Table of Contents

A Note on Scope and Report History 2

Executive Summary 3

Figure I: The Rise of China in HPC 5

I. Current and Expected Future Capabilities and Trends in Advanced Computing 5

Estimating Current and Future HPC Capabilities and Trends 6

Transition to Exascale 6

The Emergence of a Globalized HPC Environment 7

Figure II. Top 10 Systems and Highest Rank by Country in the Top500 List 7

Figure III: Current and Emerging HPC Capabilities 8

II. How HPC is Being Used by the United States and Other Nations 9

The United States 9

Other Nations 10

III. Estimates of the Current and Future Effects of HPC on Economic Growth and National Security 12

Figure IV: Boeing's Use of HPC 13

Technical R&D Challenges Toward Exascale and Other Advanced Computing Breakthroughs 14

Figure V: Projected HPC Energy Requirements 15

IV. Similarities and Differences: Between the United States and the Rest of the World 16

Figure VI: Similarities and Differences 16

Investment 17

Figure VII: 2011 Investment by Country 17

Figure VIII: US Investment Levels 18

Human Resources in Science and Technology (HRS&T) 18

Applications and Ability to use HPC 18

Indigenous Computing Technologies and Rate of Change 19

Publications 19

Figure IX: Scientometric Analysis of Publications in HPC 20

V. Conclusion and Recommendations 20

A Note on Scope and Report History

This report concentrates on so-called “classical computing” systems made up of commercial, off the shelf (COTS) machines, as this technology continues to underpin the bulk of present technical supercomputing used for large-scale scientific and industrial investigations. While other important growing trends, such as big data analytics, cloud, social networks, and mobile computing, are potential technical and societal game-changers of the future, those topics are beyond the scope of this report.

In order to compile this assessment, an interagency working group formed and conducted two months of research, holding weekly teleconferences to agree upon content, length, and focus. The working group was composed of a mix of Intelligence Community country area and science and technology analysts, High Performance Computing (HPC) technical subject matter experts, and national security and defense leaders.

The working group subsequently sponsored a three-day summit in early April 2012, with both unclassified and classified sessions, to review the state of HPC and technical supercomputing in the United States and the rest of the world, and to formulate recommendations for the report. The summit included US Government and HPC industry representatives. Their collective analysis and input is reflected in this assessment.

Executive Summary

Currently, the United States is the world leader in high performance computing (HPC), as measured by the size and diversity of its advanced computing human and technical resources. This leadership position is longstanding and responsible for introducing many of the most significant innovations in the development and use of HPC systems and applications codes. The United States' preeminent role in HPC use and innovation is and will increasingly be challenged by an array of domestic and international forces that are collectively reshaping the landscape of advanced computing.

In an information technology-driven world, HPC and its component technologies are the lynchpin for modern national security and economic competitiveness. Acquiring and developing advanced computing capabilities is essential for the complex war-fighting systems and sophisticated industrial base from which we can compete globally. Innovation and sustainment in HPC has long depended on interactions among human and technical resources and the dynamics among public and private stakeholders. The unfolding of a global HPC environment is ushering in an era in which all major US-based HPC companies must now also focus on growing their foreign markets and competing with HPC aspirant countries, which often employ a different mix of incentives and investments toward future HPC innovation.

The United States leads over all other countries in HPC by many measures, such as the number and diversity of key application codes, installed computing platforms, and a trained workforce to both deliver and use these computing systems. The United States is at risk of losing its leading position in HPC use and innovation and having its advanced computing intellectual capital and corporate knowledge move offshore. Such losses will hinder the United States ability to lead technological innovation in the information technology arena, and possibly allow other countries to deny the United States access to the top-end components needed to build the world-class HPC systems on which it now relies.

Beyond technical considerations knowledge, skill, and talent are critical enablers for advanced computing use and innovation. Human Resources in Science and Technology (HRS&T) is a measure of the talent pool available to undertake the use and application of advanced computing and other technologies. Historically the United States has been a leader in the training and sustainment of a strong and dynamic HRS&T talent pool in HPC use and application. This critical enabler toward HPC leadership is presently declining. A significant portion of the current US HPC workforce will be retiring over the next decade. At the same time, the United States' ability to train and attract top scientific and engineering talent from around the world is waning, along with declines in domestic science, technology, engineering, and mathematics (STEM) educational performance. The United States still possesses a key advantage over other countries in the cultivation of innovative 'out-of-the-box' thinking and entrepreneurship within its higher education system.

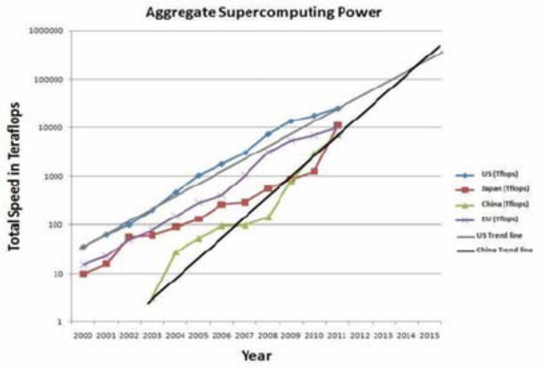
In order to retain US leadership and US knowledge of how to most effectively apply HPC, we make the following observations and recommendations:

- To remain the world's HPC leader, educating the next generation of American HRS&T to develop, use, and innovate by and through advanced computing resources is essential.
- Long-term partnerships with HPC component and system providers is needed in order to meet highly demanding national computational requirements and to ensure a robust and financially sound US HPC sector.
- The United States relies on HPC to maintain its nuclear stockpile and a growing array of advanced conventional weapons systems. HPC systems and their uses have become mainstream across all facets of the United States advanced industrial sector. HPC provides the basis from which many of tomorrow's most profound scientific breakthroughs will arise.
- Foreign and domestic forces are reshaping a global HPC environment in which various countries are jockeying for HPC leadership as US Government investments in advanced computing decline.
- Current and emerging capabilities in HPC and the economic and national security impact of those capabilities are difficult to estimate. They depend on complex interactions among human and technical resources as well as with the larger global HPC environment.
- The next generation of supercomputing system at the exascale level, if it is to be effectively applied, involves probable revolutionary changes in power management and memory utilization that are beyond the grasp of a simple scaling up of today's capabilities.

Additional findings and recommendations are contained in the conclusion to this report.

Several nations are focused on building exascale systems. The United States possesses enough knowledge to attain exascale-capable machines by 2018. However, constructing a usable machine at this scale and being able to push to systems beyond exascale requires significant investment. Technological constraints are an additional challenge. The rise of China as a country aspiring to a leadership position in HPC use and innovation is indicative of the changing global HPC environment confronting US HPC stakeholders in which the pursuit of exascale will unfold.

