ICT infrastructure and Scientific Research

Kiril Boyanov, Dimitar Todorov

Abstract: *The paper discuses the trends in contemporary science research to massively use the Information and Communication Technologies (ICT) for their regular activities. The significance of ICT infrastructure as preliminary condition for successful research is analysed. The main European ICT infrastructure projects and their status in Bulgaria are presented. The conditions for successful development of ICT infrastructure for scientific research in Bulgaria are discussed.*

Key words: *Information and Communication Technologies, ICT, ICT infrastructure, GRID, networks, supercomputing, distributed resources.*

THE EUROPEAN AND WORLD ECONOMIC AND RESEARCH LANDSCAPE

During the last years, the European economic and research landscape is in flux. Europe and the rest of the industrialised world can no longer take their technological leadership for granted. In the research and technological development Europe couldn't attract the best talents from around the world. Its performance, in terms of growth, productivity and job creation is not sufficient to maintain prosperity in the future. There is a broad consensus that research, education and innovation are at the heart of any response to these challenges.

THE IMPACT OF ICT ON CONTEMPORARY ECONOMIC DEVELOPMENT

The investment in ICT contributes to 0,3-0,8% of annual average GNP growth. The investment in software accounted for up to a third of the overall contribution of ICT investments for GNP growth. The ICT-manufacturing sector is an important driver of the acceleration in productivity growth but is highly concentrated because of large economies of scale and with high entry costs. The ICT-producing services sector has more universal potential for driving productivity growth in a broader range of countries. In Bulgaria, the growth of services due to ICT usage is not analysed but is very probable that it has happened with some lag, corresponding to the late ICT entry at all.

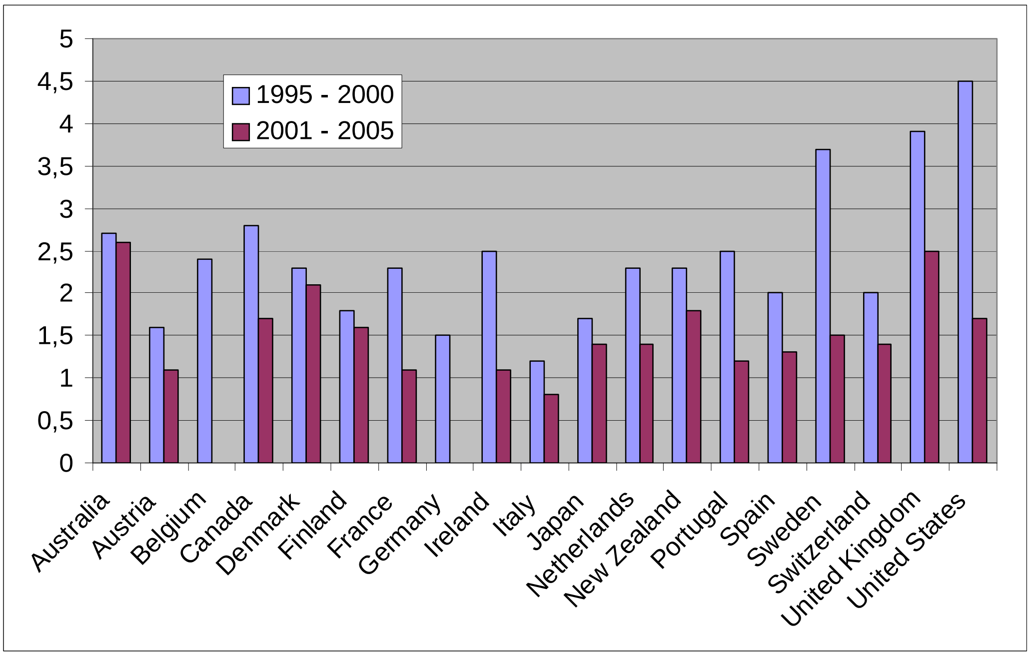

Figure 1: The impact of ICT based equipment on annual Growth of Services.

THE CHARACTERISTICS OF CONTEMPORARY SCIENCE AND RESEARCH

The realisation of the science and research projects today is related to the large volumes of row data in almost all science domains with further exponential increase. The interdisciplinary collaboration is a main driving force. Both the increasing progress on mathematical models and the complexity of simulations cause the demand for almost unlimited computing power. The infrastructure for the realisation of scientific research, more and more relies on ICT for the discoveries, developments and applications of new technologies.

