Special Study

The Economic and Societal Benefits of Linux Supercomputers

Earl Joseph, Melissa Riddle, Tom Sorensen, and Steve Conway

April 2022

HYPERION RESEARCH OPINION

Since its introduction in the 1960's, high performance computing (also called supercomputing) has made many contributions to significant scientific, engineering, and industrial advances around the world, as well as to homeland security and other critical government missions. In the 1990's, the first commodity-based supercomputers using the Linux operation system were developed through the work of key early adopters, including David Bader. By the 2000's, the revolutionary change provided by commodity Linux supercomputers redefined supercomputing and made them available and usable to a dramatically larger number of organizations around the world.

Supercomputers have played crucial roles in government agencies and departments. But that's just part of the story. Supercomputers have made cars and planes much safer, more fuel efficient and environmentally friendly. They are crucial aids in discovering and extracting new sources of oil and gas, and for developing alternative energy sources. They have enabled the weather community to create more accurate predictions of severe storms that can devastate lives and property. They are heavily relied on by industries ranging from financial services to medicine and health care, entertainment, consumer products, and more recently by Internet companies. And most recently, supercomputers were instrumental in addressing the COVID-19 pandemic.

In short, high performance computing has become indispensable for maintaining national security and economic competitiveness. That's why other nations and global regions including China, Europe, Japan, and Russia, to name a few, are racing ahead and have created national programs that are investing large sums of money to develop exascale supercomputers for use later in this decade or early in the next decade.

This special study explores the early history of how Linux supercomputers were developed and the overall returns from supercomputing in both: 1) the economic value from building and supporting supercomputers; and 2) the value from using supercomputers. Additionally, this report presents some interesting examples of how supercomputers provide returns.

Over $300 billion in revenue has been generated from selling supercomputers. This represents a sizable economic gain, especially since the use of these systems generated research valued at least ten times over the purchase price. While it is difficult to fully measure the value that supercomputers have generated, even looking at just automotives, aircraft, and pharmaceuticals supercomputers have contributed to products valued at more than $100 trillion over the last 25 years. And this doesn't count the tremendous value to new scientific discoveries in almost all disciplines. The return-on-investment (ROI) and return-on-research (ROR) examples in this report underscore the benefits of supercomputers in enabling scientific and industrial research around the world.

Note: This page is intentionally blank.

CURRENT SITUATION

This special study has three major sections:

- An exploration of the early history of how Linux supercomputers were developed
- A discussion on the economic returns from building and selling supercomputing, and the societal and economic returns from using supercomputers
- Some concrete examples of supercomputer success stories

This report provides examples of how Linux supercomputers have enabled breakthroughs of major scientific or economic importance. These achievements have already saved companies many millions of dollars and have the potential to save billions of dollars. Indeed, without Linux supercomputers, a large number of companies and research organizations would be dramatically less capable than they are today. They would have created far fewer jobs and would have made smaller contributions to national economies, both now and in the future.

THE EARLY HISTORY OF LINUX SUPERCOMPUTING

Conceived of as an open-source version of then industry mainstay Unix operating system, Linux, which began as the personal project of Linus Torvalds in 1991, has since become a ubiquitously leveraged resource in the most advanced computing environments around the world. Adopted early largely by hobbyists, it wasn't until later in the decade that Linux was considered a production level resource. In the mid 1990's, groups like NCSA with the pioneering efforts of David Bader began supplementing and replacing more traditional expensive HPC systems with cheaper, commodity off-the-shelf machines using open-source Linux operating systems.

This approach changed the HPC market very quickly and is now the foundation for 98% of all HPC systems sold, as shown in Figures 1 and 2. In the early 2000's it took only a few years for Linux to overtake all other operating systems as shown in Figure 2. The development in this design approach led to greater accessibility of compute resources across all application spaces and lowered the cost of entry for those looking to enter the space. This opened up supercomputing to a much greater set of users and applications, resulting in a doubling in the use of supercomputing over a short 5-year window (2002-2007) as shown in Figure 1.

