GRID, Virtual Observatory and Some Astronomical Applications

Georgi Petrov and Momchil Dechev
Institute of Astronomy and NAO, Bulgarian Academy of Sciences
petrov@astro.bas.org, mdechev@astro.bas.bg
(Submitted on 09.12.2014.; Accepted on 24.01.2015.)

Abstract.
The development of the GRID infrastructure gives new possibilities for treatment of heavy astrophysical calculations. Here some basic astronomical GRID projects, specialized software and applications are reviewed. Among the basic applications are N_body simulations, search for dark matter, study of large scale structure of the Universe, stellar and galaxy evolution, active processes on the Sun, near earth object discovery, specialized software, network telescopes architecture etc.
Key words: GRID, astrophysical applications

Introduction

During the 2009 - 2011 years the Bulgarian High performance computing GRID for advanced scientific applications under the National Scientific Fund DO 02-115/2009 grant was established. The first stage of the project is over. The grid system was configured and tested and it is under regular profit. The review presented here is based on the original published papers and internet information as well. The main goal is to point the possibilities of GRID for scientific investigations for the Bulgarian astronomical community.

A good start point is International Science Grid This Week (iSGTW) web site: http://www.isgtw.org/. iSGTW is an international, weekly, online science-computing newsletter that shows the importance of distributed computing, grid computing, cloud computing and high-performance computing. It reports about people and projects involved in these fields, and how these types of computing technologies are being applied to make scientific advances.

1 GRID Projects

Grid computing is a form of distributed computing whereby a "super and virtual computer" is composed of a cluster of networked, loosely-coupled computers, acting together to perform very large tasks. This technology has been applied to computationally-intensive scientific, mathematical, and academic problems and it is used in commercial enterprises for such diverse applications as drug discovery, economic forecasting, seismic analysis, and back-office data processing in support of e-commerce and web services. What distinguishes grid computing from typical cluster computing systems is that grids tend to be more loosely coupled, heterogeneous, and geographically dispersed. Also, while a computing grid may be dedicated to a specialized application, it is often constructed with the aid of general purpose grid software libraries and middleware.The picture could be fulfilled with portals and services.

In Table 1 below the basic Grid Projects, including some national and older ones, are summarised.

Table 1. Basic GRID projects

European Grid Initiative	National Grid projects	Previous projects
	Dutch Grid	EGEE
EMI - European Middleware Initiative		
gLite - European middleware distribution	GridPP	European Data Grid
Open Science Grid - the U.S scientific Grid infrastructure	INFN Grid	Datatag
Virtual Data Toolkit - provides middleware distribution	LCG France	Grid2003
	NorduGrid	GriPhyN
	WestGrid EGI	iVDGL
	GEANT	PPDG

1.1 Pan-European Grid - Enabling Grids for E-sciencE

The Enabling Grids for E-sciencE project (http://www.eu-egee.org) is no longer active. The project officially ended on April 30, 2010. The distributed computing infrastructure is now supported by the European Grid Infrastructure. This long-term organisation coordinates National Grid Initiatives, which form the Pan-European Grid (see Tabl.2).

Table 2. Pan-European GRID members

	Bulgarian Grid portal	Russian Grid
EGEE South East Europe		
Spain EGEE-III	Cyprus Grid	Portuguese Grid
Romanian Grid	Hungarian Grid	Slovak Grid

1.2 AstroGRID

AstroGrid (http://www2.astrogrid.org) is the doorway to the Virtual Observatory (VO). It provides a suite of desktop applications to enable astronomers to explore and bookmark resources from around the world, find data, store and share files in VOSpace, query databases, plot and manipulate tables, cross-match catalogues, and build and run scripts to automate sequences of tasks. Tools from other Euro-VO projects inter-operate with AstroGrid software, so one can also view and analyse images and spectra located in the VO.

AstroGrid, a UK-government funded, open-source project, helps create universal access to observational astronomy data scattered around the globe. The AstroGrid consortium, which consists of 11 UK university groups, represents astronomy and computing groups with background in handling and publishing such data. The consortium worked with international partners to agree upon standards for published observational astronomy data, so that all astronomers could interact with all data sets.