The term **e-Infrastructure** is widely used to indicate the ICT-based and network integrated Infrastructure for the implementation of Science and Research activities. The organization of the present and future research activity is shown in Figures 2 and 3.

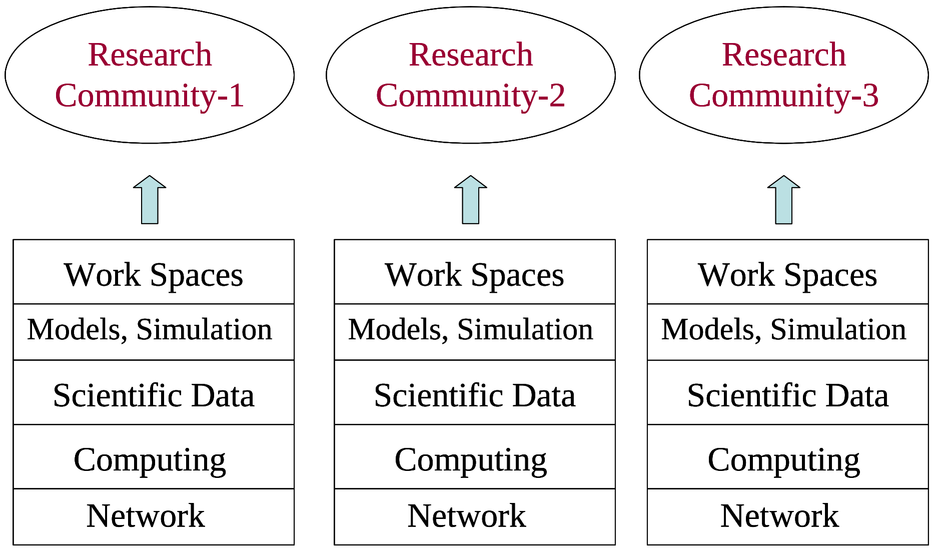

Figure 2: Organisation of Research activity till now.

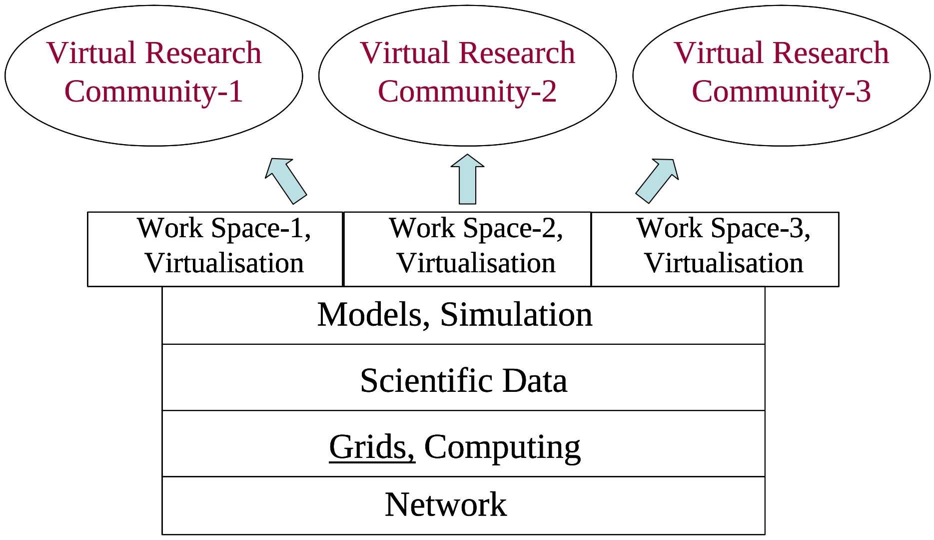

Figure 3: Organisation of Research activity in the near future.

There is no doubt that e-Infrastructure impact will be far beyond science activities as it has been already happened with Internet.

SOME ASPECTS OF EUROPEAN RESEARCH POLICY AND ICT

The well known goal of the European research policy is to establish a common European Research Area and to increase the European research effort to 3% of the European Union's GDP, two-thirds coming from private investment and one-third from the public sector. The research effort of Europe today is 2% which is far less from the United States (2.8%) and Japan (more than 3%).