100% of the Top500 supercomputers in the world are Linux HPC systems, based on Bader's technical contributions and leadership

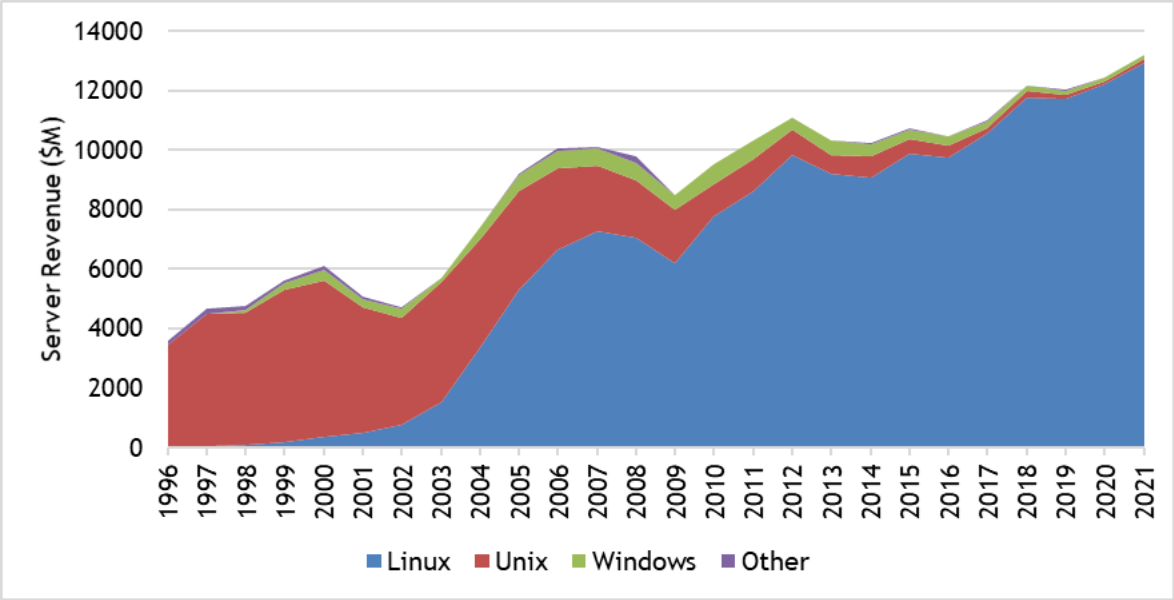
Since these origins, Linux, powered by a vast and energized community of open-source contributors, enterprise developers, and long-time users, has become inextricably linked to high-performance computing. A testament to the power of open-source economics, what began as a personal project has become the cornerstone for nearly all compute-based scientific, industrial, and academic breakthroughs.

The rise of Linux in supercomputing over the last three decades is the cumulative work of countless projects and contributors. That said, few have made such a singular contribution to the conception of this paradigm as David Bader. About Bader's impact on the modern state of supercomputing, National Academy of Engineering member Steve Wallach said, " […] 100% of the Top500 supercomputers in the world are Linux HPC systems, based on Bader's technical contributions and leadership. This is one of the most significant technical foundations of HPC." Similarly, Satoshi Matsuoka, director of RIKEN Center for Computational Science stated, "David has expanded the realm of supercomputing from

narrow sets of technical computing to be the leading edge of mainstream computing we see today in massive cluster-based supercomputers such as Fugaku, as well as hyperscale clouds. As supercomputing progresses onwards, we should further continue to observe other elements in which David has contributed to their genesis."

