◆◆ Scientific
◆◆ Research

G-Phenomena as a Base of Scalable Distributed Computing —G-Phenomena in Moore's Law

Karolj Skala, Davor Davidović, Tomislav Lipić, Ivan Sović

Centre for Informatics and Computing, Ruđer Bošković Institute, Zagreb, Croatia
Email: skala@irb.hr

ABSTRACT

Today we witness the exponential growth of scientific research. This fast growth is possible thanks to the rapid development of computing systems since its first days in 1947 and the invention of transistor till the present days with high performance and scalable distributed computing systems. This fast growth of computing systems was first observed by Gordon E. Moore in 1965 and postulated as Moore's Law. For the development of the scalable distributed computing systems, the year 2000 was a very special year. The first GHz speed processor, GB size memory and GB/s data transmission through network were achieved. Interestingly, in the same year the usable Grid computing systems emerged, which gave a strong impulse to a rapid development of distributed computing systems. This paper recognizes these facts that occurred in the year 2000, as the G-phenomena, a millennium cornerstone for the rapid development of scalable distributed systems evolved around the Grid and Cloud computing paradigms.

KEYWORDS

Historical Development of Computing; G-Phenomena; Moore's Law; Distributed Computing; Scalability; Grid Computing; Cloud Computing Component

1. Background Overview

The first computers date back to the 1822 when Charles Babbage constructed a mechanical computer that was capable of computing several sets of numbers and making hard copies of the results. This invention was far away from what we now consider as computer. The first electrical computer, named Z1, was designed in 1936 by Konrad Zuse. However, the first concept of modern computers was set by Alan Turning in 1936 with the invention of his famous Turing machine. That invention started the era of development of modern computer hardware and software. The first "modern" computers used vacuum tubes and occupied a lot of space which were hard to program and consumed a lot of energy. The most important breakthrough was the invention of transistor in 1947 by John Bardeen, Walter Brattain, and William Shockley for which the authors won the Nobel Prize in Physics in 1956. Transistors revolutionized electronics and computer science from the roots and paved the way to cheaper, smaller and therefore much more affordable computers. At that time, the rule of thumb was that more transistors on a die lead to better performance but today that is not the case. Today, the performance is achieved by other factors (e.g. number of CPU cores, vectorisation, etc.) than relying only on the number of transistors per processor.

Since the invention of the integrated circuit in 1959 and microprocessors in 1970 [1], the number of transistors that can be placed in an integrated circuit has increased rapidly. This trend was first observed by the Intel Corporation co-founder Gordon E. Moore in 1965 [2]. He observed that the number of transistors increases at a rate of roughly a factor of two per year, effectively doubling computer processor performance. The increase of the raw performance can easily be followed on the list of Top 500 supercomputers [3] in the world. The fastest supercomputer in 1993 achieved 100 GFlops (giga floating point operations per second), while in 2013 the fastest supercomputer achieved amazing 16 PFlops, which

presents an increase by a factor of 2 every 18 months. It is known that by the increasing number of transistors and the frequency of processors, the energy consumption is increased as well. As the cost of energy is constantly increasing, more attention is put on energy efficient computing. More interestingly, Koomley showed that energy efficiency in computing has also doubled every year and a half, which enabled the creation of mobile computing systems (e.g. laptops, smart phones or wireless sensors) [4]. This technology progress enabled the ubiquity of computing in our everyday life.

The progress of computing systems is usually described in decade phases [5]. In the first phase that began in the 1960s, computing was limited to large, room-sized mainframe computers, not affordable to consumer public. The second phase (1970s) is marked by the development of minicomputers, somewhat smaller, faster and more affordable. The third phase (1980s) includes the appearance of personal computers (PCs), even faster computers intended for wide consumer use. The fourth phase (1990s) is marked by the creation of Internet and the availability of cheap servers which made it possible to distribute data around the world. Finally, the fifth phase (2000s) begins the period of the ubiquitous distributed computing systems.

This period is characterized by the large-scale networking of various heterogeneous distributed computing and storage resources that can easily be accessed from anywhere in the world. The evolution of these distributed computing systems has continued through Grid computing and Cloud computing paradigms. These paradigms are the main anchors for the development of large-scale applications from many fields in science and engineering.

2. Recognition of the G-Pheniomena

We observe the phenomena that started the fifth phase of development of computing systems. The year 2000 was the first year when processor clock rate, network bandwidth and RAM (random-access memory) memory capacity aligned in the giga range. Specifically, 1-GHz processors were released by Intel and AMD, gigabit Ethernet was in use and the first computers with 1GB RAM became available. This alignment made a virtual integration of spatially distributed computers plausible, further enabling rapid development of distributed systems, thus creating conditions for space-independent concept of distributed systems. We term this phenomena the G-phenomena, which introduces new pragmatic concepts in distributed computing and communications.