The AstroGrid workbench is the main user interface for astronomers accessing the virtual observatory. The global set of standards agreed upon by the consortium and its partners allows any astronomer to query the virtual observatory to ask for information on a certain area of the sky.

Through AstroGrid, UK astronomers can also access workflows and applications for data analysis. AstroGrid has also created the "voSpace" program that allows astronomers to share their workflows. For a lot additional possibilities see the paper of A. Lawrence (2002).

AstroGRID-D enables grid science in the German Astronomical community. Figure 1 shows all the applications proposed from German AtroGRID-D.

Fig. 1. Screenshot from the AstroGrid-D webpage

2 Middleware and Frameworks

In this section The structure of the instrumentation between software and hardware will be shortly reviewed.

2.1 FALKON

FALKON is a Fast and Light-weight tasK executiON framework for Clusters, Grids, and Supercomputers.

The main goal of the FALKON is to enable the rapid execution of many tasks on compute clusters. Some tests and benchmarks show that Falkon throughput (487 tasks/sec) and scalability (2,000,000 tasks processed in just 112 minutes) are better than other systems used in production Grids. The usage of FALKON can reduce 90% in end-to-end run time in large-scale astronomy calculations, relative to versions that execute tasks via separate scheduler submissions (Raicu et al., 2007).

2.2 GRACIE

GRACIE - Grid Resource Virtualization and Customization Infrastructure. Gracie is a lightweight execution framework for efficiently executing massive independent tasks in parallel on distributed computational resources. Three optimization strategies have been devised to improve the performance of Grid system.

- Pack up to thousands of tasks into one request.
- Share the effort in resource discovery and allocation among requests by separating resource allocations from request submissions.
- Pack variable numbers of tasks into different requests, where the task number is a function of the destination resource's computability.

Gracie is a computational grid software platform developed by Peking University (Li et al., 2008).

2.3 NIMROD

The tool set, called Nimrod (http://messagelab.monash.edu.au/Nimrod), automates the process of finding good solutions to demanding computational experiments. Importantly, Nimrod is more than a job distribution system; it is a high level environment for conducting search across complex spaces. The number of jobs, and thus the parallelism, can be varied at run time, and the Nimrod scheduler places tasks on the available resources at run time.

A good description of NIMROD tools can be found in Abramson et al. (2011).

2.4 IMAGER

IMAGER (Roberts & Crutcher,1997) is an interface to parallel implementation of imaging and deconvolution tasks of the Software Development Environment (SDE) of the National Radio Astronomical Observatory (NRAO). The interface is based on the MIRIAD interface of the BIMA (Berkeley-Illinois-Maryland Association) array and it allows for interactive and batch operations – with instruments such as the VLA and the BIMA telescopes one could have spectral line data sets in excess of a gigabyte. Astronomers need access to fast processing to allow the analysis of such large data sets and to use different methods of non-linear deconvolution. Radio synthesis data reduction has been one of the most computer intensive operations in observational astronomy. In the common case of radio spectral line observations, large numbers of frequency channels lead to large amounts of data. The analysis of spectral line data, in which each channel is independent from every other channel, is an embarrassingly parallel problem.

2.5 Parallel-Processing Astronomical Image Analysis Tools for HST and SIRTF

NASA applied information system research, developed and implemented several parallel-processing astronomical image-analysis tools for stellar imaging data from the Hubble Space Telescope and the Space Infrared Telescope Facility – see, e.g. Mighell (2005). This project combines the enabling image-processing technology of the Principal Investigator's new digital PSF-fitting MATPHOT algorithm for accurate and precise CCD stellar photometry with enabling technology of Beowulf clusters which offer excellent cost/performance ratios for computational power. Data mining tools for quick-look stellar photometry and other scientific visualization tasks are also written and used in order to investigate how such tools could be used at the data servers of NASA archival imaging data like the Space Telescope Science Institute.

3 Specialized software

Here we present some selected specialized software for astronomical grid computing.