FIGURE 1

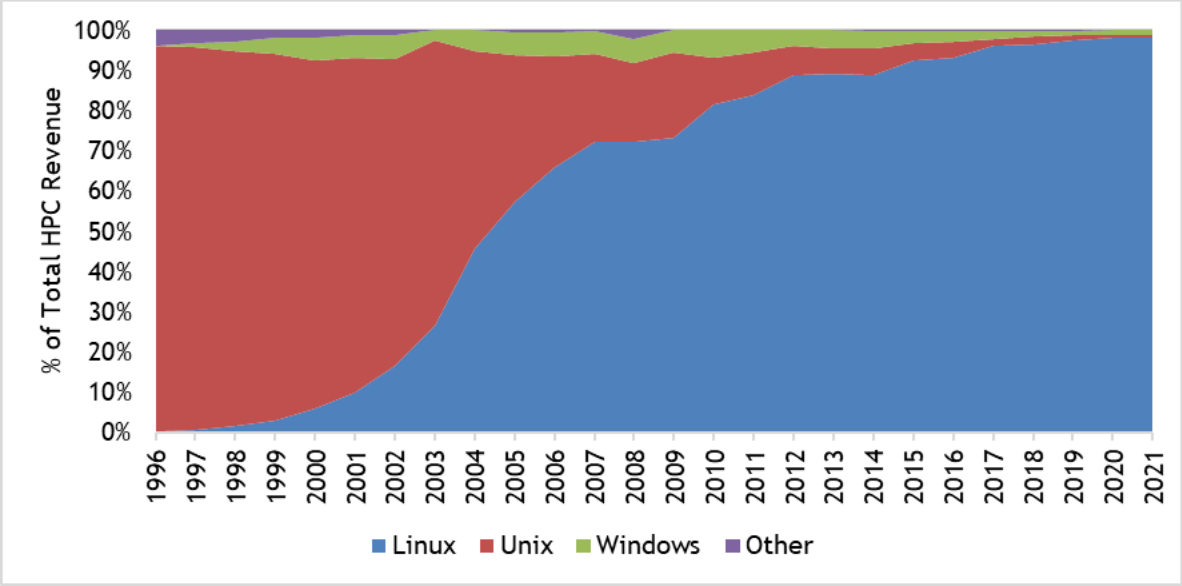
HPC Server Revenue by Operating System



Source: Hyperion Research, 2022

FIGURE 2

Percent of HPC Server Revenue by Operating System

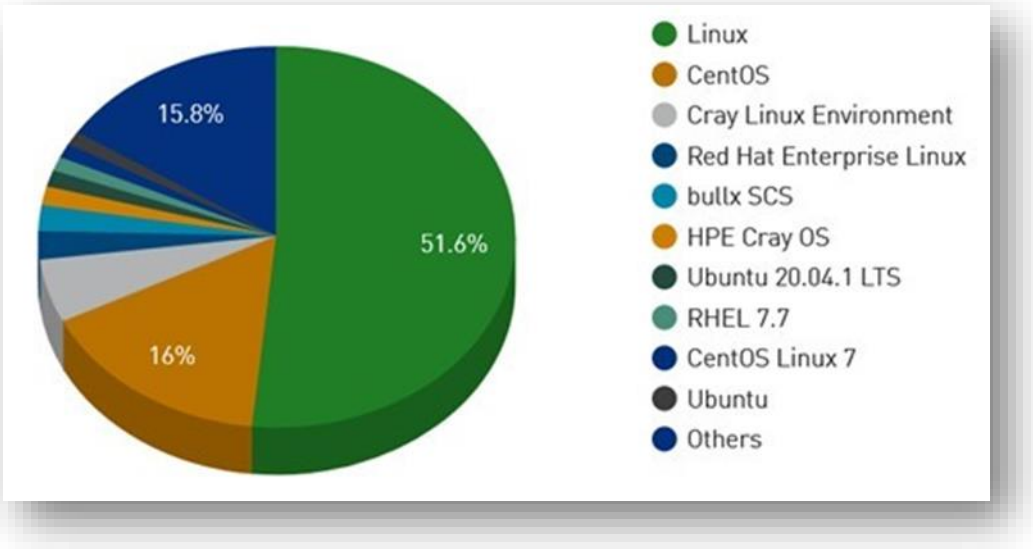


Source: Hyperion Research, 2022

A clear step change in multiple developmental facets of the technology, open-source Linux OS-based supercomputers quickly gained popularity between the years of 2002-2010, representing over 90% of the Top500 in that year, and representing 82% of all technical servers in 2010. While it has since become the only OS represented on Top500 list, there a wealth of diversity for those with different requirements or limitations. Needs for specific hardware, application, data, or user types can and are met with one of a plethora of Linux-based operating systems market offerings (as shown in Figure 3). Figure 3 shows that a number of companies have been created to sell and support Linux operating systems.

FIGURE 3

Linux Operating System Shares



Source: Hyperion Research, 2022

While an individual user or site could have any number of reasons to leverage Linux, most include one or more of the following benefits:

- Linux is generic. It maintains a relatively environment-agnostic status: a single source code can run nearly pain-free in many different hardware spaces.
- Linux is modular. This modular nature empowers users, system designers, and system administrators with customization capabilities with depth and range beyond most other options. The ability to tailor an OS closely to its use requirements is a valuable asset, and one that those seeking every bit of efficiency rarely pass up.
- Linux is scalable. Systems or sites can grow and evolve during their lifecycles. In these cases, it can be potentially difficult to manage workflow as components are added and loads increase in size and amount. Linux's scalable nature is responsible for the tremendous diversity of load types managed by the OS.
- Linux is open source. With universally free access to the complete source code, compute sites and research labs can freely alter code to any degree needed. The unnecessary and cumbersome built-in elements of many operating systems that often lead to lower performance can be deleted by users without intervention from a proprietary body.
- Linux is supported by community members. Its far-reaching use combined with its open-source code has resulted in a multitude of prominent available support resources in the community including free and proprietary offerings.
- Linux is cheap. Using and managing the OS can cost sites very little depending on their size, application types, and other factors. Even if proprietary forms of the Linux OS or support subscriptions are used, the costs remain low compared to other choices.