Figure I: The Rise of China in HPC: *Ten years ago China had no machines on the Top500 list—now it has 68.*



China's Trajectory in HPC. Over the last decade, China has become increasingly active in the HPC world, moving from the bottom of the worldwide High Performance Computing market in 2000 to number four in overall supercomputing power in 2011. China's position in HPC speed has dramatically increased relative to competitors, but this achievement does not necessarily reflect the utility of its HPC assets.

I. Current and Expected Future Capabilities and Trends in Advanced Computing

High performance computers—or supercomputers—are used for an array of problem solving, modeling, and simulation applications in industry, scientific research, large data analysis, and national security activities. Such applications provide the backbone for today's complex global financial system and strategic defense capabilities enabling breakthroughs in energy, medicine, aerospace, US nuclear forces, and many other areas. During the Cold War development of the US nuclear stockpile drove government investment in advanced HPC breakthroughs while spinning off technologies to the industrial sector. Now with global social-technological change and increasing consumer and commercial drivers, HPC innovation is decreasingly driven by national security needs.

Computer Performance

High performance computers' veracity is typically described in terms of the number of operations per second at which they operate. Today supercomputers are in the petaflop range, or one quadrillion floating point operations per second (FLOPS).

NAME	FLOPS
YottaFLOPS	10^{24}
ZettaFLOPS	10^{21}
ExaFLOPS	10^{18}
PetaFLOPS	10^{15}
TeraFLOPS	10^{12}
GigaFLOPS	10^{9}
MegaFLOPS	10^{6}
KiloFLOPS	10^{3}

The definition of HPC is somewhat contextual. While many types of powerful computing systems are used worldwide in countless commercial and civil applications, supercomputers (sometimes called capability computers) are usually designed for specific military, academic, and laboratory applications to solve very complex problems requiring massive mathematical calculations. Such complex problems range from weather forecasting to molecular modeling and from materials simulations to high-energy physics and nuclear weapons simulations.

Architecturally, the transition from terascale HPC to petascale HPC over the past decade leveraged commodity CPU and memory technology through the use of large clusters of commodity nodes linked together by specialized high-performance interconnect networks. Tailored software was developed to manage and schedule these systems and to enable application development by skilled programmers. Although complex and technically challenging, the transition was based on an evolutionary process that relied heavily on performance improvements associated with increases in transistor density and higher clock rates.

Estimating Current and Future HPC Capabilities and Trends

There is no single or universal measure for judging success or progress in HPC research and development. The most common and well-established metric for assessing machine speed in FLOPS is the Linpack Benchmark, which is used to compile the Top500 list (located at www.Top500.org) The Linpack Benchmark is simple and precise, yet has limitations: Linpack is neither representative of highly complex multiphysics codes, nor does it characterize the memory subsystem performance that is important for cryptoanalytic codes. Both of these classes of application codes are vitally important for national security purposes. Establishing a new or additional metric for ranking the Top 500 most powerful computers in the world may be in order for the near future. Other factors need to be considered if we are to make a more comprehensive assessment of the country's progress in HPC. These factors may include:

- Investments (both sovereign and private)
- Impacts of HPC for specific application focus
- Indigenous computing technology capabilities
- Impacts to economic growth and national security for uses of HPC
- Human Resources in S&T (HRS&T)
- Scientific Publications in Computer Science & Engineering
- Rate of change in HPC research and development (alternatives to the Linpack Benchmark)

Collectively, these factors matter when assessing technical supercomputing advancement because they can represent the state of a nation's overall HPC ecosystem. Where available, we have provided data on these factors for countries included in this study. Factors such as national pride, relative levels of resource commitment, and nations' declared policies to excel in the global HPC environment are not quantifiable factors in assessing similarities and differences among countries, but do have importance.

The Linpack Benchmark

The most common gauge of HPC performance is the Linpack Benchmark, a software library that measures a system's floating point computing power by solving a standard system of linear equations. The result is represented in the number of floating point operations or FLOPS. The results are published bi-annually on the Top500 list (top500.org). Notably, the Linpack benchmark measures *peak* performance of a given machine, not sustained performance over the long term running real-world applications. There are alternatives to the Linpack Benchmark. For example, a newer and important but less-established ranking is the Green500 list, which ranks the top 500 machines by energy efficiency. In addition, the Graph500 benchmark has been established to draw attention to graph-based algorithms important for analytics, an emerging area.

Transition to Exascale

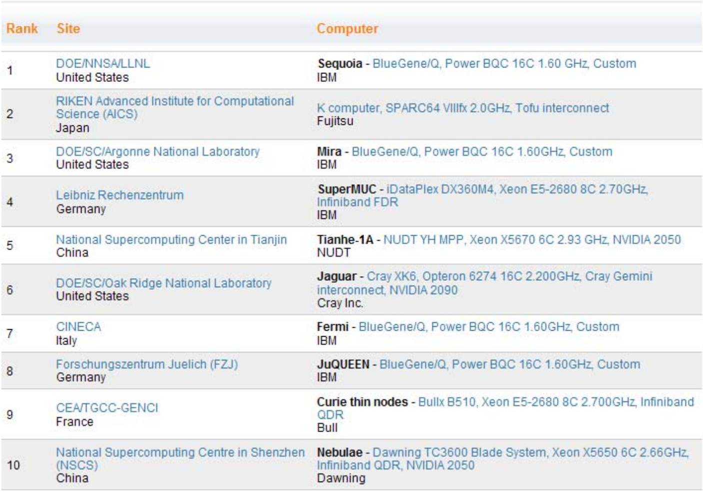
The transition to the next generation of HPC involves fundamentally different requirements in innovation than in the past. State of the art practices in programming will not scale into the future with fixed power constraints, memory subsystem limitations, and a severe lag in programming language development. Although estimates vary, by 2015 the world's fastest supercomputer is expected to reach 100 Linpack petaflops, and realization of speeds surpassing one exaflop (one billion FLOPS) may be reached by 2018. At such exascale speeds, computing levels will be approximately 1,000 times more powerful than any HPC that exists today. Technical constraints towards realizing such processing capabilities are significant and will be discussed later.

The Emergence of a Globalized HPC Environment

One central theme of HPC development that has emerged dramatically over the past decade is the ease with which any country, firm, or research facility can assemble and run an HPC center. HPC technology is widely available through COTS procurements and open-source codes that virtually any organization can obtain. Innovation in contemporary HPC commodity parts is now largely driven by worldwide public demand for faster, smaller, and more accessible electronics such as laptops, smart phones, tablets, and gaming systems. In response to this demand, private sector innovation has yielded dramatic improvements in computer processor speeds, memory speed and size, packaging, and software.