Another significant aspect of the European research policy is to strengthen the European research effort. The contemporary Research is becoming more and more expensive and the integration at European level is a must, especially in highly competitive sectors such as ICT, biotechnology and nanotechnology, aeronautics and hydrogen energy technology. The goal is to have "European added value" through establishing of a "critical mass" of resources, competition on a European level and trans-national collaboration.

The established European Framework Programmes aim to make easy existing financial needs for research activities. The realisation of the goals of the European research policy today is undoubtedly delayed by the world financial and economic crisis but continues to be without any other alternative.

Europe is in a relatively good position in some key ICT domains like microelectronics, mobile and fixed communication, consumer electronics, etc. It has, however, some weaknesses – for example almost no computer and generic software industry. The evolution of ICTs and their increasing pervasiveness offer some opportunities to catch up and develop new leaderships in the field of ICT based infrastructure.

MORE DETAILS ON E-INFRASTRUCTURE

The e-Infrastructure viewpoint allows all interrelated ICT based infrastructures to be perceived as a whole. The prime goal of the e-Infrastructure is to support e-science, e-health and e-culture but it gives opportunities for other domains like e-commerce, e-government, e-training and e-education.

A competitive e-Infrastructure is indispensable for the numerically oriented branches of the sciences like:

- climate and earth system research, water management;
- fluid dynamics, biophysics, theoretical chemistry;
- astrophysics, quantum dynamics, nanostructure physics and high-energy physics.

In the traditionally less computer oriented areas such as the social sciences, the humanities and biodiversity, there is also a strong trend towards mass deployment of ICT to manage the large variety of decentralised data sources and find novel approaches to traditional problems.

The well known e-Infrastructure diagram is shown in Figure 3. The key components of the e-Infrastructure are:

- Networking infrastructures
- Middleware and organisation
- Distributed Resources
- Data and storage environment

In the diagram, the network is at the heart of everything. The middleware and virtual organisations connect the distributed resources, data and storage facilities in a seamless way. The application domains like e-Science, e-Health and e-Commerce exemplify the parties served by this integrated infrastructure

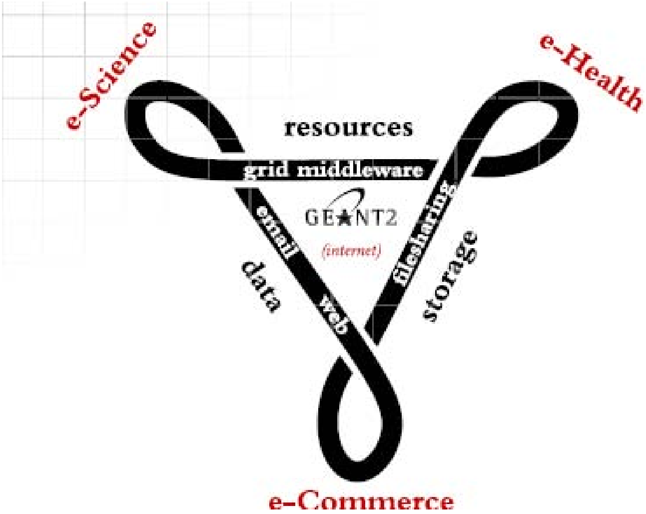

Figure 3: The e-Infrastructure diagram

In Europe the vision for e-Infrastructure is focused on computer networks and Grid technologies. The existent e-Infrastructure was initially made for research activity provision, now it provides an experimental base for realisation of similar infrastructures in the industry, administration, health etc.

Some more details on European e-Infrastructure key components are given below.

The Networking infrastructure delivers physical high speed connections through the GEANT pan-European backbone network and fine grained National Research and Education Networks (NREN). These networks form a basis for scientific communication like general purpose communications, specialised communications for ensuring of Grid computing, distributed supercomputing applications etc. Other networks could be added for linking the resources outside the scientific domain such as public resources, military or commercial resources etc.

The Middleware and organisation facilitates the integration of individual infrastructure components into European Science Grid. Grids are an evolutionary step in the way of usage of distributed computing resources and everything connected to them. European Science Grid is a vision for a pool of resources that are available to the users through their local work places and which are perceived as a whole. The middleware assumes a network and delivers a set of services for common usage of data, programmes, computers, etc.

The Resources are everything that could be of interest to science like computers, large storage facilities, telescopes, satellites, special physics equipment, weather balloons, lasers, spectrometers, visualisation means, large sensor networks, etc. A resource can also refer to a program or an artificial intelligence agent and even people as support organisations that can be shared between institutes. The only requirement to the resource properties is the information exchange to be through standard interfaces, i.e. grid protocols. The goal is to establish an ecosystem of resources that offer hardware, software, services and data spaces.

