

## Historical review and future challenges in Supercomputing and Networks of Scientific Communication

Authors:

Álvaro Fernández González – University of León / Supercomputing Center Castilla y León  
Assistant Professor University of León  
Faculty of Economics and Business Sciences / Fundación Centro de Supercomputación de Castilla y León  
Campus de Vegazana, S/N  
University of León  
24071 León  
E-mail: [aferg@unileon.es](mailto:aferg@unileon.es) // Phone: +34 987293163

Rafael Rosillo – University of León  
Assistant Professor University of León  
Faculty of Economics and Business Sciences  
Campus de Vegazana, s/n  
24071 LEÓN  
E-mail: [rosc@unileon.es](mailto:rosc@unileon.es) // Phone: +34 987293497

José Ángel Miguel Dávila – University of León  
Associate professor University of León  
University of León  
Faculty of Economics and Business Sciences.  
Campus de Vegazana, s/n  
24071 LEÓN  
E-mail: [jam.davila@unileon.es](mailto:jam.davila@unileon.es) // Phone: +34 987291715

Vicente Matellán Olivera – University of León  
Associate professor University of León  
University of León  
Computer Science Faculty.  
Campus de Vegazana, s/n  
24071 LEÓN  
E-mail: [vicente.matellan@unileon.es](mailto:vicente.matellan@unileon.es) // Phone: +34 987291743

## ABSTRACT

Supercomputing involves not only the development and provision of infrastructures of large capacity for the scientific and business community, but a new way to manage the tasks of research, development and innovation, making it necessary to use high-capacity communication networks that allow the transfer of a great volume of data between research and high performance computing facilities. When first implemented, the use of supercomputers occurred mainly in the military field, at which point they were very rudimentary, offering little possibility of communication networking. Over the years, the improvement of security, privacy and service quality in information-exchange has facilitated the creation of large networks for scientific communication, which in turn have allowed the incorporation of infrastructures for high-performance computing into the improvement of science. This paper analyzes the evolution of Supercomputing and Scientific Communications Networks by means of a critical review of its present state, as well as identifies the main uses today and predicts the challenges of the future uses of this type of advanced services.

**Keywords** Supercomputing · Simulations · Network of Scientific Communication · Scientific Collaboration · Energy consumption

-----

## 1. INTRODUCTION

Supercomputing is currently one of the three pillars, along with theory and laboratory research, on which much of the progress of science and engineering is based. Research on Distributed and Parallel Systems is one of the most developed research lines in current Computer Science [38]. They allow operations with large volumes of data and the implementation of simulation programs in the most varied fields of science through its processing elements. In order to facilitate access to these infrastructures and promote the efficient operation of the whole system of innovation [102] Supercomputing Centers have been created which are developing a new generation of professionals, companies and organizations related to Science and Technology [23].

The investment in these infrastructures has positive effects on the growth in productivity of economies [50,66,91], allowing the improvement of quality, innovation and competitiveness [36]. It is also necessary to improve the creation and exploitation of scientific knowledge and to ensure the quality of higher education [22], meaning a change in the concept of "the object of research" towards a change in the "ways of doing research" [21].

Nowadays, Supercomputing, as part of e-Science, has transformed the traditional way of scientific work, through global collaborations among researchers, the use of large quantities of data, high-speed networks and a large display capacity which allow a type of research that was not possible a decade ago [14]. The National Science Foundation (NSF) in the United States, by means of the "Atkins Report on Cyberinfrastructure" [13], the "Towards 2020 Science" report by Microsoft Research [42] and the world at large today, recognizes that no scientist can be productive or efficient in terms of global research standards if they are not able to integrate Supercomputing into their research process as a binding factor.

This article analyzes the history of Supercomputing and Scientific Communications Networks, as well as their evolution and future challenges, especially due to the significant increase of joint work in the scientific community and the process of globalization based on transnational research agreements, collaboration, resource-sharing and joint activities. For the development of the research, a collaboration of a Group of Experts on Scientific Supercomputing and Networking Communication was established with the task of clarifying concepts to fulfill the goals established. The issue will not only be analyzed technologically but mainly with respect to the different uses that have been made of it over time and which have influenced its progress, as well as what is expected in the future in various sectors. Accordingly, the study focuses on the analysis of historical facts regarding Supercomputing and Scientific Communication Networks, establishing a different consideration of the "systematic review", and allowing the establishment of a basis for future references of prospective studies on the subject. The paper is structured as follows: Section 2 describes method and research objectives; Section 3 details the major findings; Section 4 is a discussion; Section 5 details the main limitations of the study; and Section 6 relates the main conclusions and proposals for future research.

## 2. METHODOLOGY AND RESEARCH OBJECTIVES

An extensive review, with a historical perspective, of the specialized databases (Scopus, Web of Science and Science Direct) was carried out in order to detect, obtain and consult the relevant and specialized literature in relation to the topic under study. Other useful materials were also analyzed, such as websites<sup>1</sup> and relevant reports on the subject, in order to extract and compile the necessary information. This review led to the establishment of five objectives in order to ascertain the relevant aspects of Supercomputing and the relationship that the evolution of Scientific Communications Networks has in this respect. Advice on these matters was obtained from a group of experts in the management of these infrastructures.

### 2.1. OBJECTIVES AND RESEARCH QUESTIONS

The objectives of this work can be summarized in the following questions:

- Q1: Determining the historical moment considered to be the birth of each of the different stages it has gone through in its evolution. This may be done by answering the questions below. When was Supercomputing born? Which are its current major developmental milestones and the ones for the future?
  
- Q2: Establishing how and when the use of Supercomputing for scientific purposes began. The question raised by this objective is: How and when did the transition from the initial uses of Supercomputing to scientific uses occur?

---

<sup>1</sup> [www.top500.org](http://www.top500.org) and [www.green500.org](http://www.green500.org) among others

- Q3: Understanding the uses of Supercomputing and its challenges in future uses. The question aiming to meet this objective is: Which are the current uses of Supercomputing and what is the forecast for the future?
- Q4: Analyzing the development of Scientific Communications Networks. This may be done answering the question: How have Scientific Communications Networks developed?
- Q5: Determining the support that Scientific Communications Networks provide for Supercomputing. The question attempting to meet this objective is: How does the development of Scientific Communications Networks help Supercomputing?

### 3. MAJOR FINDINGS OF THE RESEARCH

In this section the results of the review are presented. The following sub-sections will present the results obtained by the five research goals previously introduced.

3.1. Q1: When was Supercomputing born? Which are its current major developmental milestones and the ones for the future?

In the development of Supercomputing two eras are distinguished: the sequential one, beginning in the 1940s, and the parallel one, beginning in the 1960s and continuing until today. Each era is composed of 3 distinct phases: one phase of architecture, in relation to the system's hardware and two phases of software, one related to compilers and the third to libraries / application packages that let users sidestep the need to write certain parts of code [18,20,34].

The history of supercomputers dates back to 1943 when the Colossus was introduced, the first supercomputer in history, designed by a group pioneering the theory of computation [54], whose aim was the decryption of communications during World War II [100]. In the same year the Electronic Numerical Integrator And Computer (ENIAC) [74] was created in the USA, one of the largest supercomputers of the time, with general large-scale purposes. In 1946, the University of Cambridge built the Electronic Delay Storage Automatic Calculator (EDSAC), considered to be the first programmable computer for general use, and the first digital, electronic and stored-program computer in the world [117]. The architecture used in that period is known as "Von Neumann architecture" [114], is still in use, and consists of a processor capable of reading and writing in a memory which stores a series of commands or instructions and performs calculations on large quantities of input data.

In the following decades, development continued at a fast pace, more so in the US than in Europe, as indicated in the 1956 report by the Department of Scientific and Industrial Research of the UK concerning High Performance Computing. In the 50s, new supercomputers were created, such as the SEAC, ERA 1101 and the ERA 1103. Later, IBM developed several models as well. The latter was responsible for the creation of much of the infrastructure of this decade [81]. 1959, a significant

milestone occurred when the University of Manchester and the Ferranti company cooperated to create the supercomputer known as Atlas [64]. It was introduced in 1962 and was 80 times more powerful than Meg / Mercury and 2400 times more powerful than Mark 1, the other large computational infrastructures of that time. In 1960, the first marketable supercomputer called the CDC 6600 [108] was launched, surpassing by far the most powerful computers at the time in computing power and cost. In the late 60s, the CDC 7600 [95] was released, which has been considered by many to be the first supercomputer in the strict sense, by current standards.

The introduction of Supercomputing in industry began in the 60s, when the first parallel computers were built. Most of these machines were mono-vector processors [48]. The multiprocessor-vector machines were created in the 70s. All machines included integrated circuit memory whose cost was very high, and the number of processors did not surpass 16 [48].

In the 80s, supercomputers increasingly attracted scientific attention [44], mainly due to the beginning of distributed computing, which produced a 16-fold increase in speed and main memory capacity of hitherto existing equipment. An example of this was when the CRAY-2 supercomputer [99] was released in 1985. It was between 6 and 12 times faster than its predecessor. The high-performance computers in the 90s were more related to architectural innovations and software levels [56,105]. In the 90s, it was possible to solve the problems of parallelizing complex tasks such as statistical procedures and the use of digitized pictures by using new algorithms such as the distributed stereo-correlation algorithm. These advances were based on multi-ring architectures with scalable topology, which allowed their use as building blocks for more complex parallel algorithms [4,24,9].

In 1991, the Congress of the United States published the High Performance Computing Act (HPCA) [55], which allowed the development of the National Information Infrastructure. In 1998 the first supercomputer that exceeded the gigaflop barrier of 10<sup>9</sup> operations per second was created according to the Linpack test [37]. In the same year, in response to a report by the Advisory Committee on Information Technology of the Presidency of the United States [59], the NSF (National Science Foundation/USA) developed several "TeraScale" initiatives for the acquisition of computers capable of performing billions of operations per second (teraflops), storage disks with capacities of billions of bytes (terabytes) and networks with bandwidths of billions of bits (gigabits) per second. Based on this initiative the TeraGrid [31] project was begun in 2001, and in 2004 it entered full production mode, providing coordinated and comprehensive services for general academic research in the US. In 2005, the NSF extended its support to TeraGrid by providing a \$150 million investment for operations, user support and improvement of the facility over the next five years.