In a global HPC environment performance rankings are often championed by individual country achievements, such as obtaining a place in the widely circulated Top500 rankings. Yet other important contributing elements such as supply chain flow, component manufacture, and commerce have few or no international boundaries. Most western HPC companies produce their components in aspiring HPC nations, especially in Asia, which are keen to begin producing their own indigenous HPCs and software while developing domestic expertise and technology. Increasingly, within any single nation, HPCs are not viewed nor used as a single stand-alone resource but instead are being integrated into a larger advanced computing infrastructure that includes a dedicated national high-speed network that shares HPC capability across a wide range of research groups, or with similar smaller and cheaper grid or cloud environments.

Figure II. Top 10 Systems and Highest Rank by Country in the Top500 List, 14 June 2012:



Rank	Site	Computer
1	DOE/NNSA/LLNL United States	**Sequoia** - BlueGene/Q, Power BQC 16C 1.60 GHz, Custom IBM
2	RIKEN Advanced Institute for Computational Science (AICS) Japan	K computer, SPARC64 VIIIfx 2.0GHz, Tofu interconnect Fujitsu
3	DOE/SC/Argonne National Laboratory United States	**Mira** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM
4	Leibniz Rechenzentrum Germany	**SuperMUC** - iDataPlex DX360M4, Xeon E5-2680 8C 2.70GHz, Infiniband FDR IBM
5	National Supercomputing Center in Tianjin China	**Tianhe-1A** - NUDT YH MPP, Xeon X5670 6C 2.93 GHz, NVIDIA 2050 NUDT
6	DOE/SC/Oak Ridge National Laboratory United States	**Jaguar** - Cray XK6, Opteron 6274 16C 2.200GHz, Cray Gemini interconnect, NVIDIA 2090 Cray Inc.
7	CINECA Italy	**Fermi** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM
8	Forschungszentrum Juelich (FZJ) Germany	**JuQUEEN** - BlueGene/Q, Power BQC 16C 1.60GHz, Custom IBM
9	CEA/TGCC-GENCI France	**Curie thin nodes** - Bullx B510, Xeon E5-2680 8C 2.700GHz, Infiniband QDR Bull
10	National Supercomputing Centre in Shenzhen (NSCS) China	**Nebulae** - Dawning TC3600 Blade System, Xeon X5650 6C 2.66GHz, Infiniband QDR, NVIDIA 2050 Dawning

Despite the prominence of the Top500 list, estimating current and future HPC capabilities and trends is complex and transcends a simple measure of the peak FLOPS performance of any one system located within a country's borders. This net assessment thus groups individual countries' advanced computing goals into three tiers: *Leaders* are those countries—starting with the United States—that have been longstanding innovators and employers of advanced computing resources. *Aspirants* are countries with significant human and technical resources as well as state industrial policies that provide a basis from which to excel in the global HPC environment. *Outliers* are nations whose ambition to leverage advanced computing resources is constrained by a lack of human and technical enablers to participate in the global HPC environment. The outlier countries are in some cases further excluded by export control and international sanction restrictions. The following chart provides current and emerging HPC capabilities for a select group representative of the three-tier approach to analysis used in this report:

Figure III: Current and Emerging HPC Capabilities

(b)(3)

[redacted]

II. How HPC is Being Used by the United States and Other Nations

In the global HPC environment the following general applications have been adopted by all HPC leaders and aspirants to varying degrees:

- Manufacturing
- Defense Applications
- New Material development
- Atmospheric science exploration and discovery
- Aerospace design and development
- Life science research
- Environmental research
- Nuclear power
- Astrophysics

The United States, and to a varying extent other mature nuclear weapons powers, depends on HPC capabilities for ensuring the safety, security, and reliability of their stockpiles in the absence of underground testing. Supercomputing is critically important in the design of advanced conventional weapons, satellite systems, countermeasures development, and defensive systems. HPC enables more efficient design, manufacturing, and testing, allowing weapon systems to reach the battlefield more quickly. HPC is also important for cryptoanalysis, battle space management and communication, and data analytics.

A Cray in One's Hands

The Cray 2 was delivered in 1985 to the US Department of Energy by Cray Research to support development of America's most advanced nuclear weapons.

Today, An Apple iPad 2, pictured below, would compete almost equally in computational power to the Cray 2, pictured above.

The United States

In the past decade, HPC vendors and chip manufacturers have moved toward civilian-oriented production over the traditional defense applications that have driven their R&D efforts. This transformation has come in response to an exponential growth in personal and home digital technologies for consumers worldwide in the form of personal computers, game consoles, tablets, smart phones, 3D Televisions, and cloud computing.

The Fortune 500 commercial industry is a major user of once exotic HPC applications that are increasingly mainstream in advanced manufacturing. Aerospace and car manufacturers use detailed HPC simulations to reduce cost of development programs by replacing expensive physical testing, such as crash safety testing in the case of automobiles, with computer modeling. Further examples of industrial HPC use are provided in the economic impact section of this report.

In the government-sponsored basic and applied science arenas, the National Oceanic and Atmospheric Administration (NOAA) uses HPC to build environmental and climate models. The National Air and Space Administration (NASA) is a significant user of HPC for spacecraft design and aerodynamic modeling, supernova explosion modeling, and studying the creation of the early universe and galaxy formation. The National Science Foundation (NSF) sponsors a great deal of academic HPC–based research in the United States.

HPC has evolved into a “third leg of science,” in addition to theory and experimentation, and is integral to leading edge research in all other scientific fields. In the biology and life science fields, which have grown exponentially in the last couple of decades, genome research, DNA sequencing, and anti-virus capabilities exploration are all big data problems that benefit tremendously from large HPC resources.

In the energy sector, detailed modeling of nuclear reactions is critical to the safety and design of modern nuclear power plants and spent fuel storage. High fidelity calculations with computational models, such as Monte Carlo Neutral Particle (MCNP), require significant computer memory and calculation speed. With the cost of developing a nuclear power plant increasing, using HPC brings significant cost savings.

Other Nations

HPC is used on a global basis for many of the same applications as in the United States, particularly by other leader countries as defined in this assessment. The following provides a survey of HPC current and future use in various leading, aspirant, and outlier countries.

