.D., Earl C. Joseph, Ph.D., and Steve Conway (IDC #240426, April 2013)

- *Changing Market Dynamics: HPC Meeting Big Data and IDC's Projected Evolution of the Market,* Steve Conway, Chirag Dekate, Ph.D., and Earl C. Joseph, Ph.D. (IDC #240365, March 2013)

- *Livermore Lab Expands Industry Partnerships: Economic Security Is Vital for National Security,* Steve Conway, Earl C. Joseph, Ph.D., and Chirag Dekate, Ph.D. (IDC #240232, March 2013)

- *Worldwide Technical Computing 2013 Top 10 Predictions,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #239421, February 2013)

- *High-Performance Data Analysis at NASA JPL,* Steve Conway, Chirag Dekate, Ph.D., and Earl C. Joseph, Ph.D. (IDC #238254, December 2012)

- *Advanced Research at the South Ural State University Supercomputing Center,* Chirag Dekate, Ph.D., Steve Conway, and Earl C. Joseph, Ph.D. (IDC #238225, December 2012)

- *The Economic Value of HPC in Science and Industry: HPC User Forum, September 2012, Dearborn, Michigan,* Steve Conway, Earl C. Joseph, Ph.D., and Chirag Dekate, Ph.D. (IDC #237182, October 2012)

- *Big Data in HPC: HPC User Forum, September 2012, Dearborn, Michigan,* Steve Conway, Earl C. Joseph, Ph.D., and Chirag Dekate, Ph.D. (IDC #237180, October 2012)

- *How Nations Are Applying High-End Petascale Supercomputers for Innovation and Economic Advancement in 2012,* Earl C. Joseph, Ph.D., Chirag Dekate, Ph.D., and Steve Conway (IDC #236341, August 2012)

- *Tokyo Institute of Technology: Global Scientific Information and Computing Center,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #236243, August 2012)

- *Shanghai Supercomputer Center,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #236245, August 2012)

- *Petascale Supercomputing at the University of Tokyo,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #236292, August 2012)
- *HPC Application Leadership at the Supercomputing Center of Chinese Academy of Sciences,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #236281, August 2012)
- *India Broadening Access to Supercomputing,* Chirag Dekate, Ph.D., Earl C. Joseph, Ph.D., and Steve Conway (IDC #lcUS23621912, July 2012)
- *Institute of Process Engineering, Chinese Academy of Sciences,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #236204, July 2012)
- *How Fujitsu Built the World's Fastest Supercomputer in Record Time and Ahead of Schedule,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #235733, July 2012)
- *Europe Sharpens Its Focus on Exascale Computing,* Earl C. Joseph, Ph.D., and Steve Conway (IDC #lcUS23555112, June 2012)
- *IBM Returns Supercomputer Crown to United States,* Earl C. Joseph, Ph.D., and Steve Conway (IDC #lcUS23547812, June 2012)
- *The Broader HPC Market: Servers, Storage, Software, Middleware, and Services,* Chirag Dekate, Ph.D., Steve Conway, and Earl C. Joseph, Ph.D. (IDC #234682, May 2012)
- *HPC End-User Site Update: RIKEN Advanced Institute for Computational Science,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #233690, March 2012)
- *National Supercomputing Center in Tianjin,* Earl C. Joseph, Ph.D., Steve Conway, and Chirag Dekate, Ph.D. (IDC #233971, March 2012)
- *Worldwide Data Intensive–Focused HPC Server Systems 2011–2015 Forecast* (IDC #232572, February 2012)
- *IDC's Worldwide Technical Server Taxonomy, 2011* (IDC #231335, November 2011)
- *High-Performance Computing Center Stuttgart* (IDC #230329, October 2011)
- *Oak Ridge $97 Million HPC Deal Confirms New Paradigm for the Exascale Era* (IDC #lcUS23085111, October 2011)
- *China's Third Petascale Supercomputer Uses Homegrown Processors* (IDC #lcUS23116111, October 2011)
- *Exascale Challenges and Opportunities: HPC User Forum, September 2011, San Diego, California* (IDC #230642, October 2011)

Copyright Notice