In 2006 the "HPC in Europe Task Force," a working group of experts that analyzes the evolution of Supercomputing in Europe, published a White Paper entitled "Scientific Case for Advanced Computing in Europe" [62]. This report was a boost for PRACE (Partnership for Advanced Computing in Europe), concluding that only through a joint and coordinated effort will Europe be able to be competitive, mainly because it is expected that the cost of the systems of Supercomputing will be of such a magnitude, that no European country alone could compete with the US and other countries in Asia or Latin America. In the same vein the IDC report [57] provides a number of

recommendations for Europe to lead scientific research and industry in 2020. In 2012, the European Commission (EC) announced a plan that included doubling its investments in Supercomputing from 630 million to 1.200 million euros [113], with focus on the development of 'exa-scale' supercomputers for 2020, capable of performing  $10^{18}$  operations per second.

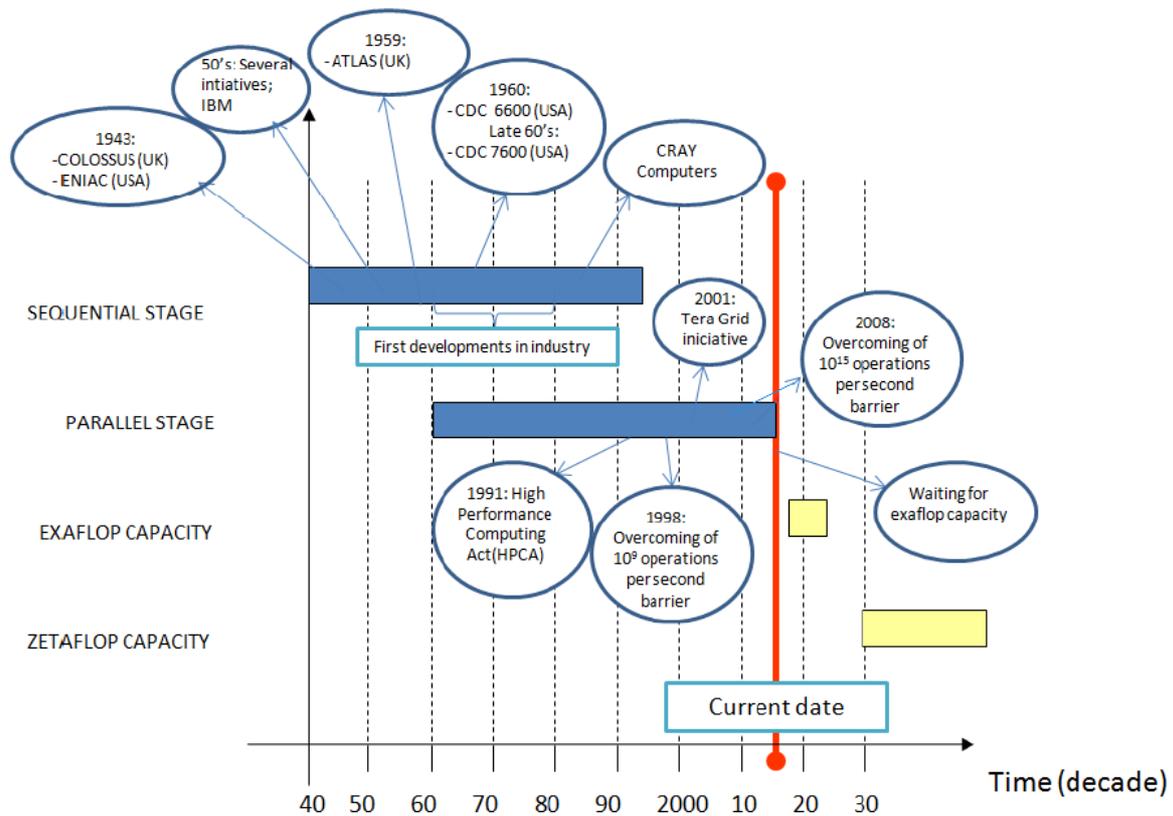
In 2008, the first supercomputer reaching the petaflop speed (1015 operations per second) was created. It was more than one million times faster than the previous [46]. This system had almost 20,000 times the number of processors of the fastest supercomputer 20 years earlier and each of the processors was almost 50 times faster. In the first decade of the 21st century, this scenario of continued exponential growth experienced an interruption, due to factors such as the effect of Moore's Law [97], which states that every 24 months the ability to integrate transistors and energy consumption will double [19]. Due to this latter relationship large refrigeration systems are required, which is a limiting factor for Supercomputing. As a result, in recent years a genuine concern for energy efficiency has arisen. This fact is reflected in the establishment of the Green 500 list<sup>2</sup> in November 2007, which established a ranking that measures the speed of calculation at lower energy consumption of the 500 most efficient supercomputers all over the world.

Answering question Q1, Figure 1 clearly indicates the main dates from the birth of supercomputing to predicted future developments, describing major milestones achieved in the literature review. Horizontal lines show major development eras, blue for the past, yellow for the future. Figure 1 shows that the use of Supercomputing has been growing from 1940, helping the development of industry and science.

---

<sup>2</sup> [www.green500.org](http://www.green500.org)

Figure 1: Development of Supercomputing

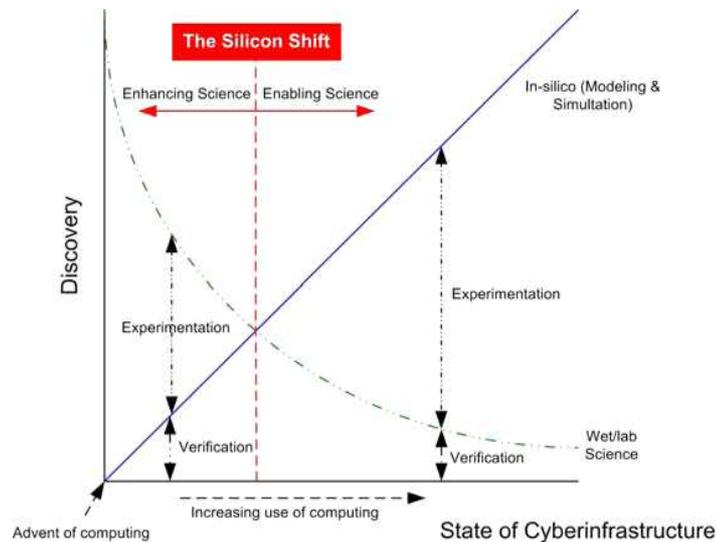


Source: the authors

3.2. Q2: How and when did the transition from the initial uses of Supercomputing to scientific uses occur?

Before the advent of Supercomputing, experimentation was basically done in a laboratory or in the field, and Information and Communication Technology (ICT) only served to assist in verification. As seen in Figure 2, over time the increase in the use of computers made it possible to create better models for scientific simulation, allowing the use of more time for the experimentation and less in verification. From the known moment as "Silicon Shift" onwards, a period began in which computers served not only to improve science but also to enable the development of science.

Figure 2: The defining moment in science: The Silicon Shift (SS). Prior to the SS, computing enhances discovery, whereas after SS it enables discovery



Source: [41]

The 1950s and 60s were years of great competition between the two blocks that divided the world, led respectively by the Soviet Union and the United States. In 1957, the Soviet Union launched the Sputnik Program [53], which consisted of a series of unmanned space missions to demonstrate the feasibility of artificial satellites within the orbit of the earth. In response, the United States created the Research Projects Agency (RPA), whose aim it was to go beyond military applications. The National Science Foundation (NSF) was created to boost basic research and education in all non-medical fields of science and engineering. This was the moment when the real impulse was given to Supercomputing [12], which ceased to serve a purely military purpose, and became a tool supporting research institutions mainly of the public type (universities and government agencies). The first private users of Supercomputers were large companies such as oil companies and banks.

In the 60s the Apollo Project [78] was launched by the United States, which made extensive use of large-scale Supercomputing. The goal of Project Apollo was to simulate a manned flyby of the moon to locate a suitable area for a possible moon landing of astronauts. This project supposed a radical change in the way Supercomputing was understood because of the highly complex simulations on physical equipment and the operational procedures used in the mission. This led to the use of simulation in large and complex systems such as models of biological systems, Artificial Intelligence, particle physics, weather forecasting and aerodynamic design.

In the 80s, new algorithms were developed for digitized images [8,6], designed to work on a transputer network with a simple topology.

In the 90s, distributed computing systems were increasingly used to solve complex problems, highlighting the improvements in evolutionary computation, that is, computational intelligence methods. These emulate the natural evolution of living beings to solve problems of optimization,

searching and learning [49]. In these years, the development of algorithms for more complex parallel operations continued, highlights being the cases of computer vision and image processing. The use of a multi-ring network called Reconfigurable Multi-Ring System (RMRS) was also developed, in which each node in the network has a fixed degree of connectivity and is shown to be a viable architecture for image processing and computer vision problems via parallel computation [3,26,25,5,7]. Supercomputing also became an indispensable tool for industry in the late 20th and beginning of the 21st century. As a result, in the 2004 study by the International Data Corporation (IDC) [57] to explore the use and impact of Supercomputing resources in industry and other sectors, almost all respondents indicated that their use was essential for their business.

In the first decade of the 21st century, new algorithms improved the use of parallel computing such as those based on edge detection in 3-D images, that are targeted at a MultiRing network [10].

In 2007, it became possible to transmit data one hundred times faster and use ten times lower energy than for technologies existing at the time, using light pulses on silicon. This allowed a substantial change in the contribution of Supercomputers to science through simulation and numerical calculation [79], as a key to running experiments.