Japan has long been a leader in HPC use and innovation. While the Japanese HPC market was hit hard in 2008 by the global recession, Japan was until recently number one on the Top500 list with its RIKEN-Fujitsu K supercomputer. Currently natural disaster prevention and environmental and life science technologies have arguably the most influence on Japanese HPC usage, notably in the wake of the massive 2011 tsunami and earthquake. K is also being employed by Japanese academic and government users for drug manufacturing, materials science, energy efficiency studies, astronomical studies, and other areas.

As with other leading HPC nations, Japan uses HPC for numerous industrial applications. Examples include automobile manufacturing applications such as Toyota’s use of HPC to support vehicle shape optimization to reduce aerodynamic drag and improve gas mileage.

> *“It still needs another decade before China-made chips meet the needs of the domestic market. Hopefully after two decades, we will be able to sell out China-made CPUs to the US just like we are selling clothes and shoes.”*
>
> - Hu Weiwu, Chief Designer of the Godson microprocessor series, March 5, 2011, addressing the Chinese National People’s Conference.

The European Union uses HPC for automotive industry simulation and modeling and for other applications used by HPC leader nations. Likewise, European countries use HPC for all sorts of scientific research similar to that noted in the United States and Japan. Europe is a world HPC lead user in weather and climate research, biosciences, and physics to industrial applications such as automotive and aerospace design engineering.

China is expanding its HPC usage in the same general fields as other parts of the world while increasingly placing more supercomputers on the Top500 list. According to Xue-Bin Chi, director of the Supercomputing Center of the Computer Network Information Center of the Chinese Academy of Sciences, China has run the following projects, for example, on two of their best supercomputers. The DeepComp6800 has been used for sandstorm prediction and drug screening for Avian Influenza, while the Tianhe-1A has been employed to perform a galactic wind simulation and parallel software research for an eigenvalue problem solver. Likewise, Chinese industry has been employing Tianhe-1A for oil exploration modeling, biomedical research, automotive engineering, civil engineering applications, and for other commercial purposes.

Although the Russians currently use HPC in a similar manner to the United States, historically Russia lagged in supercomputing development during the Cold War anywhere from 10 to 15 years in terms of computational speed and reliability. As in the United States, the then-Soviet nuclear weapons establishment typically drove the state of the art of Soviet supercomputing. The Soviet centralized economy and autocratic rule gravely stifled innovation. The end of the Cold War and the 1990s economic and social upheavals have forced the Russian government to examine more seriously the cost-benefit of information technology at large. Today Russia clearly recognizes that as a nation it needs to catch up in HPC R&D and join the drive to exascale.

Russia is expanding its usage of HPC systems into a broad swath of military and civilian industries, including aviation and space exploration, geosciences, materials science, medicines, and vaccines. HPC systems at Lomonsov Moscow State University have been used toward the development of Sukhoi Corporation's next generation of commercial aircraft. Russian-owned Gazprom, the world's largest supplier of natural gas, was as of 2011 considering acquiring its own petascale computer for seismic data processing and reservoir simulations. Additionally and in general, Russia is using HPC to study industrial fuel and energy matters, modernize its consumer industry, improve weather forecasting, model flight conditions, and design advanced artillery weapons.

Outlier nations, which would encompass countries such as Iran, Pakistan, and North Korea, are seeking to build some form of indigenous HPC capability by whatever means are available to them. Such countries are constrained by a lack of human and technical resources, and in some cases international sanctions regimes, from fully leveraging the economic and defense opportunities enabled by advanced computing. Consequentially, their pursuit of advanced computing is often narrowly focused on military applications, including development of WMD and advanced conventional weapons systems. (b)(3) [redacted]

(b)(3) [redacted]

III. Estimates of the Current and Future Effects of HPC on Economic Growth and National Security

Advanced computing capabilities continue to generate profound and often unanticipated changes that span the modern economic and national security spheres. In a global HPC environment, leaders will strive to maintain their cutting-edge positions in HPC innovation, utilization, and market share, HPC aspirants will continue to benefit from a global HPC environment and potentially leapfrog and gain market share while striving to develop indigenous capabilities. Outliers will likely remain fast followers of HPC technologies, leveraging computational resources where possible, often with a narrow military and defense focus. Overall, worldwide HPC advances at the high-end of manufacturing will likely continue to yield enormous benefits to those societies best able to leverage them and maximize the use of those applications in their industrial processes.

The advanced computing and HPC enabling sub-component sector is a growing segment of international trade in high technology. US vendors currently enjoy leadership in the majority of this market, but the rise of aspirants and diminished US Government support may erode US market leadership over time. For example, the Japanese plan to promote the use of the K supercomputer for resource development in the medical technology field to win 200B Yen worth of international orders.

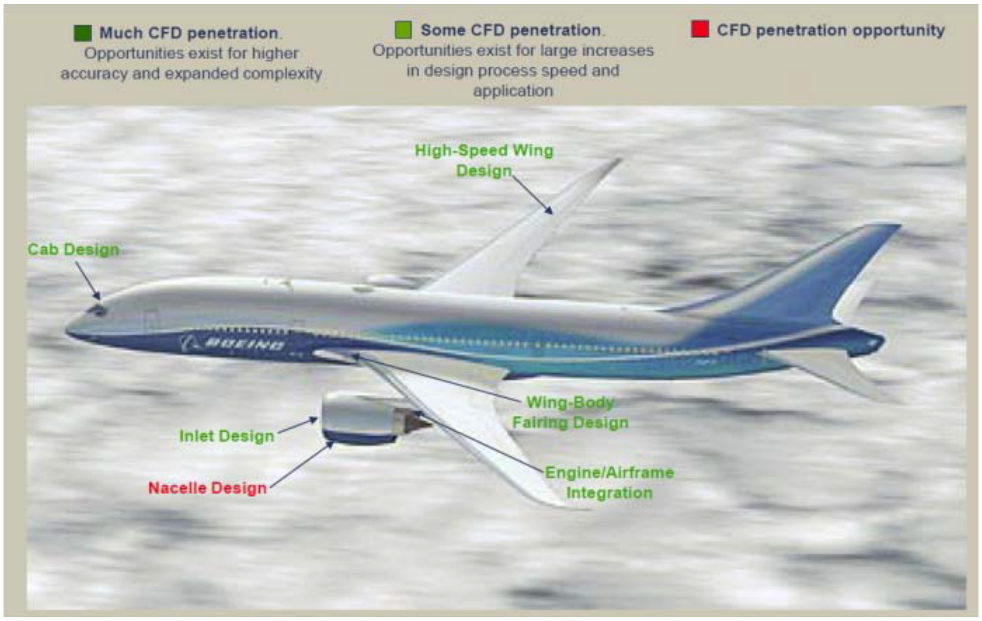
In the national security arena, the United States relies on supercomputing for its most advanced nuclear and conventional weapons capabilities. The nuclear weapons stockpile stewardship mission remains a national security driver for advancement in HPC. Moving toward predictive capability provides the confidence required to replace aging components in nuclear weapons, ensuring the safety, security, and reliability of the nuclear deterrent without the need for underground testing. HPC advancement to useable exascale capabilities is critically important to provide researchers and engineers an understanding of:

- Multiphysics of complicated systems under extreme conditions
- Massive amounts of data generated either through simulations, observations, or experiments
- Simulations for multiscale phenomena

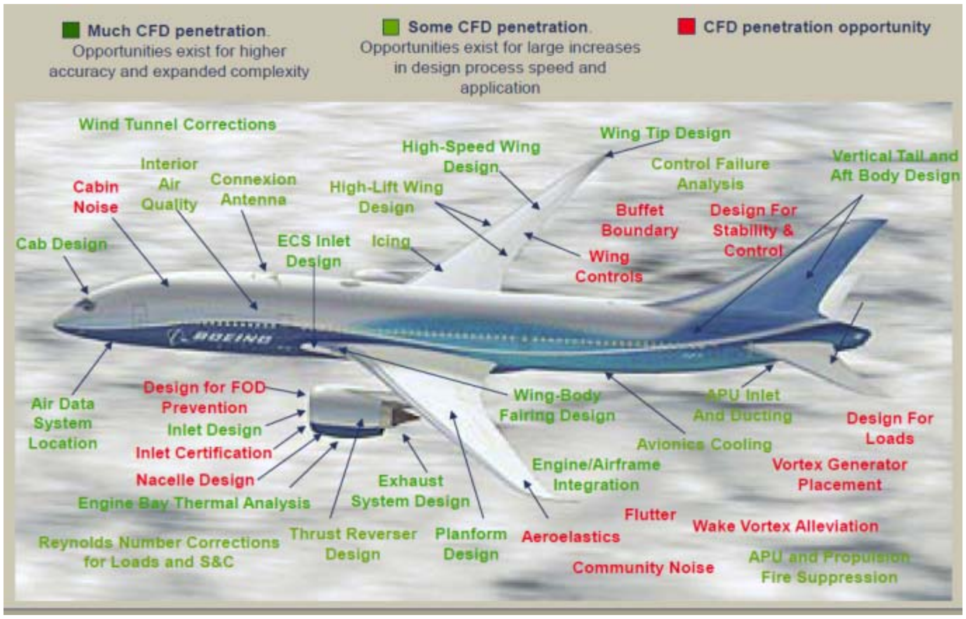
Another key US Government national security driver for HPC advancement is cyber defense. As US adversaries move to more sophisticated techniques to mask malicious cyber activities, the use of HPC in cyber defense applications will become critical.

HPC applications have become a ubiquitous force in the financial, manufacturing, and biomedical sectors among other components of the modern world economy. US capabilities in HPC have helped to retain American industrial leadership in some key competitive industries. In aerospace, for instance, Boeing successfully applied high-end computational fluid dynamics (CFD) simulations to the design of the 787. By integrating wind tunnel experimentation with computer simulation, the company was able to realize greatly reduced design-cycle time and increased data fidelity in early development phases. The illustrations on the next page compare HPC applications between Boeing's 787 and previous generations of aircraft, showing a vast increase in the use of HPC.

Figure IV: Boeing’s Use of HPC



1979 CFD contributions to the 767



2005 CFD contributions to the 787

A recent survey of American hi-tech manufacturing reinforces the conclusion that HPC systems and applications have become mainstream across the spectrum of business operations conducted by those corporations.

- HP has been investing an average of $1B/month, on service, networking, and security acquisitions to address the top tier of IT space.
- IBM has been acquiring software companies at about 1/month.
- Google, Microsoft, and Cisco have combined investment in R&D of more than $21B/yr.
- In 2009 Cisco introduced a server product line to be tightly integrated with networking and storage. Cisco also launched the Acadia joint venture with EMC and VMware to "virtualize the entire hardware stack."

HPC matters not only in manufacturing and consumer goods but also in public welfare, an increasingly sophisticated transportation infrastructure, and many other sectors necessary to support a dynamic modern economy. NOAA, for instance, generates millions of weather and climate information products daily for the protection of life, property, and resulting enhancement of the national economy. The realization of exascale computing levels will potentially revolutionize model-based forecasting for natural disaster preparedness, weather prediction, and air traffic control that impact approximately one-third of the nation's economy. Technical challenges loom in obtaining such computational capabilities that complicate efforts to estimate the full likelihood of their economic and national security impact.

Technical R&D Challenges Toward Exascale and Other Advanced Computing Breakthroughs

It is highly doubtful that current architecture and technologies can be scaled up 100-fold to deliver a useable exascale computer without breakthrough innovations in key technology areas including:

- Large scale on-chip parallelism will likely introduce a need for new algorithms and application formulations. Many current applications rely on weak-scaling algorithms in which the data set size grows proportionately with the number of central processing units (CPU), and the amount of work per CPU is relatively constant. As the number of computational cores per chip increases, more difficult strong scaling algorithms (in which the data set is relatively constant as the number of cores increase), will be required to maintain computational efficiency.