The decade phases of development of Information and Communication Technologies (ICT) are given in **Figure 1**. The figure shows projections of the increase in processor clock frequency, network bandwidth and memory capacity throughout the five phases, in concordance to the Moore's law.
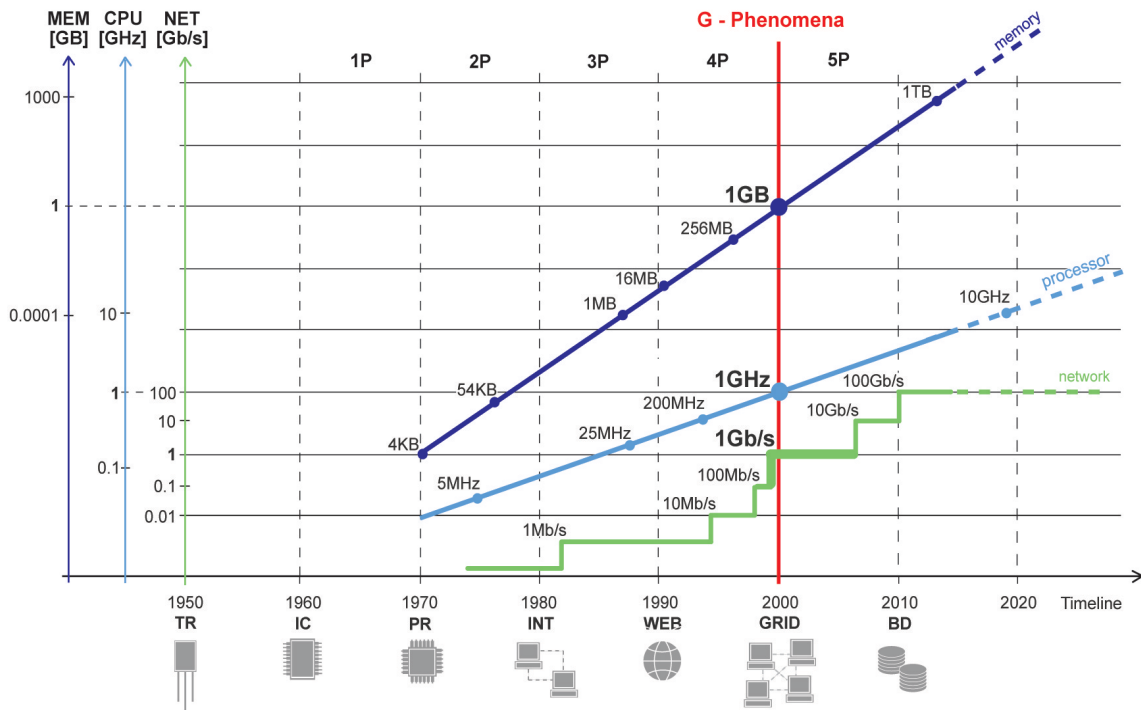
G (GIGA) - Phenomena, TR - Transistor, IC - Integrated Circuit, PR - Processor, INT - Internet, BD - Big Data, 1P...5P - Development Phases

Figure 1. **The decade phases of the development of ICT and its relations to the G-phenomena.**

Post G-Phenomena Era: High-Performance and Scalable Distributed Computing Systems

After the phenomena that occurred in 2000, which we have denoted as the G-phenomena, the era of the fast development of high-performance and scalable distributed computing systems started. The primary predisposition for that development was achieving substantial speed improvements of processors and their interconnections, and the ability to process more data in memory. High-performance distributed computing systems were founded on Grid computing paradigm [6] while scalable distributed computing systems evolved through Cloud computing paradigm.

Grid Computing is a distributed paradigm suitable for scientific high performance computing. It emerged as an alternative to expensive supercomputers by virtually connecting a large number of computational clusters formed from computers connected through fast local network [7,8]. It is especially suitable for enforcement of large-scale many-task applications including High-Throughput and Data-Intensive applications. Although the first grid infrastructures were started in late 1990s through several Grid-oriented projects in the United States, the two major European projects, started in early 2000s, the UK's e-Science program and European Union, Data Grid project [9], gave the shape of the European grid initiative. The availability of public Grid infrastructures stimulated the expansion of scientific computing research, and progressively through time Grid computing became one of the major paradigms for scientific applications. However, high utilization of public Grid infrastructures by many different groups with different applications, their technical and bureaucratic issues limited their widespread usage.