3.1 AMEEPAR

AMEEPAR is parallel processing for hyperspectral imaging. The wealth of spatial and spectral information provided by hyperspectral sensors (with hundreds or even thousands of spectral channels) has quickly introduced new processing challenges. In particular, many hyperspectral imaging applications require a response in (near) real time in areas such as environmental modeling and assessment, target detection for military and homeland defense/security purposes, and risk prevention and response.

At the time being only a few parallel processing algorithms existed in the open literature – Plaza (2006).

To address the need for integrated software/hardware solutions in hyperspectral imaging, a highly innovative processing algorithms on several types of parallel platforms, including commodity (Beowulf-type) clusters of computers, large-scale distributed systems made up of heterogeneous computing resources, and specialized hardware architectures are developed.

Several parallel algorithms to analyze the AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) data were implemented. Among them is the automated morphological extraction (AMEEPAR). This is one of the few available parallel algorithms that integrate spatial and spectral information

(Wozniak, 2009). Using 256 processors, AMEEPAR provided a 90% accurate debris/dust map of the full AVIRIS data in 10s, while the P-ATGP algorithm was able to detect the spatial location of thermal hot spots in the WTC area in only 3s.

3.2 GADGET-2

GADGET-2 is a code for cosmological simulations of structure formation. Gadget is a freely available code for cosmological N-body/SPH simulations on massive parallel computers with distributed memory. GADGET uses an explicit communication model that is implemented with the standardized MPI communication interface. The code can be run on almost all supercomputer systems presently in use, including clusters of workstations or individual PCs. All details – in http://www.mpa-garching.mpg.de/gadget/.

GADGET computes gravitational forces with a hierarchical tree algorithm (optionally in combination with a particle-mesh scheme for long-range gravitational forces) and represents fluids by means of smoothed particle hydrodynamics (SPH). The code can be used for studies of isolated systems, or for simulations that include the cosmological expansion of space, both with or without periodic boundary conditions. In all these types of simulations, GADGET follows the evolution of a self-gravitating collisionless N-body system, and allows gas dynamics to be optionally included. Both the force computation and the time stepping of GADGET are fully adaptive, with a dynamic range which is, in principle, unlimited.

GADGET can, therefore, be used to address a wide array of astrophysically interesting problems, ranging from colliding and merging galaxies, to the formation of large-scale structure in the Universe. With the inclusion of additional physical processes such as radiative cooling and heating, GADGET can also be used to study the dynamics of the gaseous intergalactic medium, or to address star formation and its regulation by feedback processes.

GADGET provides also small examples:

- A pair of colliding disk galaxies (collisionless).
- A spherical collapse of a self-gravitating sphere of gas.
- Cosmological formation of a cluster of galaxies (collisionless, vacuum boundaries).
- Cosmological structure formation in a periodic box with adiabatic gas physics.

A review of GADGET can be found in Springel (2005).

3.3 CRBLASTER

CRBLASTER (Mighell, 2008) is a fast parallel-processing image-analysis program for cosmic ray rejection. Many astronomical image-analysis programs are based on algorithms that can be described as being embarrassingly parallel, where the analysis of one subimage generally does not affect the analysis of another subimage. Yet few parallel-processing astrophysical image analysis programs exist that can easily take full advantage of todays fast multi-core servers costing a few thousands of dollars. A major reason for the shortage of state-of-the-art parallel-processing astrophysical image-analysis codes is that the writing of parallel codes has been perceived to be difficult.

The code can be used as a software framework for easy development of parallel-processing image analysis programs using embarrassing parallel algorithms: the biggest required modification is the replacement of the core image processing function with an alternative image-analysis function based on a single-processor algorithm.

3.4 N_body-sh1p

N_body-sh1p is a parallel direct N_body code (Gualandris et al., 2007). This is an Educational N-body integrator with a shared but variable time step (the same for all particles but changing in time), using the Hermite integration scheme (Hut & Makino, 2003). The source code has been adapted for a parallel ring algorithm using the MPI library
(www-unix.msc.anl.gov/mpi/mpich).

Typical command line (generates : n24body.out)

$nbody\_sh1p < n24body.in > n24body.out$

Small timing test (perfomed by A. Gualandris) for 128, 256 and 512 particles with up to 32 processors on the Blue (Boewulf) linux cluster is presented at www.sara.nl.