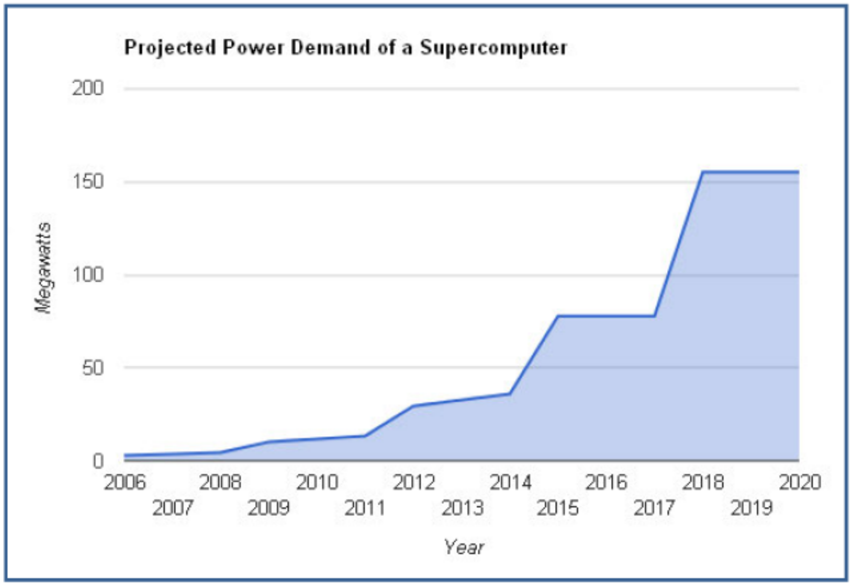
HPC and Modern Consumer Goods

HPC is playing an ever-increasing role in how businesses develop, produce, and distribute modern consumer goods. Since the 1970s, nearly every car, truck, and aircraft in the world has been designed with the aid of supercomputers. Proctor & Gamble's innovative approach to production of its signature potato chip brand Pringles is illustrative of the penetration of advanced computing into all facets of modern manufacturing. P&G makes one billion Pringles chips in two hours. With production levels at this scale during packaging the chips are moving so fast aerodynamic considerations become relevant. The chips have the characteristic of a wing, which creates some interesting challenges when moving in a high speed production process. Another challenge is evenly dispersing seasoning on the chips while they are moving at high speeds. HPC applications of *virtual prototyping* are employed to meet the scale of P&G's Pringles production and the consistent seasoning of each chip involved in these challenges.

- System reliability and resiliency will be challenges for exascale systems with hundreds of thousands of nodes and millions of cores. Current memory and memory subsystem problems will need to be solved. Today, checkpoint/restart is the main strategy for enabling a computation to continue through system failure. Due to the technical challenges of mammoth Input/Output (I/O) rates and energy costs of disk storage, it is entirely likely that new approaches will be required to assure that long-lived computations can progress through multiple system failures.
- Current programming models, development and performance tools, and system software will likely need substantial revision to achieve necessary clarity of problem expression, application development and performance optimization for next-generation systems. It is entirely possible that almost all parts of the current petascale software stack will require substantial revision.

The energy required to power an exascale computer that is a simple scale up/scale out of existing systems is prohibitively expensive (estimated at >100 megawatts [MW]). Improving the op per joule, as well as the databyte transferred per joule metrics are essential and require multiple lines of attack. The Chinese Tiahne 1A, for example, requires 4.04MW of power.

Figure V: Projected HPC Energy Requirements



It is difficult for existing environmental modeling frameworks to extend onto the petascale-class architectures and they need to be re-engineered for the coming exascale platforms. Aside from the challenges posed by the highly threaded, heterogeneous nature of coming computer node architectures, this infrastructure was developed in an era where I/O and memory footprint were essentially free and reliability of individual computational elements was a given. At exascale, memory and I/O use will need to be carefully controlled and the software framework will bear increased responsibility for resilience in the presence of system component faults.

IV. Similarities and Differences: Between the United States and the Rest of the World

As noted, depending on the metrics employed there are quantitative and qualitative differences in how to assess innovation, use, and development of HPC. The following section explores various indicators of how the United States and the rest of the world compare in HPC innovation, use, and development.

Figure VI: Similarities and Differences

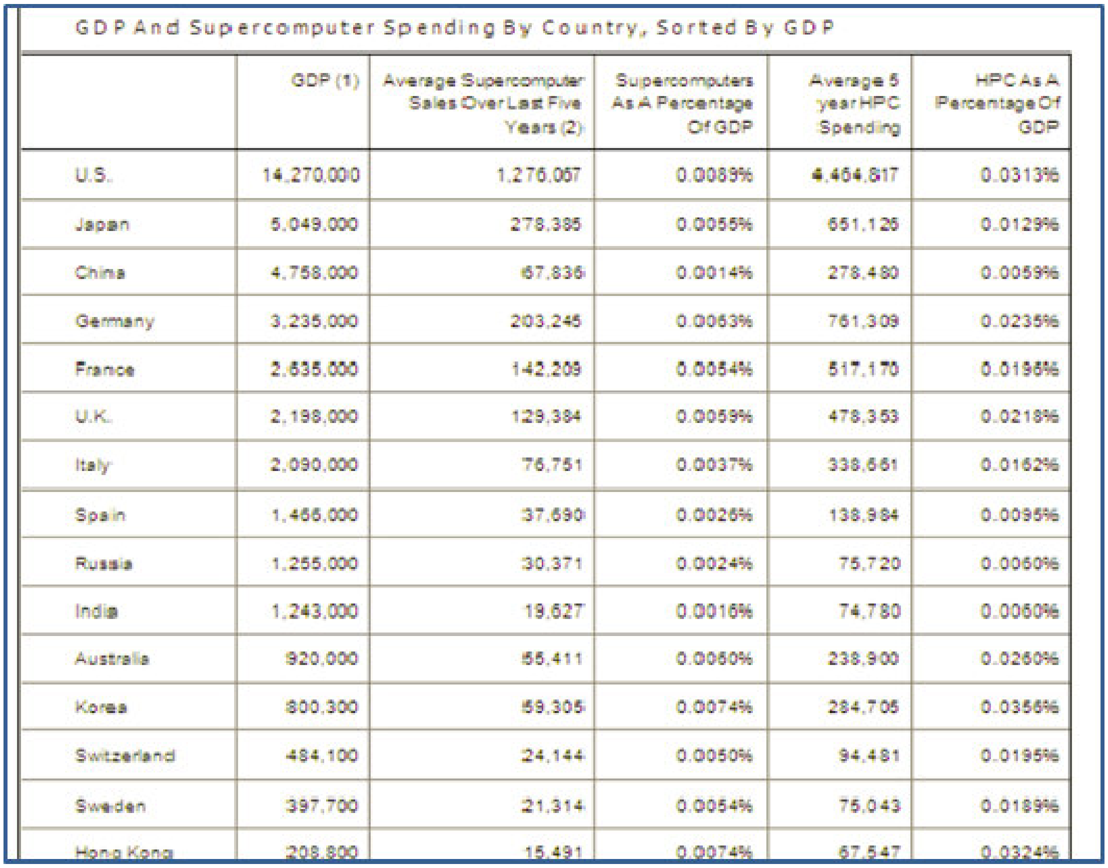
Similarities and Differences in HPC Innovation, Development, and Use Among the United States and Other Nations