The Data is a special type of resource that is in the same time a fundamental property of information age. For effective data handling and managing with their exponential growth a proper data life cycle management infrastructure is essential.

THE DEVELOPMENT OF E-INFRASTRUCTURE IN EUROPE

There is e-Infrastructure Reflection Group (**e-IRG**), which main objectives are to support the creation of a political, technological and administrative framework for the easy and cost-effective shared use of distributed computer resources across Europe like grid computing, storage and networking. E-IRG includes government representatives and experts from European countries and identifies the fundamental needs for resources and services to enable pan-European e-Science. At the same time, it recommends resource sharing policy guidelines to National Grid initiatives and Regional & European e-Infrastructure projects. E-IRG contributes to International policy in that domain and gives input to other policy drafting bodies on European level. It focuses first on e-Science application user groups but also address wider application domains like e-Learning, e-Education etc. E-IRG identifies, informs and promotes GRID usage among communities who can benefit from sharing of their resources and addresses Governance issues of Grid deployment and usage.

There is a consensus that the next e-Infrastructure services should be delivered with a sustainable quality of Network connectivity, Supercomputer resources usage and Grid resources usage. For that reason the e-Infrastructure funding schemes are transforming from project oriented to more time independent types that rely on coordination on international and national level and on usage of new funding models based on pan-European practice. The principal challenge for the e-infrastructure service delivering is the planning of their transformation in paid utility services. A special attention should be paid that the premature transformation leads to slowdown of e-infrastructure development and distribution and the late transformation leads to unnecessary exploitation expenses.

SOME SIGNIFICANT EUROPEAN INFRASTRUCTURE INITIATIVES

The **GÉANT** project is the infrastructure foundation of the scientific research in Europe. GÉANT is a multi-gigabit research and education network which provides state-of-the-art data communications to the European research and educational community with leading-edge standards of reliability and innovation and the most advanced services and widest geographical reach of any network of its kind in the world. The GÉANT network connects 30 European NRENs across 34 countries. The NRENs connect research and educational institutions within their respective countries, providing connectivity to more than 30 million research and education end users in over 3,500 institutions across Europe. The network backbone throughput is 10 Gbps with plans to be 40 Gbps.

The Bulgarian Research and Educational Network (BREN) is building up and developing high speed communicational network infrastructure for the BAS institutions, the Universities and the Schools in Bulgaria. In 2008 as part of GÉANT2 project in IPP-BAS is deployed a GEANT backbone Point of Presence (PoP) which is connected by 10 Gbps communication lines with Athens, Bucharest and Budapest and by 2.5 Gbps communication line with Istanbul. The line speed to the BREN is 1 Gbps which is 10 times below the speed of other European Research and Educational Networks. The BREN line speed to the Bulgarian Universities and BAS Institutes is 100 Mbps and is several 10 times below the speed of scientific organisations of Europe.

The **EGEE** (Enabling Grids for E-sciencE) project is funded by the European Commission and its main aims are to build a secure, reliable and robust Grid infrastructure, to attract and engage a wide range of users from science and industry and to provide the users with extensive technical and training support. The EGEE project scale and scope (http://gridportal.hep.ph.ic.ac.uk/rtm/) is:

- Computer Clusters – 267;
- Countries – 54;
- Processor cores – about 114000;
- Volume of Storage environment – 20 PB;
- Registered virtual Organisations – more than 152;
- Registered users – more than 16000;
- Jobs per day – more than 150000;
- Application domains – more than 15.

The Bulgarian Grid infrastructure includes 6 Grid clusters with a total of 390 processor cores. The clusters ensure free usage access for Bulgarian Research Community for applications demanding significant computing resources.

In the second half of 2009, a new Grid cluster will be delivered and deployed in BAS with 40 worker nodes and a total of 320 processor cores, 30 TB disk storage and Infiniband connectivity between cluster units.

The location and processor cores of Bulgarian GRID clusters are summarized in Figure 4. For reference, a numbers of processor cores of Grid infrastructures in some European countries are shown in Figure 5.