Cloud Computing has been defined as a "distributed computing paradigm where the boundaries of computing will be determined by economic rationale rather than technical limits". Cloud computing enables efficient management of data centers, time-sharing, and virtualization of resources with a special emphasis on the business model [10]. Through this paradigm, users and businesses are able to consume Infrastructure-as-a-Service, Platform-as-a-Service and Software-as-a-Service service models on demand from anywhere in the world [11]. When Cloud's services are made available by Cloud provider in a pay-as-you-go manner to the general public it is referred to as Public Cloud while Private Cloud refers to internal Cloud platforms owned and utilized only by particular organizations. Unlike Grids, Clouds can scale up or down according to the requirements of users and at the same time enable more flexible and isolated execution environments for applications.

Over the past few years an increasing number of companies, especially from telecommunication and IT sector, are moving from static, centralized cluster environments to more elastic, scalable and potentially cheaper Cloud platforms. The fast growing SMEs have needs to quickly set-up required resources with minimal costs and time spent, but do not have the capital or the time to invest in deploying their own resources. In other words, by moving from the usual capital upfront investment model to an operational expense, cloud computing promises, especially to SMEs and entrepreneurs, to accelerate the development and adoption of innovative solutions.

Similarly, many branches of modern research are computation- and data-intensive forms of discovery, encompassing the generation, analysis and interpretation of vast amounts of data against catalogues of existing knowledge in complex multi-stage workflows. These workflows, or analyses, are enabled by a combination of analysis platforms and computational infrastructure, which a Cloud can provide, while additionally offering scalable resources on demand.

3. Conclusion

G-phenomena, as an observation of Giga measure of different parameters at the turn of the millennial transition, reaches new changes in the emergence of distributed systems. Started by the Grid and Cloud computing paradigms, it has led to the current development of ICT technologies. This has dramatically increased the generic role of electronics and digital network computing in nearly every segment of the world, in science, economy and everyday life. Thus, G-phenomena can be observed as a basis of scalable distributed computing systems experienced as a millennium cornerstone.

G-phenomena is the result of the exponential development rate of computing and related technologies over the past 50 years. This development is the most prominent driving force of the human society. It is expected that new computing technologies will continue to emerge, such as today's researched photonic and quantum computing paradigms, that will make computers even more powerful than all the other current computers combined.

Acknowledgements

REFERENCES

[1] A. Osborne, "An Introduction to Microcomputers. Volume 1: Basic Concepts," 2nd Edition, Osborne-McGraw Hill, Berkely, 1980.

[2] G. E. Moore, "Cramming More Components onto In-

tegrated Circuits," *Electronics Magazine*, 1965.

[3]   Top500, "Top 500 Supercomputer Sites," 2013.
      http://www.top500.org/

[4]   J. Koomey, S. Berard, M. Sanchez and H Wong, "Implications of Historical Trends in the Electrical Efficiency of Computing," *IEEE Annals of the History of Computing*, Vol. 33, No. 3, 2011, pp. 46-54.
      http://dx.doi.org/10.1109/MAHC.2010.28

[5]   J. M. Gallaugher, "Moore's Law & More: Fast, Cheap Computing, and What It Means for the Manager," 2008.
      http://www.gallaugher.com/Moore%27s%20Law%20&%20More.pdf

[6]   I. Foster, C. Kesselman and S. Tuecke, "The Anatomy of the Grid: Enabling Scalable Virtual Organizations," *International Journal of High Performance Computing Applications*, Vol. 15, No. 3, 2001, pp. 200-222.
      http://dx.doi.org/10.1177/109434200101500302

[7]   I. Foster and C. Kesselman, "The Grid: Blueprint for a New Computing Infrastructure," Elsevier, Amsterdam, 2003.

[8]   F. Berman, G. Fox and T. Hey, "Grid Computing, Making the Global Infrastructure a Reality," Wiley, Hoboken, 2003.

[9]   The DataGrid Project, "The DataGrid Project," 2004.
      http://eu-datagrid.web.cern.ch/eu-datagrid/

[10] A. Fox, R. Griffith, A. Joseph and R. Katz, "Above the Clouds: A Berkeley View of Cloud Computing," University of California, Berkeley, 2009.

[11] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg and I. Brandic, "Cloud Computing and Emerging IT Platforms: Vision, Hype, and Reality for Delivering Computing as the 5th Utility," *Future Generation Computer Systems*, Vol. 25, No. 6, 2009, pp. 599-616.
      http://dx.doi.org/10.1016/j.future.2008.12.001