3.5 Parallel processing algorithms

Papers and books present the parallel processing including algorithms, architectures etc. Among the basic books is Parhami's (1999). Cosmological problems, solved with parallel processing are presented in Bode & Bertschinger (1995) and in Ferrell & Bertschinger (1995).

3.6 Robotic Telescopes

Global networks of robotic telescopes provide important advantages over single telescopes. Independent of daytime and weather, they can perform more efficiently multiwavelength observations and continuous long-term monitoring, as well as react rapidly to transient events such as GRBs and supernovas. With the number of currently existing robotic telescopes a very powerful network could already be built - this is the idea of OpenTel - an Open Network for Robotic Telescopes. OpenTel provides the means for interconnecting single robotic telescopes to a global network for sharing observation time, observation programs and data. OpenTel is an open network. Open means open standards, open source and open for telescopes to join - http://www.gac-grid.org/project-products/Software/RoboticTelescopes.

Grid technology provides an ideal framework. It provides solutions for the management of Virtual Organizations, grid resources, computational jobs and observation, data and metadata. The architecture is built on two technologies: the grid middleware of the Globus Toolkit and the Remote Telescope Markup Language (RTML) for the exchange of observation requests.

3.7 The Networked Telescope

Pipeline processing systems for modern telescopes are widely considered critical for addressing the problem of ever increasing data rates; however, routine use of fully automated processing systems may discourage the typical user from exploring processing parameter space or trying new techniques. This issue might be particularly important with regard to radio interferometer data in which the post-calibration processing required to create an image for scientific analysis is not well defined. A good review of Grid architecture for Pipeline processing is given in Plante et al. (2001). The pipeline by default is automated and uses NCSA supercomputers to carry out the processing. This same system can also be used by the astronomer to create new processing projects using data from the archive.

Here are some ways that allow users to interact with the pipeline:

(a) prior to observations: the astronomer can override default processing parameters to better suit the scientific goals of the project;

(b) during observations: the astronomer can monitor the telescope and data via the web;

(c) after observations: the astronomer can browse the archive's holdings using customizable displays;

(d) prior to processing: the astronomer can create his/her own scripts for reprocessing archival data;

(e) during processing: optional viewers can be opened up to monitor, and possibly steer, the deconvolving process.

The processing is carried out using AIPS++, which employs the Glish scripting language to glue processing objects together. Its event-driven programming model (combined with the toolkit nature of AIPS++) makes it ideal for building automated processing in a distributed environment. An important role for NCSA, as a member of the AIPS++ development consortium, is to enable support for parallel processing on a range of mildly to massively parallel machines, with a particular emphasis on Linux clusters. The Intel Itanium-based supercluster that will be brought online at NCSA this year will handle the bulk of the imaging and deconvolution chores for the pipeline, while smaller machines will handle the serial processing.

4 Modelling and Simulations

In this part selected astronomical simulations, modelling and experiments are presented.

4.1 SkyMaker

SkyMaker is a program that simulates astronomical images. It accepts object lists in ASCII generated by the *Stuff* program (terapix.iap.fr/soft/stuff) to produce realistic astronomical fields. SkyMaker is part of the EFIGI development project (http://www.efigi.org). The authors are Emanuel Bertin (Bertin, 2009) and Pascal Fouque.

4.2 Cosmic simulation

A lot of things of cosmic simulation could be found in http://www.igstw.org. Cosmic structure formation theory has passed test after test, predicting how many galaxies will form, where they will form, and what type of galaxy they will be. But for almost 20 years, its predictions about the central mass of dwarf galaxies have been wrong.

Worldwide, there are many teams working on their own versions; each attacks the problem from a different angle.

E.g. Governato et al. (2010) imply that the dark matter particle that we think is the correct one is not the correct one, or maybe that gravity works differently than we think it does.

To create the simulation about a million computer hours were used, which means that it would have taken close to a hundred years to run the same simulation on the average desktop.

Klypin et al.(1999, 2007), Trujillo-Gomez et al.(2015) explored the large-scale effects of energy released by young stars. Stars are forming, and young stars release large amounts of energy into the gas that surrounds them. That energy finds its way to larger scales, affecting the motion of gas in the whole galaxy even the way it is being accreted in the galaxy.