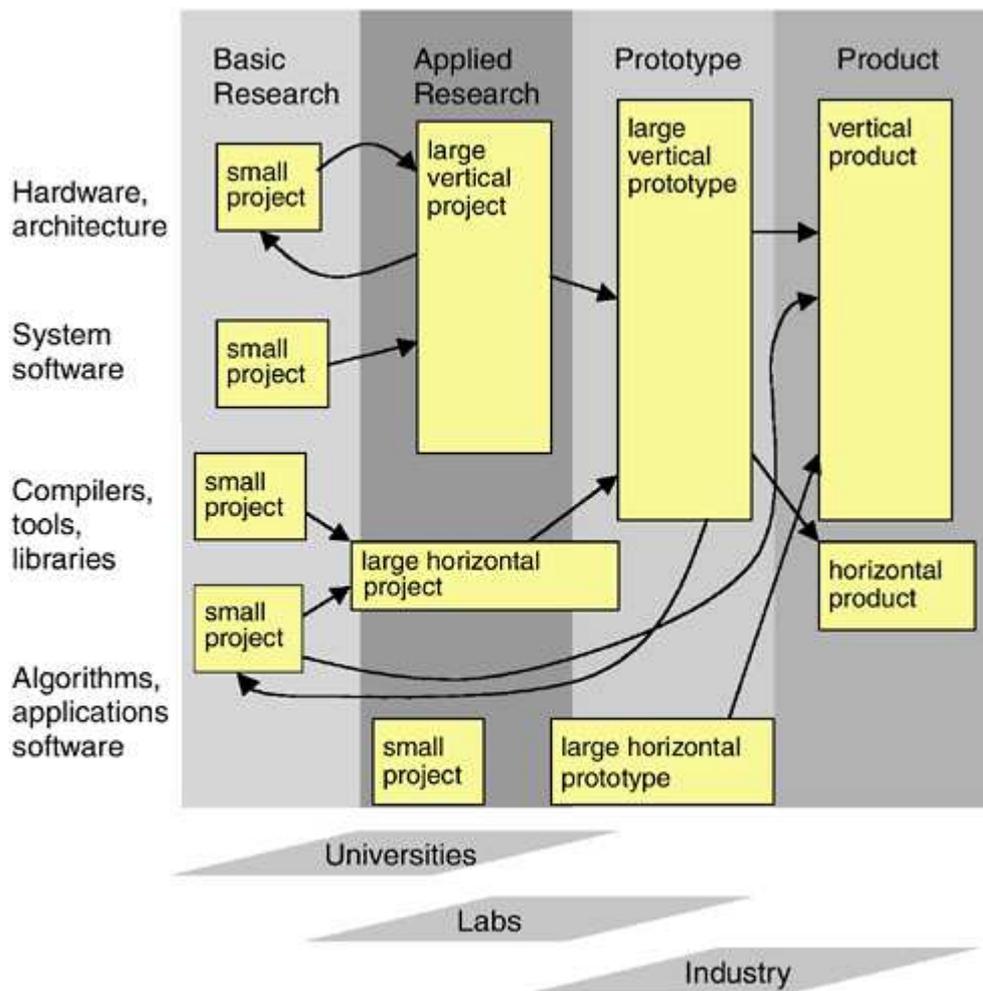
In response to the question Q2, it can be concluded that the 60s were the turning point for the use of Supercomputing in the scientific field.

### 3.3. Q3: Which are the current uses of Supercomputing and the forecast for the future?

Supercomputing has revolutionized design and manufacturing, allowing better products to be manufactured, and reduced risks by means of better analysis and appropriate design decisions. There is a reduction of time and cost not only in design but also in production [96], as simulations of the final product lessen the need for making prototypes. This fact is presented graphically in Figure 3.

Figure 3 represents the process from basic research to the creation of the final product, carried out by the main actors involved (universities, laboratories and industry), using the appropriate supercomputing tools (hardware, software, compilers and algorithms). Generally, basic research is done in small projects, exploring a multitude of ideas. Applied research projects normally validate ideas from basic research and often involve larger groups. When it is possible to develop a prototype that may become a product, the integration of multiple technologies (e.g., hardware, software, new compilers and algorithms) becomes necessary, thereby validating the design by showing the interplay of these technologies. Such development includes many interactions, whereby projects inspire one another, moving quickly from basic research to final products, sometimes requiring multiple iterations of applied research. In this context, failures are as important as successes in motivating new basic research and the search for new products.

Figure 3: The research-to-production continuum



Source: [86]

The improvement of Supercomputing resources provides new capacities for managing and analyzing information, as well as facilities for archiving, conserving and exploiting many kinds of data through which the researchers interpret scientific phenomena. The devices of future supercomputers will provide a new way for application developers to tackle new challenges through the use of open languages and other tools [35] in various scientific and economic sectors.

The improvement of data acquisition devices, the availability of networks for distribution and the increased storage capacity of computers have made it possible for supercomputers to acquire and manage large quantities of data, in the range of terabyte (a trillion characters) or petabytes (a quadrillion characters), and even greater (exa-, zetta-, yotta- etc.). This fact has been highlighted in various scientific publications, as for instance (omitir coma) in a special edition of Nature in 2008 under the title "Big Data: Welcome to the petacentre, science in the petabyte era." In recent years a great number of international initiatives also stand out, which were based on the anticipation of using exa-scale hardware in the coming years [15,47], specifically around 2018 according to some authors [80].

In fact, the most important current challenges of science [28] and engineering, both in simulation and in data analysis, are beyond the capacity of petaflops and are quickly approaching the needs of exaflop computing [39]. Processing these volumes of data poses problems whose computing requirements are beyond the scope of a single machine [83], thereby marking the need to improve the design of High Performance Computers through a mathematical process that allows adequate use of the whole infrastructure of Supercomputing.

Some uses of Supercomputing in various industries and sectors are detailed in the following table:

Table 1: Different industries and sectors where Supercomputing is used.

Automotive industry	<ul style="list-style-type: none"> <li>- Optimization of aerodynamics.</li> <li>- Simulation of combustion processes in engines.</li> <li>- Improvement of vehicle structural integrity in car crashes.</li> <li>- Development of new batteries and motors for the manufacture of electric cars.</li> </ul>
Airline industry	<ul style="list-style-type: none"> <li>- Simulation for efficient use of aircraft fuel.</li> <li>- Real-time simulation of an aircraft in full flight.</li> <li>- Development of accurate turbulence models. In the future flights could be made under the most extreme weather conditions.</li> </ul>
Aerospace industry	<ul style="list-style-type: none"> <li>- Simulation of the re-entry of a spacecraft into the earth 's atmosphere.</li> <li>- Analyzing gas flow in rocket engines.</li> <li>- Processing of data streams received from large-scale satellites.</li> </ul>
Oil industry	<ul style="list-style-type: none"> <li>- Simulations for discovering oil and gas pockets, through analysis of analyzing seismic waves.</li> <li>- Improvement in the accuracy of determining geological structures and the likely location of oil reserves.</li> </ul>
Electrical industry	<ul style="list-style-type: none"> <li>- Real-time management of electricity grids.</li> <li>- Planning of assets replacement and renovation of electricity distribution networks.</li> <li>- Measurement of the network and its technical condition, as well as monitoring tasks and instrumentation.</li> </ul>
Film industry	<ul style="list-style-type: none"> <li>- Progress in the field of computer generated images (CGI).</li> <li>- Improvement of graphic processing for the development of video games.</li> <li>- Playback (rendering) of frames to reduce the time for making an individual frame.</li> </ul>
Financial sector	<ul style="list-style-type: none"> <li>- Estimation of the value of assets, commodity prices and risk assessment.</li> <li>- Responses in real time to data stream inputs, combining high reliability and security.</li> <li>- Applications for development, through integration with third-party software, including databases.</li> </ul>
Health care sector	<ul style="list-style-type: none"> <li>- Improvement of simulations for diagnosis and treatment of diseases.</li> <li>- Optimization of time-sequencing of a genome, with a reduction in time (from 13 years to 1 day at present) and cost (from 3 billion to approximately 1,000 dollars at present).</li> <li>- Creation of tools for storage and analysis of data for genome-based medicine.</li> <li>- Development of bio-simulators, simulating human organs, with the goal of finding cures for diseases.</li> </ul>
Transport sector	<ul style="list-style-type: none"> <li>- Simulation of traffic flows in real time to reduce congestion, travel times and fuel consumption.</li> <li>- Improvement of road network design, using intelligence tools.</li> <li>- Development of a support infrastructure for electric vehicles, for simulation and the design of electrical systems.</li> </ul>
Construction industry	<ul style="list-style-type: none"> <li>- Gathering of aggregate data from various sources (eg. smoke detectors, CCTV, air- conditioning systems, etc.) while combining it with the results of the simulations (for example, to predict how fire might spread using computational fluid dynamics, or how a human being might behave in an evacuation) to produce and evaluate firefighting scenarios of, thereby mitigating risks to people and property.</li> </ul>
Emergencies	<ul style="list-style-type: none"> <li>- Analysis of decision support in anticipation of rescue and relief operations through real-time simulation of systems that need to be controlled and enable rapid evaluation and implementation of action plans. Examples include traffic accidents, floods and fires.</li> </ul>

	- Creation of virtual scenarios to simulate emergencies.
--	--

Source: the authors, adapted from [96]

The table above is based on a major study by the University of Edinburgh in 2011. The analysis proposed in this paper compares supercomputing applications described in table 1 with the cases describing the state-of-the-art for the period 2012-2014 in order to analyze the development of uses in a field exhibiting very rapid advancement.

In the last few years, new uses and applications of Supercomputing have been described that will shape future trends in this discipline. In the following we provide a description of uses of supercomputing in the early 21st century in various fields from a historical viewpoint, by analyzing bibliographical references concerning the use of supercomputing in the Web of Science in the period 2012 to 2014. Details of these challenges are as follows:

Health Care Sector:

- Development of parallelization techniques for the analysis of multiple-concurrent genome, which not only greatly reduces computation time, but also results in an increased usable sequence per genome [88]. Also, new techniques for the sequencing of DNA, such as translocation of molecules through biological nucleotides and synthetic nanopores [72], are using, among others, genome assembly by means of Next-Generation Sequencing techniques (NGS) [76]. This allows personalized cancer treatments by developing virtualization techniques of and by improving the utilization of resources and the scalability of NGS [111]. Furthermore, the uses of tools for large-scale phylogenetic inference with supercomputers of maximum probability [103] are highlighted.
- Cardiology has managed to build models through the use of complex algorithms [121], which show the full three-dimensional interaction of blood flow with the arterial wall. It has also improved the understanding of cardinal function regarding integrated health and disease, using anatomical multistate computer models that are realistic and biophysically detailed. These require a high level of computational capacity and highly scalable algorithms to reduce execution times [82].
- The Human Brain Project (HBP) will develop a new integrated strategy to understand the human brain and a novel research platform that will integrate all data and knowledge about the structure and function of the brain in order to build models that are valid for simulations. The project will promote the development of Supercomputing in the field of life sciences and will generate new neuroscientific data as a reference point to model. It will develop new tools for computing, modeling and simulation, and will allow the construction of virtual laboratories for basic and clinical studies, the simulation of drug use and the creation of virtual prototypes of brain function and robotic devices [71].
- In oncology, systems of diagnosis for colon cancer based on virtual colonoscopies have been described, processed by computationally intensive algorithms, in order to study aspects such as bowel preparation, with computer-assisted screening and examination of colon cancer and

computer-assisted detection in real-time, with the aim of improving sensitivity in detecting colon polyps. There are also mobile systems with high-resolution displays connected to the virtual colonoscopy system to allow visualization of all intestinal lumens and the diagnosis of colon lesions anytime and anywhere [125].