In a technical space which places a high priority on maximizing and optimizing in every way possible, Linux is an evolving but reliable platform for scaling up and out on a huge variety of devices and environments. Linux, now in its mature stage, is what it always was since its origin: the embodiment of the hard work and ingenuity of people and organizations looking to forward their science and the computing community as a whole. While the rise of Linux-powered supercomputing is a cumulative effort of many individuals, the early work done by David Bader to develop what many consider the first Linux supercomputer demonstrated what would eventually become the modern paradigm of advanced computing.

ESTIMATING THE VALUE OF LINUX HPC SERVERS

There are two main areas where Linux computers have contributed great value to the world:

1. The value from selling Linux HPC computers
2. The value resulting from using Linux computers

The Value from Selling Linux Computers

Hyperion Research's HPC QView tracking Linux servers and the broader HPC ecosystem includes many sectors like bio-life sciences, manufacturing/CAE, chemical engineering, DCC & distribution, economics/financial modeling, EDA and IT, oil/gas, geosciences, mechanical design, defense and homeland security, government laboratories, university/academic, weather and climate modeling, AI, ML, DL, marketing, transportation, and other industries. From this data, the total economic value from Linux HPC computers has been over $300 billion including the servers, the supporting software, storage, and services over the last 25 years.

Based on consistent high growth rates of HPC systems purchases to address some of the most challenging national issues and to increase corporate competitiveness, the next five years are expected to show high revenues. Looking just in 2022, the direct economic returns are projected to exceed $16 billion in servers and around $30 billion including the supporting infrastructure.

The direct economic returns from selling Linux computers in 2022-2026 are projected to exceed $90 billion in servers and an additional $90 billion for the supporting infrastructure. This results in nearly $200 billion in revenue generated from selling Linux supercomputers over just a five-year period. This represents a sizable amount of economic gain, especially since the use of these Linux systems generates research valued at least ten times over the purchase price.

Over $300 billion of revenues were generated from selling Linux computers, but the total economic value is even greater.