Over time, scientific understanding of processes such as star formation has evolved, yielding new equations. The equations can in turn be used to refine the computational model. Bertshinger (1995) presented the COSMICS codes. Mayer et al. (2008) presented the formation of disk galaxies (Mayer et al., 2008) and on http://hydra.susx.ac.uk HYDRA consortium presented its N_body hydrodinamical simulations.

4.3 Flip-flopping of black hole accretion disks

The accretion disk of a black hole forms from gas attracted by the black hole's massive gravitational pull. For the last 20 years, astrophysicists have debated whether the whirlpool-like motion of the accretion disk will periodically reverse motion, a behavior called "flip-flop". According to a simulation powered by TeraGrid, the whirlpools of gas flip-flop as they are sucked into black holes – Blondin & Pope, 2009.

When flip-flopping first turned up in a 1988 numerical simulation, some scientists argued that it explains recurrent x-ray flares observed by the European X-Ray Observatory in 1985. But in subsequent years, although some simulations showed flip-flop, others did not, casting doubt on the existence of the phenomenon. The earlier work was criticized for a wide variety of reasons, but the chief among them was the lack of computer power and, hence, accuracy of the computation.

The most basic form of the equation used in the simulation was originally formulated by Fred Hoyle and Ray Lyttleton in 1939. The simulation showed that the accretion disk reversed direction repeatedly, confirming that at least in this model of black hole accretion disks, flip-flop does occur[1].

[1] For more details see http//www.igstw.org.

4.4 Millennium Simulation Project

The Millennium Simulation Project (http://www.mpa-garching.mpg.de/galform/virgo/millennium) is helping to clarify the physical processes underlying the buildup of real galaxies and black holes. It has traced the evolution of the matter distribution. The Millennium Run used more than 10 billion particles to trace the evolution of the matter distribution in a cubic region of the Universe over 2 billion light-years on a side.

It loaded the main supercomputer at the Max Planck Society's Supercomputing Centre in Garching, Germany for more than a month. By applying sophisticated modelling techniques to the 25 Tbytes of stored output, Virgo scientists have been able to recreate evolutionary histories both for the 20 million or so galaxies which populate this enormous volume and for the supermassive black holes which occasionally power quasars at their hearts. By comparing such simulated data to large observational surveys, one can clarify the physical processes underlying the buildup of real galaxies and black holes.

Among them are:

- A journey through the simulated universe.
- The dark matter distribution in the universe at the present time
- The galaxy distribution in the simulation on very large scales for a rich cluster of galaxies
- Slices of the dark matter distribution
- Halo and semi-analytic galaxy catalogues
- How did the universe evolve into the structures we know?

The very early universe consisted of homogeneous gas with tiny perturbations. As the gas cooled over time, it collapsed under gravity into clumps and then galaxies. Bode et al. (2001) simulation is one example for large scale simulations. The Millennium project researchers ran one of the largest simulation of a cosmological structure. In the simulation, the region of study collapses from about 2 billion light years across to form a region of galaxy clusters only 25,000 light years across.

The distribution of galaxy clusters in the universe can actually help us to learn things about dark energy, how much matter there is in the universe, and how fast the universe is expanding.

5 Astrophysical Applications and Projects

In this last part we shall make a short review of the basic astrophysical projects and applications.

5.1 A virtual universe - GIMIC and Millennium simulations

With the aid of the grid, researchers are conducting the largest-ever calculation to follow the formation of the dark haloes that seed galaxies.

Virgo is an international consortium of cosmologists that performs large numerical simulations of the formation of galaxies. Its Millennium Simulation is the largest ever calculation to follow the formation of the dark haloes that seed galaxies.

To understand the properties of the galaxies themselves, it is necessary to simulate how gas cools and forms stars in such haloes.