	US	EU	Russia	Japan	China	India
Investment	Requires dedicated funding commitment to exascale.	EU doubling HPC investments to $1.6B, plus large investments from each nation.	Committed $1.5B toward exascale, but spread over time.	Willing to put $1B into a single system.	Spending a large amount. Very motivated to BUY leadership.	Increasing investments in HPC, but from a low base.
Human Resources in S&T	Graying of HPC talent. Low graduate rates.	Experienced & talented workforce, but shortages everywhere.	Strong programming and software talents poorly directed.	Strong today, but uncertain future.	Largest number of graduates. Not yet fully trained.	Market forces driving graduates to non HPC alternatives.
Applications	All major areas in industry, government, and academia.	All major areas in industry, government, and academia. Focus varies by individual country.	Strong with military applications but much weaker in industry and new areas.	All major areas in industry, government, and academia, very little in defense areas.	Increasing HPC usage and investments faster than any other nation.	Lack of application usage strength due to dated infrastructure and market forces.
Indigenous HPC Computing Technologies	Market leader for almost all HPC technologies.	Bull/CEA in France. Several world-class software firms.	T Platforms and RUS main systems integrators.	Covers almost every HPC technology area.	Developing processor capabilities, multiple system vendors, custom interconnect capability.	Indigenous producers are limited in global marketplace participation.
Rate of Change HPC Ecosystem	A slowing in R&D investments.	Major recent growth after a slowdown from 3 years ago.	Major growth over last 3 years from a small base.	Unclear rate of growth.	Fastest rate of growth across the world.	Almost steady state and no growth.

Investment

The level of current and planned funding a country is devoting to HPC use and innovation is a ready indicator for that country's potential leadership in the global advanced computing environment.

Figure VII: 2011 Investment by Country

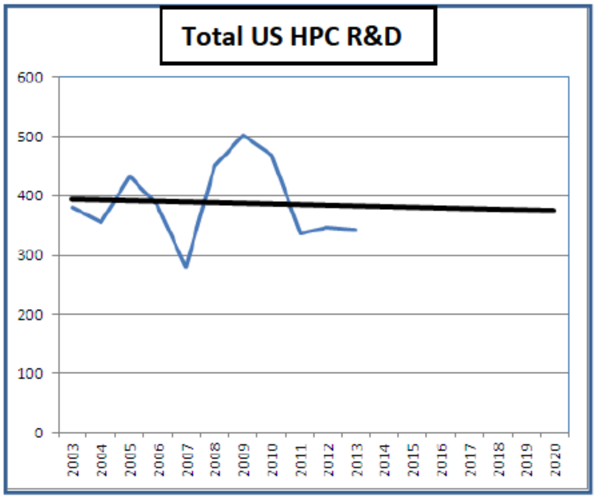


GDP And Supercomputer Spending By Country, Sorted By GDP

	GDP (1)	Average Supercomputer Sales Over Last Five Years (2)	Supercomputers As A Percentage Of GDP	Average 5 year HPC Spending	HPC As A Percentage Of GDP
U.S.	14,270,000	1,276,067	0.0089%	4,464,817	0.0313%
Japan	5,049,000	278,385	0.0055%	651,126	0.0129%
China	4,758,000	67,836	0.0014%	278,480	0.0059%
Germany	3,235,000	203,245	0.0063%	761,309	0.0235%
France	2,635,000	142,209	0.0054%	517,170	0.0196%
U.K.	2,198,000	129,384	0.0059%	478,353	0.0218%
Italy	2,090,000	76,751	0.0037%	338,661	0.0162%
Spain	1,466,000	37,690	0.0026%	138,984	0.0095%
Russia	1,255,000	30,371	0.0024%	75,720	0.0060%
India	1,243,000	19,627	0.0016%	74,780	0.0060%
Australia	920,000	55,411	0.0060%	238,900	0.0260%
Korea	800,300	59,305	0.0074%	284,705	0.0356%
Switzerland	484,100	24,144	0.0050%	94,481	0.0195%
Sweden	397,700	21,314	0.0054%	75,043	0.0189%
Hong Kong	208,800	15,491	0.0074%	67,547	0.0324%

While the United States remains clearly the largest investor in HPC-related technology and systems, an analysis of international trends in such investment indicates that US investment levels are stagnant or declining when compared to the rest of the world. US HPC investments traditionally led by the government sector are declining over the last five years relative to the rest of the world.

Figure VIII: US Investment Levels



China's approach to HPC investment provides an interesting contrast to the US approach. Chinese investment streams in HPC come from multiple sources of funding, making it difficult to estimate overall Chinese investment in HPC: Chinese investment involves a complex mix of sovereign wealth, state-owned conglomerates, small- and medium-sized enterprises, and regional and local contributors. The complex nature of China's HPC investments makes comparison with the traditional US public and private sector bifurcation difficult.

Human Resources in Science and Technology (HRS&T)

HPC advances are unattainable without enough trained personnel, including computational scientists, programmers, system administrators, technologists, and all the others who support the HPC ecosystem. The United States is experiencing low graduation rates at both the undergraduate and graduate levels in HPC-related disciplines. The European Union has an experienced and talented but small HPC workforce. Russia historically has possessed strong mathematical and algorithm development skills, but a lack of resources in Russian technical education threatens this talent pool. Japan has a strong and steady HPC community. China has the largest number of graduates in HPC-related disciplines, but they will require many years to become competitive in HPC innovation. India has a fairly robust pool of talent to draw from, but market forces are driving them into non-HPC areas.

Applications and Ability to use HPC

To some extent, the United States, European Union, Russia, Japan, China, and India all pursue HPC for similar purposes such as weather modeling, aerospace and automotive industries, oil and gas exploration/exploitation, and material and life sciences. The United States hosts the greatest array of disciplines in which HPC is used effectively, followed closely by Japan (for all areas other than offensive military applications). Japan's world-class capability has been developed over decades, and includes the ability to apply HPC to complex industrial applications. The European Union and increasingly China are also adept at applying HPC for large-scale challenges. Russia has a history of applying HPC to small-scale military R&D, and recently has begun to apply large-scale HPC to areas outside of defense applications. India lags in its ability to apply HPC due to dated infrastructure and market pressures driving other uses of HPC-related talent and capabilities.

Indigenous Computing Technologies and Rate of Change

A country's ability to develop indigenous high-end computer technologies provides a key index of its potential for innovation in the global HPC environment. Calculating indigenous capabilities and their impact on the rate of change for a country's HPC innovation potential is complex.

- Japan has a long history of HPC software innovation and plans to allocate $35 million to $40 million for exascale software development.
- The European *action plan for leadership in HPC* involves further development of the European HPC infrastructure and a pooling of HPC national investments. Specific objectives include: providing a world-class HPC infrastructure; ensuring access to HPC technologies, systems, and services; establishing a pan-European governance to pool resources and increase efficiency.
- Some key advances in HPC in the United States have been accelerated by DARPA investments in basic research and technologies. This included initiatives in the development of fundamental networking technologies, large-scale and fast-storage technologies, HPC programmability capabilities, display technologies, and fast and energy-efficient HPC hardware.