The Value Resulting from Using Linux Computers

While it is difficult to fully measure the value that Linux computers have generated, a low estimate would greatly exceed $3 trillion over the last 25 years. The value is both in direct economic returns as well as in major scientific R&D returns and discoveries. Computing has become a pillar of scientific research in almost all scientific disciplines.

Economic Returns from Linux Supercomputers

The use of Linux supercomputers for research can generate substantial return on investment. Table 1 shows the average financial impacts by sector for 175 projects analyzed for the ROI study. Based on the data collected, the average financial ROI was $509.3 dollars in revenues/sales per dollar invested in HPC, and $47.2 dollars in profits or cost savings per dollar invested in HPC. In addition, there were 2,335 new jobs created from these projects.

TABLE 1

Financial ROI Projects

	Average of Revenue $ per HPC $ Invested	Average of Profit or Cost Saving $ per HPC $ Invested
Total	$509.3	$47.2

Note: This study analyzed ROI for 26 academic projects, 6 government projects, and 143 industry projects.

Source: Hyperion Research, 2022

Table 2 shows the financial returns by industry for Linux supercomputers. For profits & cost savings, insurance, oil/gas, finance/insurance and the bio/life sciences had the highest returns. In revenues generated, transportation, finance/insurance and oil/gas have the highest returns, followed by manufacturing, telecommunications, and the life sciences.

TABLE 2

Financial ROI by Industry

Industry	Average of Revenue $ per HPC $ Invested	Average of Profit or Cost Saving $ per HPC $ Invested
Insurance	$175.7	$280.0
O&G	$416.0	$53.7
Financial	$641.7	$47.4
Life Sciences	$205.6	$40.9
Telecomm	$210.7	$30.4
Manufacturing	$216.5	$28.4
Defense	$75.0	$18.8

TABLE 2

Financial ROI by Industry

Industry	Average of Revenue $ per HPC $ Invested	Average of Profit or Cost Saving $ per HPC $ Invested
Transportation	$1,804.3	$15.6
Retail	$30.3	$12.3
Total	$452.1	$37.6

Source: Hyperion Research, 2022

Societal and Economic Returns from Using Linux Supercomputers

Many research projects have societal benefits beyond revenues. The return on research (ROR) for these projects can take many forms. Table 3 shows the main types of innovations and ROR projects in the study. Creating a new approach is the most frequent, followed by creating better products, and then discovering something new.

TABLE 3

Areas of the Projects

Primary Innovation/ROR Area	Count of Projects
Created New Approach	297
Better Products	193
Discovered Something New	94
Scientific Breakthrough	59
Helped Society	51
Support Research Programs	47
Cost Savings	22
Total	763

Source: Hyperion Research, 2022

LINUX SUPERCOMPUTERS ARE USED IN A LARGE NUMBER OF DISCIPLINES

This is an example list of the many areas where Linux supercomputers are a fundamental tool for discovery and research.

Bio-Sciences

- Genomics
- Proteomics
- Drug discovery
- Bioinformatics
- Agricultural research
- Epidemiology/public health
- Precision medicine

CAE: Product Design

- Structural analysis
- Fluid-structure analysis
- Crashworthiness
- Combustion design/testing
- Noise/vibration/harshness
- Aerodynamics
- Environmental friendliness
- Materials science
- Autonomous vehicles
- Smart cities

Mechanical Design

- CAD and drafting
- 3D wireframe
- Storage and access of mechanical designs
- Civil engineering design

Chemical Engineering

- Chemical process engineering
- Molecular modeling
- Chemical analysis

DCC & Distribution

- 3D animation/special effects
- Film editing/production
- Advanced rendering
- Film/video distribution

- Advanced gaming with central servers

Defense

- Surveillance & signal processing
- Encryption
- Communications & intelligence
- Anti-terrorism
- Geospatial image management, analysis
- Weapons development
- Warfighter support
- Law enforcement and homeland security
- Battle simulations