A key project of the Virgo Consortium – the Galaxies Intergalactic Medium Interaction Calculation (GIMIC)(Frenk, 2008), simulates the formation of galaxies in five key regions of the universe, allowing Virgo members to obtain unprecedented insight into how galaxies form. GIMIC simulates the formation of galaxies in several regions selected from the Millennium Simulation, but now including hydrodynamics. This allows Virgo members to obtain unprecedented insight into how galaxies form on truly cosmological scales. GIMIC has revealed that astrophysical processes separate the ordinary or "baryonic" matter from dark matter even on large scales. For details see http://www.deisa.eu – project GIMIC.

5.2 Scientific Applications of AstroGRID-D

Scientific Applications of AstroGRID-D can be found at http://www.gac-grid.org.

Dynamo scripts

The ***Dynamo*** scripts are designed to operate on a large number of compute nodes and they are an easy way to run many independent jobs on them. The package was originally an application for a magnetohydrodynamic simulation, but it has been developed further so it can be generally used.

Nbody6++

Nbody6++ is a member of a family of high accuracy direct N-body integrators used for simulations of dense star clusters, galactic nuclei, and problems of star formation. It is a special version of Nbody-6 optimised for massively parallel computers. Some of the most important applications are simulations of rich open and globular clusters with a large number of binaries and galactic nuclei with single and binary black holes.

Clusterfinder

Clusterfinder is a case used within the AstroGrid-D project that tests the deployment and performance of a typical data-intense astrophysical application. The algorithm for any point in the sky depends only on the data from nearby points, so the data access and calculation can easily be parallelized, making Clusterfinder well-suited for production on the grid. The scientific purpose of Clusterfinder is to reliably identify clusters of galaxies by correlating the signature in X-Ray images with that in catalogs of optical observations.

Astronomy in recent years turned from the study of individual or unusual objects to the statistics of large numbers of objects, observed at a variety of wavelengths across the electromagnetic spectrum, so that the techniques developed for Clusterfinder are applicable to many cutting edge astronomic studies, e.g. cosmology and galaxy clusters.

Cactus

The Cactus Computational Toolkit (http://cactuscode.org) is an open source problem solving environment designed for scientists and engineers. Cactus is used to numerically simulate extremely massive bodies, such as neutron stars and black holes. An accurate model of such systems requires a solution of the full set of Einstein's equations for general relativity.

ProC

ProC is an acronym for the Planck Process Coordinator Workflow Engine. The vast amount of data produced by satellite missions is a challenge for any data reduction software in terms of complex job submission and data management. The demands on computational power and memory space qualify the satellite data reduction to be prototype of grid applications. Therefore, the Planck Process Coordinator Workflow Engine ProC for the Planck Survey or satellite has been chosen as a Grid Use Case within the AstroGrid-D project. The ProC is interfaced to the Grid Application Toolkit (GAT), which allows the execution of jobs on the submission host, on clusters via the PBS and SGE GAT adapters, and on the Grid, using the Globus Toolkit 2 and 4 GAT adapters for process-to-process communication.

Below – Fig.2 is the block-scheme of the ProC interfaces to the Grid Application Toolkit (GAT) which via its set of adapters offers job execution on the local host, on worker nodes of a local cluster, and on remote Grid hosts. Details in http://www.gac-grid.org/project-products/Software/ProC/proc.pdf .

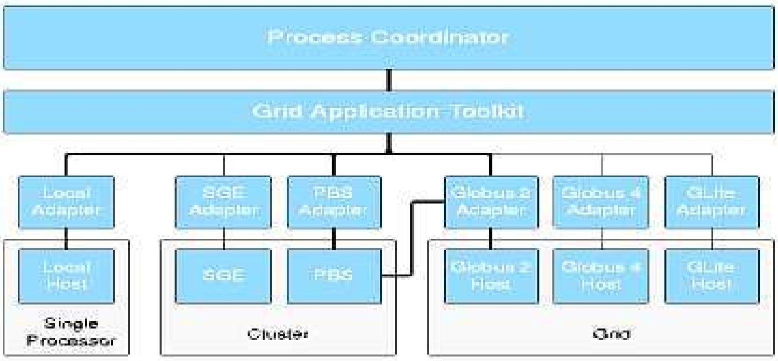

Fig. 2. Block-scheme of the ProC interfaces. Credit: http://www.gac-grid.org/project-products/Applications/ProC.html

5.3 Black holes and their jets

Jets of particles streaming from black holes in far-away galaxies are a highly discussed topic in astrophysics.