- A study from 2014 uses computational intelligence to analyse large-scale genetic next-generation sequencing data. This allows using approaches to identify genetic diseases which can be utilized in the identification of regulators, which is important in effective biomarker identification for early cancer diagnosis and treatment planning with therapeutic drug targets for kidney cancer [123].

- The area of pharmacy has seen the development of polypharmacy, which studies the ability of drugs to interact with multiple factors, thus attacking the current problems of a rise in the cost of drug development and decreased productivity, incorporating applications such as high performance virtual screening (docking) [40].

- In relation to data-processing in the health care sector, a study of high-resolution displays has been implemented by means of self-organizing maps (SOM) based on a corpus of over two million medical publications. The results of this study show that it is possible to transform a large corpus of documents into a map that is visually appealing and conceptually relevant for experts [101]. Also, an increase in the volume of biomedical data must be met, including next-generation sequencing in clinical histories, which will require large storage-capacity and calculation-methodologies [29]. The use of Supercomputing in complex statistical techniques will allow the accumulation of data on epidemiology, survival and pathology of the world to discover more about genetic and environmental risk, biology and etiology [90].

- High-performance applications will be useful for large-scale projects of virtual screening, bioinformatics, structural systems biology and basic research in understanding protein-ligand recognition [58].

- It will be possible to estimate biologically realistic models of neurons, based on electrophysiological data, which is a key issue in neuroscience for the understanding of neuronal function [69].

#### Aerospace Sector:

- The new generation of radiotelescopes offers a vision of the universe with greater sensitivity [115]. The latest generation of interferometers for astronomy will conduct sky surveys, generating petabyte volumes of spectral line data [116].

- Simulations of core-collapse supernovae in galaxies [70] are being developed. This is a difficult phenomenon to analyze, even after the extensive studies done over many decades. This unresolved issue involves nuclear and neutrino physics in extreme conditions, as well as hydrodynamic aspects of astrophysics [107], thus creating an interesting field of study for the future.

- Supercomputing is used as a fundamental tool for NASA-missions [27] and for scientific and

engineering applications of NASA [93].

- Another important application is the study of the properties of core convection in rotating A-type stars and their ability to create strong magnetic fields. 3-D simulations can serve to provide data regarding asteroseismology and magnetism [43], as does NASA's Kepler mission, which is currently collecting data on a frequent timetable on the asteroseismology of hundreds of stars [77]. This will allow the Sun to be understood in a broader context than it is nowadays, providing comparable structural information on hundreds of solar-type stars. Simulations of emerging data of solar-magneto flow are likewise being carried out [104]. Recent advances in asteroseismology and spectropolarimetry are beginning to provide estimates of differential rotation and magnetic structures for G-type stars and core convection in A-type stars [109].

- Research in astronomy will soon pose serious computational challenges, especially in the Petascale data era, and not every task (e.g. calculating a histogram and computing minimum/maximum data), may be achievable without access to a supercomputing facility that provides an unprecedented level of accuracy and coverage. The analysis of GPU and many-core CPUs is important in this context because it provides a tool which is easy to use for the wider astronomical community, and enables a more optimized utilization of the underlying hardware infrastructure [52].

#### Aeronautical Sector:

- Development of new aerodynamic designs via a simulation that consists of three subparts: core geometry, a computational fluid dynamics (CFD) flow analysis, and an optimization algorithm [63]. Calculations are made on the aerodynamics of the vertical stabilizer as well as an accurate estimation of its contribution to the directional stability and control of aircraft, especially during the preliminary design phase [84]. Another remarkable application is the study of airflow using Supercomputing [65].

#### Meteorology:

- The running of simulations on HPC platforms creates a climate model for conducting climate research at 24 academic institutions and meteorological services in 11 European countries [11].

- Using CFD to model windfarms on land, the prediction and optimization of farm production through the assimilation of meteorological [17] data is possible.

- Simulations of atmospheric dust storms [2], based on the data of an experiment using lasers for remote sensing of aerosol layers in the atmosphere above Sofia (Bulgaria), during an episode of Saharan dust storms [106].

#### Environment:

- Development of climate-modeling through international multi-institutional collaboration on global climate models and prior knowledge of the climate systems inspired by the World Modeling

Summit 2008 [60].

- Modeling of chemical transport emissions (MCTs) in order to estimate anthropogenic and biogenic emissions for Spain with a temporal and spatial resolution of 1 hour and 1 km<sup>2</sup>, taking 2004 as the reference period [51].
- Assistance with the generation of clean energy [61].

#### Biological Sector:

- Creation of databases for the analysis of plant genes [124].
- Development of parallel Supercomputing systems for solving large-scale biological problems using protein-protein interaction (PPI) [73].

#### Emergencies:

- Development of algorithms related to seismic tomography [67].
- Investigation of both tropical cyclones and the impact of climate change through modern models based on Supercomputing work done by NASA [98].

#### Naval Sector:

- Forecasting of real situations on three-dimensional models set up for the Navy and using virtual simulation [30].

#### National Security:

The support of Supercomputers will be essential in national security, one of the main users of Big Data in a wide range of case studies and application scenarios, such as in the fight against terrorism and crime, necessitated by high-performance analysis [1].

Data-intensive applications will gain importance in the future. The volume of measurements, observations and results of simulations will increase exponentially, so that future research efforts should be focused on the collection, storage and exploitation of data as well as on knowledge extraction from these databases.

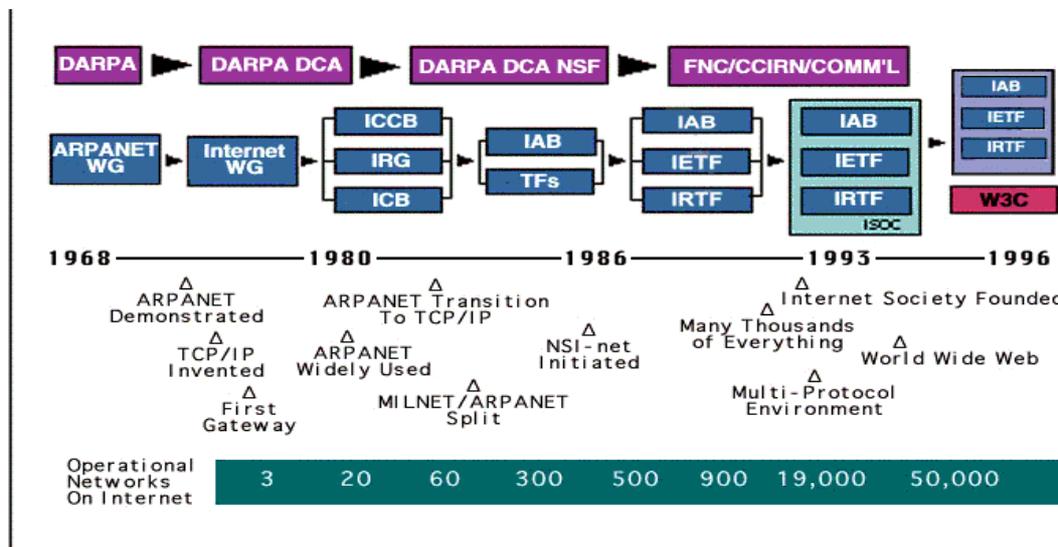
In summary, in response to question Q3, it can be observed, by means of detailing supercomputing applications that virtually all fields of science and industry will experience a breakthrough by using supercomputing.

### 3.4. Q4: How have Scientific Communications Networks developed?

The development of Scientific Communications Networks began in the United States in the 1960s, when ARPANET came into existence. This computer communication network was created by the United States Department of Defense, whose first node was opened in 1969 at the University of California. This network was funded by the Defense Advanced Research Projects Agency (DARPA) and can be considered to be the first scientific communication network in history [85]. One of its alleged origins lies in the space race between the United States and the Soviet Union in the 1950s and 60s, especially after the launch of the Soviet 'Sputnik' satellite in 1957 [53].

1983 is considered the year in which the Internet, as it is known currently, resulted in a separation of the military and civilian parts of the network. 1984 was an important milestone for the interconnection of supercomputers when the US National Science Foundation (NSF) started to design a high-speed successor to ARPANET that would create a backbone network to connect its six Supercomputer centers in San Diego, Boulder, Champaign, Pittsburgh, Ithaca and Princeton. In 1986 the NSF permanently established its own network, called NSFnet, motivated by the bureaucratic impediments to using ARPANET, which disappeared from general traffic as such in 1989. At that time many institutions already had their own networks and the number of servers in the network exceeded 100,000. The aforementioned developments can be seen in Figure 4 [credited to the Internet Society].

Figure 4: Timeline of the ARPANET evolution



Source: [68]

The High Performance Computing Act (HPCA) was passed in the United States in 1991, which allowed the funding of a National Research and Education Network (NREN). The law was popularly referred to as "the information superhighway," and primarily allowed the development of high performance computing and advanced communication, giving a boost to many important technological developments. The experts concluded that if the development of the areas covered by the Act had been left to private industry, it would not have been possible to reach the scientific development achieved through the Act [87].

From the early 90s onward, the Supercomputing Centers of Illinois, Pittsburgh and San Diego all contributed to the development of high-capacity networks through their participation in the Gigabit Network Project [92], supported by the NSF and Defense Advanced Research Projects Agency (DARPA). In 1994, this support was extended again for another two years and in 1995, after the end of the NSFnet project, these centers became the NSF's first nodes of high performance, Backbone Services for research and education. Finally on April 30, 1995, the NSFnet closed down. Since then, the Internet has consisted entirely of various commercial ISPs and private networks (including networks between universities).

In 1996, Internet 2 was created, based on a consortium that emerged as an idea similar to those of the Scientific Communications Networks of the 70s, bringing together over 200 universities, mainly American, in cooperation with 70 leading corporations, 45 government agencies, laboratories and other institutions of higher education in addition to more than 50 international partners [16]. The project's main objectives were to provide the academic community with an extended network for collaboration and research among different members, thereby enabling the development of applications and protocols that can then be commercialized through the Internet and to develop the next generation of telematics applications, facilitating research and education as well as promoting a generation of new commercial or non-commercial technologies.