Economics/Financial

- Portfolio optimization
- Pricing exotic instruments
- Global risk management
- Economic modeling

EDA/IT/ISV

- Schematic capture
- Logic synthesis
- Computer system modeling/testing
- Software development, ISV and other software
- CSP (cloud service providers) design and modelling
- IOT devices and systems
- Telecommunications

Geosciences

- Seismic processing
- Reservoir modeling
- Pollution modelling
- Other types of earth modeling

Government Lab

- Scientific research
- Stockpile stewardship
- Industrial partnering
- National programs
- New technologies

Academic/University

- Scientific research

- Industrial partnering

Weather/Climate

- Weather forecasting
- Climate research

Other

- Transportation and logistics
- Retail and marketing
- Social media design and applications
- New non-traditional applications like AI

EXAMPLES OF THE IMPACT OF THE USE OF LINUX SUPERCOMPUTERS

Linux supercomputers have played crucial roles in companies, universities, and government agencies worldwide. The development of advanced aircraft and spacecraft relies on Linux supercomputers. Linux supercomputers enable national and international weather and severe storm predictions that help save lives and billions of dollars in property. Medical researchers use Linux supercomputers to discover lifesaving medicines and model dangerous microbes.

Examples of Linux supercomputer achievements include:

- Seismic simulations are used to generate advanced hazard maps and this research team developed software that significantly advances this map-making ability. The potential saving is in billions of dollars in preventing property loss/damage, along with protecting lives.
- Researchers from the Centers for Disease Control (CDC) created a far more detailed model of the hepatitis C virus, a major cause of liver disease. Annual health care costs associated with this virus are $9 billion in the U.S. alone.
- Linux supercomputers enabled the development of a computer model that comprehensively simulates the human heart down to the cell level. This innovation has significant potential for saving health care costs by reducing heart disease and improving heart health. This research has strong potential for helping to reduce coronary heart disease, which costs the United States over $100 billion each year.
- GE used a supercomputer to reveal a new aspect of turbine behavior that is already providing GE with a competitive advantage in fuel efficiency. Every 1% reduction in fuel consumption saves users of these products $2 billion/year.
- BMI utilized supercomputers and computational models to design components that could save 1.5 billion gallons of fuel and $5 billion in fuel costs per year.
- Supercomputers were used to generate more oil and gas from US reserves. This promises billions of dollars per year in savings and reduction of US dependence on foreign energy.
- Automotive and engine manufacturers are using high performance computing to develop next-generation engines that use less fuel. These fuel savings are estimated at more than $1 billion per year.

But that's just part of the story. Without supercomputers, detecting today's sophisticated cyber security breaches, insider threats and electronic fraud would be impractical. In short, Linux HPC systems have become indispensable for maintaining both national security and economic competitiveness. That's

why many nations and global regions including USA, China, Europe, and Japan have invested large sums of money to develop exascale Linux supercomputers.

Case Study Examples of the Value from Using Linux Computers

These examples illustrate the importance of continuing to provide this diverse community of HPC users with access to first class Linux supercomputers whose capabilities are on a par with those of other leading nations and global regions, including exascale supercomputers as soon as those begin to appear around the world. Anyone who doubts the importance of Linux supercomputers for scientific and economic competitiveness should consider the examples in this report:

- The Centers for Disease Control (CDC) used a supercomputer to advance understanding of the hepatitis C virus, a major cause of liver disease. This paves the way for researchers to discover new therapies for combating the virus. Annual health care costs associated with this virus are estimated to be $9 billion in the U.S. alone.
- The Mary Bird Perkins Cancer Center (Baton Rouge, LA) made important advances that could lower the incidence of second cancers caused by radiation in children receiving radiation therapy. The collaborators saved more than $12 million by using high performance computing. The researchers estimate this achievement has accelerated radiation toxicity research by more than a decade.
- Researchers at the Salk Institute (San Diego) are using supercomputers at the nearby NSF-funded San Diego Supercomputer Center 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.
- About 660,000 people die of heart disease in the United States every year – that's 1 in every 4 deaths. Scientists from DOE's Lawrence Livermore National Laboratory (LLNL) modeled the human heart in much greater detail than before, using one of the world's most powerful supercomputers. These advances lay the foundation for progress in preventing and treating heart disease.
- Turbines are responsible for keeping the lights on, since most of the world's electricity is generated by turbines. And jets couldn't stay aloft without them. Recently, GE used a supercomputer to reveal a new aspect of turbine behavior that is already providing GE with a competitive advantage.
- Thanks to new research on a supercomputer, Ramgen will begin testing a 13,000-horsepower CO2 compressor this year. This compressor is projected to reduce the capital costs of CO2 compression by 50 percent and produce a minimum of 25 percent savings in operating costs. Applying these cost savings to a new 400-megawatt clean coal plant would result in capital cost savings of approximately $22 million and an annual operating cost savings of approximately $5 million.
- A BMI Corp. SmartTruck technology developed on a supercomputer could save 1.5 billion gallons of diesel fuel and $5 billion in fuel costs per year. This technology is now in use.
- Boeing Corporation saved many millions of dollars by using supercomputers. Boeing physically tested 77 prototype wing designs for the 767 aircraft, but for the new Boeing 787 Dreamliner only 11 wing designs had to be physically tested (a 7 fold reduction in the needed amount of prototyping), mainly because over 800,000 hours of computer simulations on supercomputers had drastically reduced the amount of needed physical prototyping.
- Seismic simulations were used to generate a hazard map, with the potential saving of many lives and properties. Upon completion of California state-wide seismic hazard maps, savings

will be many billions of dollars. These National Seismic Hazard Maps will help set building codes and insurance rates, as well as provide short-term forecast of the frequency of damaging earthquakes in California over a specified time span.

- Supercomputing has led to significant improvements in the Navy's effort to improve tropical cyclone intensity prediction.
- Manufacturers will have valuable new information that will ultimately help them design better engines more quickly and at a lower cost. These new models will allow researchers to stretch uses of the models beyond what can currently be done experimentally, testing out theoretical innovations such as low temperature combustion.
- RENCI used supercomputing to compute new data in order to update coastal floodplain maps for North Carolina. These maps are required by FEMA for local communities and municipalities to be eligible for flood insurance coverage under the National Flood Insurance Program (NFIP). The modeling program was so successful in North Carolina that it was used by FEMA and the US Army Corps of Engineers to provide data to update coastal floodplain maps from the Virginia-North Carolina state line to the Delaware Bay.
- The low-cost, post-processing system powered by supercomputing reduces the need for observation wells and has demonstrated commercial success in oil and gas recovery, carbon capture and sequestration and geothermal energy. The system is already in use to track injected carbon dioxide in several energy exploration projects
- Researchers are using simulations to decrease the materials scrapped during the continuous casting process. Decreasing the material scrapped due to defects such as cracks, even by a small percentage, results in a large net savings to steel manufacturers and customers. Based on the roughly 100 million tons of steel produced each year in the U.S. and approximately $400 per ton net cost of scrapping, a one percent reduction in yield loss would save about $400 million per year.

About Hyperion Research, LLC

Hyperion Research provides data-driven research, analysis and recommendations for technologies, applications, and markets in high performance computing and emerging technology areas to help organizations worldwide make effective decisions and seize growth opportunities. Research includes market sizing and forecasting, share tracking, segmentation, technology and related trend analysis, and both user & vendor analysis for multi-user technical server technology used for HPC and HPDA (high performance data analysis). Hyperion Research provides thought leadership and practical guidance for users, vendors and other members of the HPC community by focusing on key market and technology trends across government, industry, commerce, and academia.

Headquarters

365 Summit Avenue
St. Paul, MN 55102
USA
612.812.5798
www.HyperionResearch.com and www.hpcuserforum.com