Hayashida and al. (2010) suggest that the magnetic field lines must somehow help the energy travel far from the black hole before it is released in the form of gamma rays. The data suggest that gamma rays are produced not one or two light days from the black hole (as was expected) but closer to one light year.

This simulation (Fig.3) depicts a black hole with a dipole as a magnetic field. This system is sufficiently orderly to generate gamma ray bursts that travel at relativistic speeds of over 99.9% the speed of light.

The black hole pulls in nearby matter (yellow) and sprays energy back out into the universe in a jet (blue and red) that is held together by the magnetic field (green lines).

The simulation was performed on the Texas Advanced Computing Center resources via TeraGrid, consuming approximately 400 000 service units. (Picture from the Internet).

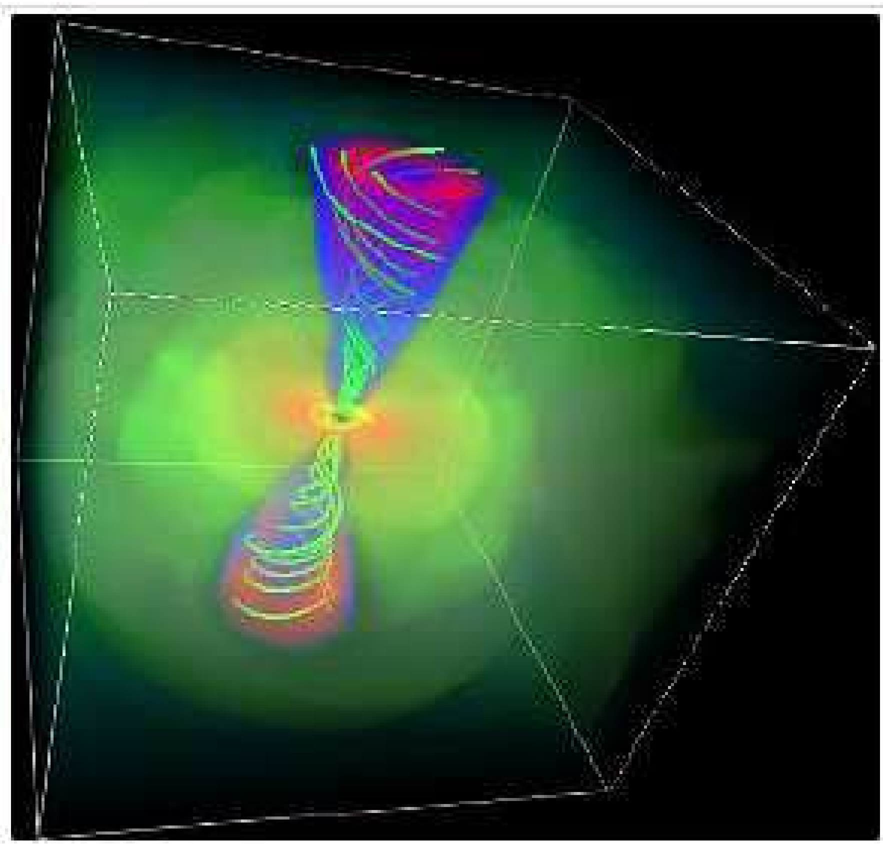

Fig. 3. Simulation of the black hole and its jet in a magnetic field. Credit:Jonathan McKinney (KIPAC) http://kipac.stanford.edu/kipac/projects/computational_astrophysics

This new understanding of the inner workings and construction of a blazar jet requires a new working model of the jet's structure, one in which the jet curves dramatically and the most energetic light originates far from the black hole.

5.4 Distant Galaxy Search Applying Astrogrid-RU

The first astronomical problem that has been experienced by Institute of Informatic Problems(IPI) RAN together with the Special Astrophysical Observatory of RAS (SAO RAS) applying AstroGrid and Aladin is a distant radio galaxy search in the sky strip investigated in the deep survey with the RATAN-600 (large Russian radio telescope). Details are in http://synthesis.ipi.ac.ru/synthesis/.