The national cyberinfrastructure in the United States [28] was the result of the Next Generation Internet Research Act of 1998, the HPCA of 1991 [55], the American Competitiveness Initiative (ACI) and TeraGrid created in 2001. In 2003, TeraGrid capabilities were expanded through high-speed network connections to link the resources of the University of Indiana, Purdue University, Oak Ridge National Laboratory, and the Texas Advanced Computing Center at the University of Texas, Austin. With this investment, the TeraGrid facilitated access to large volumes of data and other computing resources within the scope of research and education. Early in 2006, these integrated resources included more than 102 teraflops of computing power and more than 15 petabytes (a quadrillion bytes) of online and file data- storage with an access and retrieval system using high performance networks. Through the TeraGrid, researchers could access more than 100 databases specific to each discipline.

It must be noted that at the early stages of ARPANET few attempts were made in Europe to join the new network, with the exception of the National Physics Laboratory (NPL), the University College of London in England and the Royal Radar Establishment in Norway [94]. However, despite these limited early initiatives, real interest in the technology developed in the United States did not begin until the second half of the 80s, when there was a large number of TCP / IP networks operating in Europe in an isolated fashion. Some of them began to enjoy the first transatlantic connections to the Internet, usually via dedicated lines financed by US agencies such as the NSF, NASA and the Department of Energy (DoE), which were very interested in cooperating with certain European research centers. Thus, in 1988 and 1989, prestigious European institutions in the Nordic countries (through NORDUnet / KTH), France (INRIA), Italy (CNUCE), Germany (Universities of Dortmund and Karlsruhe), the Netherlands (CWI, NIKHEF) and the UK (UCL) became connected. Some supranational organizations also established dedicated links to the Internet in those years, such as the European Laboratory for Nuclear Research (CERN), the European Space Agency (ESA) and the European UNIX Users Group (EUUG).

In order to coordinate the various initiatives for academic and research networks appearing on a national level in most Western European countries, both economic investment and possible technical solutions were rationalized. Thus emerged such organizations as: JANET (UK), DFN (Germany) and SUNET (Sweden) in 1984, SURFnet (the Netherlands) and ACONet (Austria) in 1986, SWITCH (Switzerland) in 1987, RedIRIS (Spain) and GARR (Italy) in 1988. These networks were interdisciplinary: their aim was to serve the whole of the academic and research community, regardless of their area of activity, by using a single centralized infrastructure, which meant joint forces and benefits from the resulting synergies and economies of scale.

In order to optimize the use of these networks, the European Union is currently promoting the technological development by establishing a network for the joint use of Supercomputing resources by its member countries and the support for studies related to high-performance computing [15]. Through these advanced networks, Europe makes Supercomputing resources more accessible to the projects of scientific and industrial research, and participates in important world-class collaborations that improve productivity by providing Supercomputing resources for general and scientific researchers.

Over the years, the development of Scientific Communications Networks has been increasing in many countries and continents, in addition to the cases cited in the US and Europe. For instance, Latin America has been developing of such networks since the 90s [89]. Currently, there is a network called CLARA (Latin American Cooperation of Advanced Networks), which supports research networks in Latin America and the Caribbean, and a project called ALICE for the interconnection between Latin America and Europe to create an infrastructure for research networks using the Internet Protocol (IP). Likewise, there is a pan-European research network called GÉANT16, whose objective is to lead the operation through the partnership with four European NRENs (National Research and Education Networks) with close historical and social ties to Latin America. Figure 5 shows the details:

Figure 5: RedCLARA 2013



Source: [112]

Since the 90s, the government in other countries such as China has effectively used the public-sector research potential to boost the knowledge-based economy [110], funding virtually unlimited highly skilled human resources, and becoming the fifth leading nation in terms of its share of the world's scientific publications with exponential growth in the rate of papers published, thus making it a major player in critical technologies like nanotechnology. The construction of networks of scientific communication [126] have been outlined, and many studies have been carried out in China from the perspective of how to develop an effective national system or environment for innovations and for increased collaboration between industry and higher education, leading to knowledge transfer between two [120].

In Japan, the development of a new research system throughout the 90s has led to the emergence of new innovation systems in which university-industry linkages have been sought as a means of stimulating regional economic growth. The idea of a regional innovation system (RIS) is relatively new and did not receive much attention in policy frameworks until recently. In 2004, a 'radical' change [122] was introduced to Japanese national universities through the National University Incorporation Law (2003), which meant a change of roles for the universities because of the concentration of resources in 'elite' institutions, and the 'regionalisation' of science and innovation policies. This included 'cluster' initiatives and policies promoting wider university-industry links at a regional level and the promotion of networks among industry, universities and public research institutes, by supporting the creation of new businesses and new industries [119]

### 3.5. Q5: How does the development of Scientific Communications Networks help Supercomputing?

For many years the management and analysis of data produced by Supercomputing applications

were a minimal component of the process of modeling and simulation, in which the management of user data was neglected. With the growing complexity of systems, the complexity of input and output data has also increased. In the future, the volume of data will greatly exceed the current volume, so processing it will become very important, while privacy must be preserved. At the current rate of progress it is projected that exaflop-capacity systems (EFLOPs) will be available around 2019 and zetaflop- capacity (ZFLOPs) in 2030 [75]. To achieve this predicted increase, a highly effective memory will be required, as well as the development of effective programming methodologies, languages and new algorithms capable of exploiting the new, massive, heterogeneous parallel systems with multiple cores. Irregular non-local communication patterns might cause bottlenecks in multi-core supercomputers given the increased data volume. New efficient parallelization algorithms are being developed, but this problem still remains as one of the most complex issues of Supercomputing. [118].

The NSFnet allowed a large number of connections, especially from universities. Although its initial objective was to share the use of expensive Supercomputing resources, the organizations connected soon discovered that they had a superb medium for communication and collaboration amongst each other. Its success was such, that successive enlargements of the capacity of the NSFnet and its trunk lines became necessary at a multiplication rate of 30 every three years: 56,000 bits per second (bps) in 1986, 1.5 million bps in 1989 and 45 million bps in 1992. In 1993 the National Information Infrastructure (NII) was announced, one aspect of which is the National Research and Education Network (NREN), a billion bps "backbone," completed in 1996. Currently the Internet2 Network offers 8.8 Terabits of capacity and 100 gigabit Ethernet technology on its entire footprint, and connection to an international 100 Gbps network backbone.

It is significant to note that the assessment of the effectiveness of research communities must be addressed, not only considering quantitative and scientific production factors but also qualitative factors which influence the successful or unsuccessful integration of research communities. Usually, these platforms are geographically dispersed and interconnected by communication systems that allow implementing new grid and cloud computing platforms [45]. For this reason task scheduling becomes very important in order to manage different users and avoid long delays in queues for computing resources [83].

In 2005 the Council and the Commission of the European Union agreed, through resolutions, to promote and encourage the growth of innovation, research and joint work to attract researchers and encourage trans-disciplinary research projects from global research networks [33]. Currently, new collaborative research projects are being launched, enabling a new way of doing research by linking research communities remotely, via e-science, or e-knowledge.

A clear example of the necessity of using Scientific Communications Networks connected to large computing capabilities is the SKA Project (Square Kilometer Array), considered as an unprecedented global project of science<sup>3</sup> in terms of its size and scale in the field of radio astronomy, and whose mission is to build the world's largest radio telescope, with a square kilometer collecting area. It will constitute the largest array of radio telescopes ever built, as well as

---

<sup>3</sup> [www.skatelescope.org](http://www.skatelescope.org)

represent a qualitative leap in the fields of engineering and research. It will result in increased scientific capacity, so that it is expected to revolutionize fields such as astronomy, astrophysics, astrobiology and fundamental physics. Radio telescopes will be located in South Africa and Australia and the processing of data by supercomputers will be conducted mainly in Europe and the United States. The large volume of data to be handled demonstrates the need for good communication networks that allow optimal data transport from the point of collection to the places of processing, thousands of miles away. The project will run from January 2013 to 31 December 2023. Another proof is the project of the European Laboratory for Nuclear Research (CERN) [32] with a particle accelerator that generates large amounts of information per second. It has become necessary to turn to new sources of analysis, situated in various countries.

In line with this strategy of large-scale development of communications, it must be noted that, according to Robert Vietzke, executive director of Internet 2 [16], in the future, the United States will be interconnected by a network using 100 Gbps wavelengths. Currently, organizations can work with Internet2 and advanced regional networks, based on 100 gigabit Ethernet (GE) technology Layer 2 connection, support for software-defined networking (SDN), and implementation of a model developed by the Department of Energy's ESnet, called Science DMZ. Thus, more than 200,000 academic centers, libraries, health centers, government and research organizations will be connected, which will permit the transport by network of special applications for health, safety and public administration, and improve the transport of data to be analyzed by Supercomputers [16].

In response to the question Q5, we can conclude that the increase in the volume of data to be processed by supercomputers nowadays, and that expected in the future, require harmonized development not only of supercomputers, as estimated by some authors [80], but in the capacity of scientific communication networks for transporting data and its reliability, as well as the consolidation of research communities.

#### 4. DISCUSSION

This study is based on an extensive historical review of the literature of the evolution of Supercomputing and Scientific Communications Networks, infrastructures that help to carry out simulations [96], essential for scientific work [13,42], which in turn will promote the development of various sectors.

Supercomputing will be the driving force in the development of the most important milestones of science. The development of Supercomputing is based on processing large volumes of data, especially when the exaflop capacity arrives in few years' time. It will become necessary to implement parallel processes that require complex algorithms, as well as to improve and expand the capacity offered by Scientific Communications Networks with improved network connectivity that will enable a new generation of applications to interact with machines based on cloud computing. The big volume of data used in various fields, will create new challenges and opportunities in modeling, simulation and theory supercomputing. Computational challenges make possible new opportunities for research. In many topical areas (as it can be seen on the section 3) will be essential

to have a specialist in every area of knowledge for modeling, because if not, the facilities of Supercomputing will not be enough to satisfy the future challenges for the advance of science and technology.