Site Name	Location	Processor Cores
BG01-IPP	IPP-BAS	22
BG02-IM	IM-BAS	20
BG03-NGCC	IPP-BAS	200
BG04-ACAD	IPP-BAS	80
BG05-SUGrid	Sofia University	24
BG06-GPHI	GPHI-BAS	44
BG-INRNE	INRNE-BAS	20
Total Processor Cores:		410

Figure 4: Clusters in the Bulgarian Grid infrastructure.

Country	Number of Processor Cores
UK	15 551
France	13 326
Italy	10 814
Germany	7 918
Poland	7 092
Spain	4 172
The Netherlands	3 995
Greece	1 187
Romania	801
Serbia	743
Czech Republic	672
Bulgaria	410

Figure 5: Number of processor cores of some national Grid infrastructures in Europe.

DEISA (Distributed European Infrastructure for Supercomputing Applications) is a EC funded consortium of leading national super computer Centres. This is an open research infrastructure for usage in Grid environment which uses an internal network provided by

GEANT and the NRENs for connecting the supercomputers with reserved bandwidth. The DEISA main objectives are:

- To stimulate supercomputer usage for scientific research in Europe;
- To deploy and operate a persistent, production quality, distributed supercomputing environment with continental scope;
- To enable scientific discovery across a broad spectrum of science and technology;
- The real scientific impact is the only criterion for the initiative success.

The state-of-the-art of the supercomputing infrastructure in Bulgaria and its perspectives could be summarized as follows:

- In 2008 an IBM Blue Gene/P supercomputer was deployed in the State Agency for Information Technology and Communications (SAITC) in Sofia;
- The supercomputer IBM Blue Gene/P is and will be used for scientific research the field of nuclear physics, material research, medicine, drug discovery, education, etc.
- The supercomputer performance achieved on 23.42 TFLOPS which is ensured by 8192 processor cores, 4 TB random access memory, 12 TB disk storage in the near future will be deployed. At present 1 Gbps connectivity to the BREN is used but this will be increased in the near future.

SUMMARY OF BULGARIAN E-INFRASTRUCTURE PROBLEMS

- There is no budget funding for BREN;
- Low connectivity throughput – the trend for sustainable connection speed lags from 4 to 10 times;
- There is no budget funding for development and support of national Grid infrastructure.

The possible solutions for the above problems could be:

- SAITC to place part of state administration network backbone to BREN disposal;
- One-time funding for BREN backbone build up;
- Provision of 3 to 4 mill. leva annual budget funding for e-infrastructure support.

CONCLUSIONS

The e-infrastructure is in the base for scientific research and has substantial impact on economy of the country. Today the scientific results depend on the e-infrastructure level and appropriate strategy planning. At present, the e-infrastructure in the country has a chance to be developed on an acceptable level by reasonable state investments. This funding will enable active participation of the Bulgarian scientists in the EC projects, which will result a contemporary level of scientific activity. The present financial and economic crisis forces many young people to look on science as real alternative for their professional realisation, which preserves and increases their qualification and the presence of contemporary e-infrastructure gives them chances to choose Bulgaria for their scientific activity realisation.

REFERENCES

[1] OECD Information Technology Outlook 2008, OECD 2008, http://browse.oecdbookshop.org/oecd/pdfs/browseit/9308041E.PDF

[2] Boyanov, Kiril and Martikainen Olli,"Some Trends in Information and Communication Technology – the challenges for Europe, International Workshop "Next Generation Network Technologies", Rousse, Bulgaria, 11-12 October 2002

[3] The Global Information Technology Report 2008-2009 © 2009 World Economic Forum, http://www.insead.edu/v1/gitr/wef/main/fullreport/index.html

[4] Boyanov K., D. Todorov, Hr. Turlakov, "Computing Utilities", International Conference "Automatics and Informatics '08", 1-4 October 2008, Sofia, Bulgaria.

[5] GÉANT - a multi-gigabit pan-European data communications network project, http://www.geant.net/

[6] Enabling Grids for E-sciencE (EGEE) project, http://www.eu-egee.org/

[7] Distributed European Infrastructure for Supercomputing Applications (DEISA) project, http://www.deisa.eu/

ABOUT THE AUTHOR

Acad. Kiril Boyanov, Director of IPP-BAS, Sofia, e-mail: boyanov@acad.bg, http://www.bas.bg/clpp/en/indexen.htm

Assoc.Prof. Dimitar Todorov, Department DDSN, IPP – BAS, e-mail: dttod@acad.bg, http://www.acad.bg/rismim/en/indexen.htm