Complexities of Assessing HPC Development and Rates of Change

系统主要技术指标

峰值性能: 1.07PFlops

LINPACK效率: 74.4%

主存储器容量: 150TB

外存总容量: 2PB

多核处理器数: 8704

主机系统功耗: 1074KW

性能与功耗比: 741MFlops/W

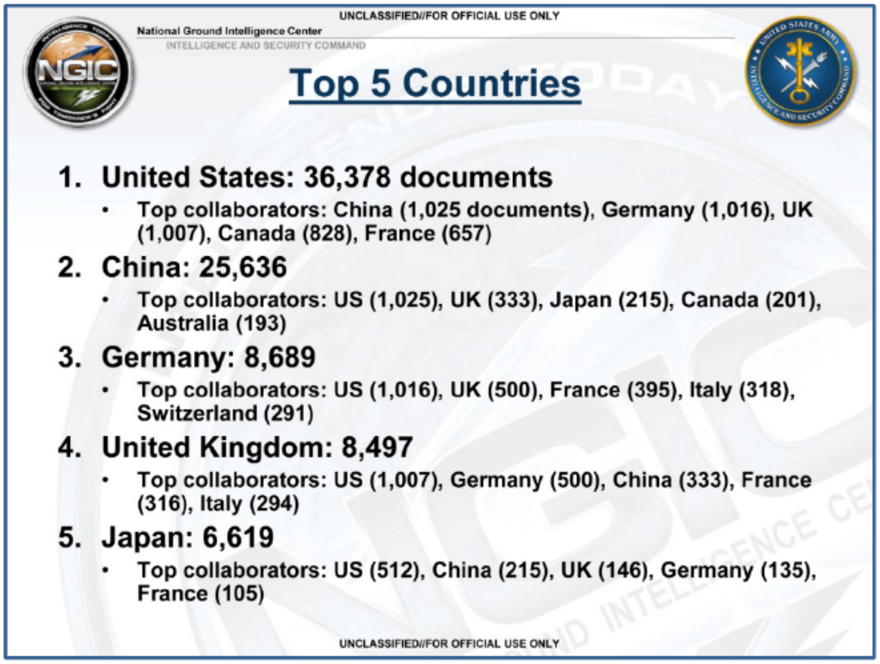
神威蓝光千万亿次计算机系统

China's first indigenous HPC system represents a major achievement using several novel aspects of managing power consumption. Estimating how big a stride this represents in the larger context of China's HPC ecosystem illustrates the difficulties of assessing indigenous innovation and rates of change in HPC capabilities. The Sunway BlueLight MPP, installed in September 2011 at the National Supercomputer Center in Jinan, is being powered by 8,704 ShenWei SW1600 processors. The resulting machine delivers just over a petaflop of performance, with a Linpack rating of 796 teraflops. Still, China remains largely dependent on US-made microprocessors—the brains of the computer—which puts it behind the leading edge. While the Chinese claim that the Sunway BlueLight is a success of China's indigenous innovation effort, experts have shown that the lineage of the ShenWei processor can be traced to the US-developed DEC Alpha 21164 processor.

Publications

The size and breadth of research papers accepted for publication provides an indication of the vigor of a country's academic and scientific participation levels in the global HPC environment. The international research linkages and new collaborative possibilities enabled by modern information communications complicate the association of specific researchers with specific countries. How does one describe the origins of a Chinese post doc funded by the government of Japan conducting research at a US university? Nonetheless, grounds exist from which to undertake a comparison of trends in publication involving individual countries publication levels. Analysts at the National Ground Intelligence Center (NGIC) undertook such a scientometric study for this net assessment. The results of that effort are captured below and reveal the global dynamism of HPC-centered research.

Figure IX: Scientometric Analysis of Publications in HPC



UNCLASSIFIED//FOR OFFICIAL USE ONLY

National Ground Intelligence Center
INTELLIGENCE AND SECURITY COMMAND

Top 5 Countries

1. **United States: 36,378 documents**
   - Top collaborators: China (1,025 documents), Germany (1,016), UK (1,007), Canada (828), France (657)
2. **China: 25,636**
   - Top collaborators: US (1,025), UK (333), Japan (215), Canada (201), Australia (193)
3. **Germany: 8,689**
   - Top collaborators: US (1,016), UK (500), France (395), Italy (318), Switzerland (291)
4. **United Kingdom: 8,497**
   - Top collaborators: US (1,007), Germany (500), China (333), France (316), Italy (294)
5. **Japan: 6,619**
   - Top collaborators: US (512), China (215), UK (146), Germany (135), France (105)

UNCLASSIFIED//FOR OFFICIAL USE ONLY

V. Conclusion and Recommendations

The United States has long been a pioneer and leader in the uses, development, and innovation of HPC. Foreign and domestic forces are reshaping the global HPC environment and will increasingly challenge this position. Nowhere is this more recognizable than in the global pursuit of exascale processing capabilities. Realization of breakthroughs on the exascale level will require sustained attention and investments across the intersections of the financial, human, and technical spheres that underpin the entire US HPC ecosystem. A key finding of this report is that the US HPC ecosystem must be viewed holistically, not in simple terms of a bifurcation between the public and private sectors. For the United States and other countries, the overall health of the HPC ecosystem will be the essential factor in overcoming the technical and financial challenges of attaining exascale capabilities.

Genuine exascale capabilities hold the potential to profoundly impact the United States' economic competiveness, national security, and societal well-being. Exascale computing speeds are projected to be reached by the end of this decade, with multiple countries likely to achieve this milestone. While such speeds may be within reach in the near term, their effective utilization to work simulation and modeling problems is the truly revolutionary potential of exascale. If one country were to reach useable exascale capability several years head of others, that country might capture intellectual property rights and possibly control global HPC R&D, patents, and supply chains.

The growing global availability and capability of scientific computing coupled with advances in computer modeling and simulation are creating opportunities around the world in the commercial and defense sectors for substantial gains in end-use performance, efficiency, manufacturing costs, and product turn-around time. To retain dominance in HPC, we make the following recommendations for the US drive toward exascale and beyond:

(b)(3), (b)(5)

[redacted]

As a nation, the United States has pioneered the greatest uses and innovations in advanced computing, making it the current world leader in this field. Domestic and international forces have developed that will challenge the United States flagship role in HPC uses and innovation.

China and other countries are employing complex mixes of public and private HPC R&D funding as they benefit from the opportunities arising from a new global HPC environment. US efforts are needed to ensure we do not lose our technological, economic, and competitive edge in HPC. The United States should advance HPC technologies addressing both human and technical factors in order to maintain its leadership in advanced computing.