It must not be forgotten, however, that the exponential growth in the processing power of Supercomputers, which requires constant technological advances [97] and occurs every few months, is limited by the considerable increase in power consumption connected with the new infrastructures. Nowadays, the growth in capacity is linked to concerns about finding a range of services with the lowest possible energy consumption. The future development scope scheduled for 2018 will involve "exa-scale" Supercomputers [80], capable of processing a volume vastly superior to current limits as well as the necessary Scientific Communications Networks that allow the transport of huge data volumes.

In summary, the analysis of the five research questions demonstrates how, anticipating the future evolution of scientific infrastructures, it is possible to improve their present use and gain extensive insight regarding future use. Furthermore, it is clear that with knowledge of the different uses and future possibilities, performance will improve, not only in those areas of greatest use described, but also in new fields yet to develop.

## 5. LIMITATIONS OF THE STUDY

This study has a number of limitations that should be considered when interpreting its results and conclusions:

- Only those academic publications which relate to Supercomputing and Scientific Communications Networks in indexed journals have been examined, as well as those which are presentations from seminars and conferences relating to technical matters. In any case, in this and other matters related to technologies, there is a wide range of relevant, albeit informal, information where experiences and projects are detailed in blogs and technical reports that could also provide very important information and which could be used to as a supplement to the basis of the study.
- Some relevant issues might remain unanswered, though of interest to further research. In order to know whether the questions accurately meet the objectives, the collaboration of a Group of Experts on Scientific Supercomputing and Networking Communication was requested, with the aim of verifying that the approach taken was consistent with the responses needed to fulfill the objectives.
- We have tried to analyze the largest possible number of studies on the subject matter, based on a historical perspective of the analysis of the main milestones, but it has been impossible to guarantee a 100% inclusion of all the studies that could be of interest, as set out in systematic reviews. The reason for this limitation is the large amount of existing information and the excessive workload this would imply and which would not guarantee quality references.
- The search was conducted primarily through digital databases and the constraints encountered

were that in some cases the searches were done based on authors. In other cases it was only possible to search for the content inferred from keywords or the purpose of the study, with the advice of the Group of Experts on Scientific Supercomputing and Networking Communication, whose profile is more technical than academic. The number of references of the word 'Supercomputing' note quotes was 1,627 in the Web of Science, of which 10% have been used to conduct the study of this article.

- Due to the numerous fields in which Supercomputing and Scientific Communications Networks are used, there are a large number of studies that analyze the infrastructures as a means, without taking into account the ultimate goal of either this research project or the aspects covered by the objectives and questions in this article. Therefore, in many cases the information obtained was not relevant.

## 6. CONCLUSION

This paper provides a historical review of Supercomputing and Scientific Communications Networks, as well as of their current and future uses, optimizing the work of organizations, while observing that the progress made by the academic and research community has historically contributed decisively to paradigm shifts. During the development of this study, an in-depth analysis of existing literature was conducted, identifying five research questions to demonstrate the importance of Supercomputing and Scientific Communications Networks in the advancement of science, thus enabling new paradigms that will allow doing qualitative and competitive research.

In particular, we have observed throughout the overall data collected, that Supercomputing has progressed on a broad scale since its inception in the 1940s, then exclusive to the military field, until today, when, a part from being applied more intensively to science and in various fields of knowledge, issues such as energy efficiency are matters of great importance. What poses a challenge for the future is the processing of large volumes of information, which need large communication networks.

Based on the above-mentioned historical analysis, we may highlight the following conclusions about the past, the present and the future of Supercomputing services and Scientific Communications Networks, through answers elicited by five explorative research questions: (1) the review of the main milestones of the past, can afford the challenges of the future, especially with the prevision of creation, in next years of exaflop supercomputers; (2) the use of the supercomputers for scientific purposes has a wide background and we can conclude that no scientific research in the future can be ruled without a use of supercomputing tools; (3) in practically in all fields of science and industry will experience a breakthrough in by using supercomputing, so new projects and business can consider supercomputing as a base for its research; (4) the rise of high-speed networks, differentiated from commercial Internet, creates new spaces for sharing, discussions and joining forces without restriction of space, time or distance and for transferring large amounts of data across regions, countries and continents, so it is essential a harmonized development of Supercomputing and the Scientific Communication Networks; and (5) it is clear that, in general, the more capability of Scientific Communications Networks, lets more optimal performance of supercomputing

services.

This study has found that the available supercomputing facilities must meet the purpose of being suitable instruments for simulation processes in various fields, especially when increasingly often the vast majority of problems must be solved by a joint effort of multiple scientific disciplines. The development of collaborative research should be sought to optimize the use of Supercomputers.

The models for simulations in Supercomputers, due to the large volume of data, will be algorithmically and structurally complex and will contain large amounts of information. Therefore, the hardware should be efficiently used, while simultaneously trying to minimize elevated power consumption. The design of interconnection networks among processors in each chip and among system nodes are issues that require new ideas, as do communication networks for the exchange of data. It is essential to have the means to train personnel adequately in the use of these technologies.

Finally, we must note that it will be necessary to use the current knowledge related to these matters to provide a starting point for further research and to explore new fields where the use of Supercomputing will be helpful.

**Acknowledgment:** The authors acknowledge financial support from the Spanish Ministry of Science and Competitiveness grant (ECO2012-25439). The authors would like to thank the anonymous reviewers for their valuable comments and suggestions to improve this paper.

## BIBLIOGRAPHY

1. Akhgar B, Saathoff G B, Arabnia, H R, Hill R, Staniforth A, and Bayerl P S (2015) Application of Big Data for National Security: A Practitioner's Guide to Emerging Technologies. Butterworth-Heinemann
2. Alonso-Pérez S, Cuevas E, Pérez C, Querol X, Baldasano J M, Draxler R and De Bustos J J (2011) Trend changes of african airmass intrusions in the marine boundary layer over the subtropical Eastern North Atlantic region in winter. *Tellus Series B-Chemical and physical meteorology* 63(2): 255-265
3. Arabnia H R (1990) "A Parallel Algorithm for the Arbitrary Rotation of Digitized Images using Process-and-Data-Decomposition Approach". *Journal of Parallel and Distributed Computing* 10(2): 188-193
4. Arabnia H R (1995) "A Distributed Stereocorrelation Algorithm". *Proceedings of Computer Communications and Networks (ICCCN'95)*, IEEE: 479-482
5. Arabnia H R and Bhandarkar S M (1996) "Parallel Stereocorrelation on a Reconfigurable Multi-Ring Network". *The Journal of Supercomputing (Springer Publishers)* 10(3): 243-270
6. Arabnia H R and Oliver M A (1987) "A Transputer Network for the Arbitrary Rotation of Digitised Images". *The Computer Journal* 30(5): 425-433
7. Arabnia H R and Oliver M A (1987) "Arbitrary Rotation of Raster Images with SIMD Machine Architectures". *International Journal of Eurographics Association (Computer Graphic Forum)*

8. Arabnia H R and Oliver M A (1989) "A Transputer Network for Fast Operations on Digitised Images". *International Journal of Eurographics Association (Computer Graphics Forum)* 8(1): 3-12
9. Arabnia H R and Smith J W (1993) "A Reconfigurable Interconnection Network For Imaging Operations And Its Implementation Using A Multi-Stage Switching Box". *Proceedings of the 7th Annual International High Performance Computing Conference. The 1993 High Performance Computing: New Horizons Supercomputing Symposium, Calgary, Alberta, Canada: 349-357*
10. Arif W M and Arabnia H R (2003) "Parallel Edge-Region-Based Segmentation Algorithm Targeted at Reconfigurable Multi-Ring Network". *The Journal of Supercomputing* 25 (1): 43-63
11. Asif M, Cencerrado A, Mula-Valls O, Manubens D, Doblás-Reyes F and Cortés A (2014) Impact of I/O and Data Management in Ensemble Large Scale Climate Forecasting Using EC-Earth3. *Procedia Computer Science* 29: 2370-2379
12. Aspray W, and Williams B O (1994) Arming American scientists: NSF and the provision of scientific computing facilities for universities, 1950-1973. *Annals of the History of Computing, IEEE* 16(4): 60-74
13. Atkins D (2003) Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. National Science Foundation
14. Atkins D, Borgman C, Bindoff N, Ellisman M, Feldman S, Foster I and Ynnerman A (2010) Building a UK Foundation for the Transformative Enhancement of Research and Innovation. Report of the international panel for the 2009 review of the UK research councils e-science programme. Swindon, UK
15. Attig N, Gibbon P and Lippert T (2011) Trends in Supercomputing: The European path to exascale. *Computer Physics Communications* 182(9): 2041-2046
16. Aulkner L (2006) Internet 2 About us. URL: <http://www.internet2.edu/about>. 2014. Retrieved 12 Nov 2014
17. Avila M, Folch A, Houzeaux G, Eguzkitza B, Prieto L and Cabezón D (2013) A Parallel CFD Model for Wind Farms. *Procedia Computer Science* 18: 2157-2166
18. Bacon J (1998) "Concurrent System. Operating Systems, Database and Distributed Systems: an integrated approach". Addison Wesley
19. Ballardini J, Grosclaude E, Hanzich M, Suppi R, Rexachs del Rosario D and Luque Fadón E (2010) Incidence of Parallel Models and Scaling CPU Frequency in the Energy Consumption of the HPC Systems Programming In: XVI Congreso Argentino de Ciencias de la Computación (CACIC' 2010)
20. Banerjee P (1994) "Parallel Algorithms for VLSI Computer-Aided Design", Prentice-Hall

21. Bermeo H P, De los Reyes E and Bonavia T (2008) Dimensions of the Scientific Collaboration and its Contribution to the Academic Research Groups' Scientific Quality. *Research Evaluation* 18(4): 301-311
22. Bernhard A (2009) A knowledge-based Society Needs Quality in Higher Education. *Problems of education in the 21st century* 12: 15-21
23. Bethel E W, Van Rosendale J, Southard D, Gaither K, Childs H, Brugger E and Ahern S (2011) Visualization at Supercomputing Centers: The Tale of Little big Iron and the Three Skinny Guys. *IEEE Computer Graphics* 31(1): 90-95
24. Bhandarkar S M and Arabnia H R (1995) "The REFINE Multiprocessor: Theoretical Properties and Algorithms". *Parallel Computing (journal), Elsevier*, 21(11): 1783-1806
25. Bhandarkar S M and Arabnia H R (1995) "The Hough Transform on a Reconfigurable Multi-Ring Network". *Journal of Parallel and Distributed Computing* 24(1): 107-114
26. Bhandarkar S M, Arabnia H R and Smith J W (1995) "A Reconfigurable Architecture For Image Processing And Computer Vision". *International Journal of Pattern Recognition And Artificial Intelligence (IJPRAI) (special issue on VLSI Algorithms and Architectures for Computer Vision, Image Processing, Pattern Recognition And AI)* 9(2): 201-229
27. Biswas R, Dunbar J, Hardman J, Bailey F R, Wheeler L and Rogers S (2012) The Impact of High-End Computing on NASA Missions. *IT Professional* 14(2): 20-28
28. Bollen J, Fox G and Singhal P (2011) How and Where the TeraGrid Supercomputing Infrastructure Benefits Science. *Journal of Informetrics* 5(1): 114-121
29. Brown J R and Dinu V (2013) High performance Computing Methods for the Integration and Analysis of Biomedical Data using SAS. *Computer Methods and Programs in Biomedicine* 112(3): 553-562
30. Bub F L, Mask A C, Wood K R, Krynen D G, Lunde B N, DeHaan, C J and Wallmark J A (2014) The Navy's Application of Ocean Forecasting to Decision Support. *Oceanography* 27(3): 126-137
31. Catlett C, Allcock W E, Andrews P, Aydt R A, Bair R, Balac N and Marsteller J (2006) TeraGrid: Analysis of Organization, System Architecture, and Middleware Enabling New Types of Applications. *High Performance Computing Workshop* 16:225-249
32. CERN - European Organization for Nuclear Research (2011) web, jul. 2011, URL: <http://public.web.cern.ch/public/>
33. Comunidad Europea (2010): *Assessing Europe's University-Based Research*, 2010
34. Cosnard M and Trystran D (1995) *Parallel Algorithms and Architectures*. International Thomson Computer
35. Davis N E, Robey W, Ferenbaugh C R Nicholaeff, D and Trujillo D P (2012) Paradigmatic shifts for exascale Supercomputing. *Journal of Supercomputing* 62(2): 1023-1044
36. De Filippo D, Morillo F and Fernández M T (2008) *Indicadores de colaboración científica del*

CSIC con Latinoamérica en base de datos internacionales. [Indicators of Scientific Collaboration between CSIC and Latin America in International Databases]. *Revista Española de Documentación Científica* 31(1): 66-84

37. Dongarra, J J, Bunch J R, Moler C B and Stewart G W (1979). LINPACK users' guide. 8
38. Dongarra J J, Foster I, Fox G and Gropp W (2002) *The Sourcebook of Parallel Computing*. Morgan Kaufmann
39. Dongarra J J and Van Der Steen A J (2012) *High-performance Computing Systems: Status and Outlook*. *Acta Numérica* 21: 379-474
40. Ellingson S R, Smith J C and Baudry J (2014) Polypharmacology and Supercomputer-based Docking: Opportunities and Challenges. *Molecular Simulation* 40(1): 10-11
41. Elmagarmid A K, Samuel A and Ouzzani M (2008) Community-Cyberinfrastructure-Enabled Discovery in Science and Engineering. *IEEE Computing in Science and Engineering* 10(5): 46-53
42. Emmot S and Rison S (2005) *Towards 2020 Science Report*. Microsoft Research
43. Featherstone N A, Browning M K, Brun A S and Toomre J (2009) Effects of Fossil Magnetic Fields on Convective Core Dynamos in A-type Stars. *The Astrophysical Journal* 705(1): 1000
44. Fernbach S (1984) Supercomputers—Past, Present, Prospects. *Future Generation Computer Systems* 1(1): 23-30
45. Foster I, Kesselman C and Tuecke S (2001) “The Anatomy of the Grid: Enabling Scalable Virtual Organizations”. *International Journal of High Performance Computing Applications* 15(3): 200–222
46. Gaudiani A (2012) *Análisis del rendimiento de algoritmos paralelos de propósito general en GPGPU [Performance Analysis of Parallel Algorithms for General Purposes in GPGPU]* (Doctoral dissertation, Facultad de Informática). URL:  
[http://sedici.unlp.edu.ar/bitstream/handle/10915/22691/Documento\\_completo\\_\\_.pdf?sequence=1](http://sedici.unlp.edu.ar/bitstream/handle/10915/22691/Documento_completo__.pdf?sequence=1)  
Retrieved 2015 10 08
47. Geller T (2011) Supercomputing's exaflop target. *Communications of the ACM* 54(8): 16-18
48. Gengler M, Ubeda S and Desprez, F (1996) "Initiation au parallélisme: concepts, architectures et algorithmes" [Introduction to Parallelism: Concepts, Architectures and Algorithms]. Masson ISBN 2-225-85014-3
49. Goldberg D (1989) *Genetic Algorithms in Search, Optimization and Machine Learning*. Addison-Wesley
50. Guellec D and Van Pottelsberghe de la Potterie B (2001) *R&D and Productivity Growth: Panel Data Analysis of 16 OECD Countries*. STI Working Papers 2001/3. OECD
51. Guevara M, Martínez F, Arevalo G, Gassó, S and Baldasano J (2013) An Improved System for Modelling Spanish Emissions: HERMESv2.0. *Atmospheric Environment*. 81: 209-22
52. Hassan A H, Fluke C J and Barnes D G (2011) *Unleashing the Power of Distributed CPU/GPU*

architectures: Massive Astronomical Data Analysis and Visualization Case Study. arXiv preprint arXiv:1111.6661

53. Hauben M (2010) History of ARPANET. URL: <http://pages.infinet.net/jbcoco/Arpa-Arpanet-Internet.pdf> Retrieved 2010 23 13
54. Hermes H (1969) Enumerability. Decidability. Computability. Die Grundlehren der Mathematischen Wissenschaften in Einzeldarstellungen. Band 127. Springer-Verlag. Berlin, Heidelberg, New York, 1969
55. High Performance Computing Act of 1991 (HPCA). Act of Congress promulgated in the 102nd United States Congress as (Pub.L. 102–194) on 1991-12-09
56. Hwang K (1993) Advanced Computer Architecture: Parallelism, Scalability, Programmability. McGraw-Hill 1993
57. IDC (2004) White Paper: Council on Competiveness Study of U.S. Industrial HPC Users. URL: [http://www.compete.org/storage/images/uploads/File/PDF%20Files/HPC\\_Users\\_Survey%202004.pdf](http://www.compete.org/storage/images/uploads/File/PDF%20Files/HPC_Users_Survey%202004.pdf)
58. Kantardjiev A A (2012) Quantum.Ligand.Dock: protein-ligand docking with quantum entanglement refinement on a GPU system. Nucleic Acids Research 40(W1): W415-W422
59. Kennedy K and Joy W (1998) Interim Report to the President, President's Information Technology Advisory Committee (PITAC). National Coordination Office for Computing, Information and Communication, 4201 Wilson Blvd, Suite 690, Arlington, VA 22230, August 10, 1998
60. Kinter J L, Cash B, Achuthavarier D, Adams J, Alshuler E, Dirmeyer P and Wong K et al. (2013) Revolutionizing Climate Modeling with Project Athena. A Multi-Institutional, International Collaboration. Bulletin of the American Meteorological Society 94(2): 231-245
61. Kramer D (2011) Supercomputing has a future in clean energy. Physics Today 64(7): 27-29
62. Kupczyk M and Meyer N (2010) PRACE World-Class Computational Facilities Ready for Polish Scientific Community. Computational Methods in Science and Technology 2010 57-62
63. Kwon H I, Kim S, Lee H, Ryu M, Kim, T and Choi, S (2013) Development of an Engineering Education Framework for Aerodynamic Shape Optimization. International. Journal of Aeronautical and Space Sciences 14(4): 297-309
64. Lavington S H (1978) The Manchester Mark I and Atlas: a Historical Perspective. Communications of the ACM 21(1): 4-12
65. Lawson S J, Woodgate M, Steijl R and Barakos G N (2012) High Performance Computing for Challenging Problems in Computational Fluid Dynamics. Progress in Aerospace Sciences 52: 19-29
66. Lederman D and Maloney W (2003) R&D and Development. World Bank Policy Research Working Paper 3024
67. Lee E, Huang H, Dennis J M, Chen P and Wang L (2013) An Optimized Parallel LSQR Algorithm for Seismic Tomography. Computer & Geosciences. 61: 184-197

68. Leiner B M, Cerf V G, Clark D D, Kahn R E, Kleinrock L, Lynch D C and Wolff, S. (2009). A Brief History of the Internet. *ACM SIGCOMM Computer Communication Review*, 39(5), 22-31
69. Lepora N F, Overton P G and Gurney K (2012) Efficient Fitting of Conductance-based Model Neurons from Somatic Current Clamp. *Journal of Computational Neuroscience* 32(1): 1-24
70. Lingerfelt E J, Messer O E, Desai, S S, Holt C A and Lentz E J (2014) Near Real-time Data Analysis of Core-collapse Supernova Simulations with Bellerophon. *Procedia Computer Science* 29: 1504-1514
71. Markram H, Meier K, Lippert T, Grillner S, Frackowiak R, Dehaene S and Saria A (2011). Introducing the Human Brain Project. *Procedia Computer Science* 7: 39-42
72. Martin H S C, Jha S and Coveney P V (2014) Comparative Analysis of Nucleotide Translocation through Protein Nanopores Using Steered Molecular Dynamics and an Adaptive Biasing Force. *Journal of Computational Chemistry* 35(9): 692-702
73. Matsuzaki Y, Uchikoga N, Ohue M, Shimoda T, Sato T, Ishida T and Akiyama Y (2013) MEGADOCK 3.0: A High-performance Protein-protein Interaction Prediction Software using Hybrid Parallel Computing for Petascale Supercomputing Environments. *Source code for biology and medicine* 8: 18
74. Mc. Cartney S (1999) ENIAC: The Triumphs and Tragedies of the World's First Computer. Walker & Company. ISBN:0802713483
75. Meuer H W and Gietl H (2013) Supercomputers—Prestige Objects or Crucial Tools for Science and Industry?. *PIK-Praxis der Informationsverarbeitung und Kommunikation* 36(2): 117-128
76. Menhorn F and Reumann M (2013) Genome Assembly Framework on Massively Parallel, Distributed Memory Supercomputers. *Biomedical Engineering/Biomedizinische Technik* 2013
77. Metcalfe T S, Mathur S, Dogan G and Woitaszek M (2012) First Results from the Asteroseismic Modeling Portal. *Progress in Solar/Stellar Physics with Helio- and Asteroseismology* 462: 213
78. Milone D, Azar A and Rufiner L (2002) “Supercomputadoras basadas en "clusters" de PCs. Trabajo de desarrollo tecnológico realizado en el Laboratorio de Cibernética de la Facultad de Ingeniería (UNER)” [“Supercomputers based on PC "clusters." Technological Development Project Carried out in the Cybernetic Laboratory of the Engineering Faculty]. *Revista Ciencia, Docencia y Tecnología*, Vol. XIII, No. 25, pp. 173-208, nov. 200268
79. Moraleda A (2007) Supercomputing: a Qualitative Leap for Competitiveness. *Economistas* 26(116): 294-297
80. Munetomo M (2011) Realizing Robust and Scalable Evolutionary Algorithms toward Exascale era. *Evolutionary Computation (CEC), 2011 IEEE Congress on* (pp. 312-317). IEEE
81. National Academy of Sciences (2005) Getting up to Speed the Future of the Future of Supercomputing. Committee on the Future of Supercomputing Computer Science and Telecommunications Board Division on Engineering and Physical Sciences
82. Neic A, Liebmann M, Hoetzel E, Mitchell L, Vigmond, E J, Haase G and Plank, G (2012) Accelerating Cardiac Biodomain Simulations Using Graphics Processing Units. *IEEE*

83. Nesmachnow S (2014) Planificación de tareas en sistemas cluster, grid y cloud utilizando algoritmos evolutivos. [Scheduling in Cluster Systems, Grids and Clouds using Evolutionary Algorithms]. *Revista de la Sociedad Mexicana de Inteligencia Artificial* 18 / 36 1: Jan - April 2014
84. Nicolosi F, Della Vecchia P and Ciliberti D (2013) An Investigation on Vertical Tail Plane Contribution to Aircraft Sideforce. *Aerospace Science and Technology* 28(1): 401-416
85. O'Neill J E (1995) The Role of ARPA in the Development of the ARPANET, 1961-1972. *Annals of the History of Computing, IEEE*, 17(4): 76-81
86. Patterson C A, Snir M and Graham S L (2005). *Getting Up to Speed: The Future of Supercomputing*. National Academies Press
87. Perine K (2000) The Early Adopter - Al Gore and the Internet - Government Activity. *The Industry Standard*
88. Puckelwartz M J, Pesce L, Nelakuditi V, Dellefave-Castillo L, Golbus J R, Day S M and McNally, E M (2014) Supercomputing for the Parallelization of Whole Genome Analysis. *Bioinformatics* 30(11): 1508-1513
89. RedCLARA (2011) Compendio RedCLARA de Redes Nacionales de Investigación y Educación Latinoamericanas. [CLARA Compendium of Latin American National Research Networks and Education]. Bogotá: RedCLARA
90. Reumann M., Makalic E, Goudey B W, Inouye M, Bickerstaffe A, Bui M and Hopper J L (2012) Supercomputing Enabling Exhaustive Xtatistical Analysis of Genome wide Association Study Data: Preliminary results. *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE* (pp. 1258-1261). IEEE
91. Romer P (1990) Endogenous Technological Change. *Journal of Political Economy* 98(5): 71-102
92. Rosenberg L C (1991) Update on National Science Foundation funding of the "Collaboratory". *Communications of the ACM*, 34(12): 83
93. Saini S, Rappleye J, Chang J, Barker D, Mehrotra P and Biswas R (2012) I/O Performance Characterization of Lustre and NASA Applications on Pleiades. *High Performance Computing (HiPC), 2012 19th International Conference on* (pp. 1-10). IEEE
94. Sanz M A (1998) Fundamentos históricos de la Internet en Europa y en España. [Historical Foundations of the Internet in Europe and Spain]. *Boletín Rediris* 45
95. Saunders V R and Guest M F (1982) Applications of the CRAY-1 for quantum chemistry calculations. *Computer Physics Communications* 26(3): 389-395
96. Sawyer M and Parsons M (2011) *A Strategy for Research and Innovation through High Performance Computing*. The University of Edinburgh 2011

97. Schaller R (1997) Moore's Law: Past, Present and Future. *Spectrum*, IEEE 34(6): 52-59
98. Shen B, Nelson B, Cheung S and Tao W K (2013) Improving NASA's Multiscale Modeling Framework for Tropical Cyclone Climate Study. *Computing in Science & Engineering* 15(5): 56-67
99. Simmons M L and Wasserman H J (1990) Performance comparison of the CRAY-2 and CRAY X-MP/416 supercomputers. *The Journal of Supercomputing* 4(2): 153-167
100. Singh S (2000) *The Code Book: the Secret History of Codes and Code-breaking*: 77-85. London: Fourth Estate
101. Skupin A, Biberstine J and Boerner K (2013) Visualizing the Topical Structure of the Medical Sciences: A Self-Organizing Map Approach. *Plos One* 8(3)
102. Soete L, O'Doherty D, Arnold E, Bounfour A, Fagerberg J, Farinello U and Schiestock G (2002) Benchmarking National Research Policies: The Impact of RTD on Competitiveness and Employment (IRCE). A Strata-ETAN Expert Working Group, European Commission DG Research, Brussels
103. Stamatakis A, Aberer A J, Goll C, Smith S A, Berger S A and Izquierdo-Carrasco F (2012) RAXML-Light: a Tool for Computing Terabyte Phylogenies. *Bioinformatics* 28(15): 2064-2066
104. Stein R F, Lagerfjård A, Nordlund Å and Georgobiani, D (2012) Helioseismic Data from Emerging Flux Simulations. *Progress in Solar/Stellar Physics with Helio-and Asteroseismology* 462: 345
105. Stone H (1993) *High Performance Computer Architectures*. Addison Wesley
106. Stoyanov D, Grigorov I, Deleva, A, Kolev N, Peshev Z, Kolarov G and Ivanov D (2013) Remote Monitoring of Aerosol Layers over Sofia during Sahara Dust Transport Episode (April, 2012). In *Seventeenth International School on Quantum Electronics: Laser Physics and Applications* (pp. 87700Y-87700Y). International Society for Optics and Photonics
107. Sumiyoshi K. (2011) A numerical Challenge on the Core-collapse Supernovae: Physics of Neutrino and Matter at Extreme Conditions. *Journal of Physics: Conference Series* 302(1)
108. Thornton J E (1970) Design of a Computer—The control data 6600
109. Toomre J, Augustson K C, Brown B P, Browning M K, Brun A S, Featherstone N A and Miesch M S (2012) New Era in 3-D Modeling of Convection and Magnetic Dynamos in Stellar Envelopes and Cores. *Progress in Solar/Stellar Physics with Helio-and Asteroseismology* 462(331)
110. Turpin T, Lian Y, Tong J and Fang X (1995). Technology and Innovation Networks in the People's Republic of China. *Journal of Industry Studies*, 2(2): 63-74
111. Um J, Choi H, Song S K, Choi S P, Mook Yoon H, Jung H and Kim T H (2013) Development of a Virtualized Supercomputing Environment for Genomic Analysis. *Journal of Supercomputing* 65(1): 71-85

112. Utreras F (2014) *Visión y Proyectos: el presente y el futuro*. [Vision and Projects: Present and Future]. XI Encuentro Temático Nacional Renata – RUP. RedClara
113. Villarubia C (2012) “La Comisión Europea dobla el presupuesto para HPC” [“The European Commission Doubles the Budget for HPC”]. URL:  
<http://www.bsc.es/sites/default/files/public/about/news/hpc-22022012-datacenterdynamics.pdf>,  
 Retrieved: 27/08/2014
114. Von Neumann J (1945) First Draft of a Report on the EDVA. Between the United States Army Ordinance Department and the University of Pennsylvania Moore School of Electrical Engineering University of Pennsylvania. June 30, 1945. Contract No. W-670-ORD-4926
115. Wang R and Harris C (2013) Scaling Radio Astronomy Signal Correlation on Heterogeneous Supercomputers using Various Data Distribution Methodologies. *Experimental Astronomy* 36(3): 433-449
116. Westerlund , S. and Harris, C. (2014). A Framework for HI Spectral Source Finding Using Distributed-Memory Supercomputing . *Publications of the Astronomical Society of Australia*, 31, e023 doi:10.1017/pasa.2014.18.
117. Wilkes M V and Renwick W (1950) The EDSAC (Electronic Delay Storage Automatic Calculator). *Mathematics of Computation* 4(30): 61-65
118. Winkel M, Speck R, Huebner H, Arnols L, Krause R and Gibbon P (2012) A Massively Parallel, Multi-disciplinary Barnes-Hut Tree Code for Extreme-scale N-body Simulations. *Computer Physics Communications* 183(4): 880-889
119. Woolgar L (2007) New Institutional Policies for University-Industry Links. *Japan Research Policy* 36: 1261-1274
120. Wu W (2007) Cultivating Research Universities and Industrial Linkages in China: The Case of Shanghai. *World Development*, 35(6):1075-1093
121. Wu Y and Cai X-C (2014) A fully Implicit Domain Decomposition based ALE Framework for Three-dimensional Fluid-structure Interaction with Application in Blood Flow Computation. *Journal of Computational Physics* 258: 524-537
122. Yamamoto, K. (2004). “Corporatization of National Universities in Japan: Revolution for Governance or Rhetoric for Downsizing?”. *Financial Accountability & Management* 20(2): 153-181
123. Yang W, Yoshigoe K, Qin X, Liu J S, Yang J Y, Niemierko A, Deng Y, Liu Y, Dunker A K, Chen Z, Wang L, Xu D, Arabnia H R, Tong W. and Yang, M. Q. (2014) Identification of Genes and Pathways Involved in Kidney Renal Clear Cell Carcinoma. *BMC bioinformatics*, 15(Suppl 17), S2
124. Yin S, Luo H and Ding S (2013) Real-time Implementation of Fault-tolerant Control Systems with Performance Optimization. *Industrial Electronics, IEEE Transactions* 61(5): 2402 – 2411.
125. Yoshida H (2013) Cloud-super-computing Virtual Colonoscopy with Motion-based Navigation for Colon Cancer Screening. *Consumer Electronics?? Berlin (ICCE-Berlin)*, 2013.

ICCE Berlin 2013. IEEE Third International Conference on (pp. 165-167). IEEE.

126. Zhou P and Leydesdorff L (2006). The Emergence of China as a Leading Nation in Science. *Research Policy*, 35(1): 83-104.