It used RC catalogue as a list of initial radio sources. One had to select optical sources with certain properties taken from DR 3 SDSS and cross-matched them with the RC catalogue. The result of the cross-match may contain candidates for distant galaxies that should be analyzed further applying their images and Aladin capabilities.

Figure 4 presents an example of a result image opened in Aladin.

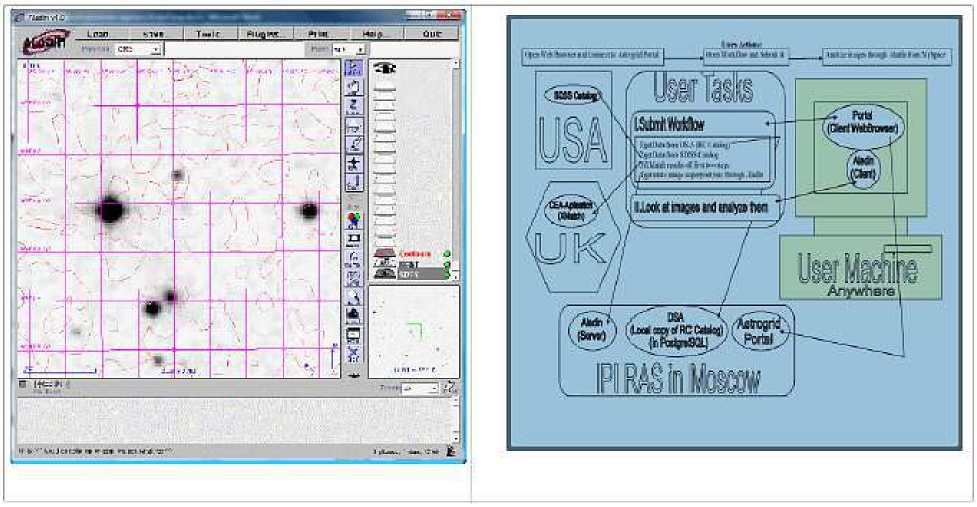

Fig. 4. Screenshot from the webpage of the Russian AstroGrid

5.5 GRID and the Virtual Observatory

SI-GRG: GRID Research Group at INAF SI in Trieste (SI-GRG) is doing research on Grid application and infrastructure development focused on Astronomical and Astrophysical problems – http://wwwas.oats.inaf.it/grid/.

Virtual Observatories (VObs) aim at federating astronomical databases in a way that they are accessible in a uniform way irrespective of peculiarities characterizing each of them (format of data, requests syntax). Virtual Observatories generally federate astronomical databases on a national basis; they in turn join other national VObs to form wider alliances on an international basis – see Passian (2004). IVOA (International Virtual Observatory Alliance) is the worldwide alliance of all VObs. The main goal of IVOA is to define a set of universally accepted standards in order to make possible a uniform vision of all federated Vobs (Rixon, 2009; Walton & Rixon, 2008). IVOA also supplies tools and software layers to practically implement this uniformity.

The concept of VObs therefore deals with data storage and retrieval. But astronomers have to process data once they have been retrieved and very often a considerable amount of computing power is requested to process such data. Because VObs offers astronomical data but not computing power a synergy between the VObs and the Grid appears as a natural choice (Shade, 2001).

The interconnection of GRID and the VO is presented in Skoda (2009) and Taffoni et al. (2009).

Among the other DRACO Project (Datagrid for Italian Research in Astrophysics and Coordination with the Virtual Observatory) is a concept aiming at providing the scientific community with a distributed multifunctional environment allowing the use of specialized (observational, computing, storage) Grid nodes. DRACO has been generated from a section of a project called "Enabling platforms for high-performance computational Grids oriented towards scalable virtual organizations" which has been approved and funded by the Italian Fund for Basic Research (FIRB).

5.6 HORIZON project

The HORIZON Project (Wozniak, 2009) is built on several research teams in different institutes, namely the CEA/SAp in Saclay, the Observatories of Paris (LUTh and LERMA laboratories), Lyon and Marseilles. The scientific objective is specifically oriented towards studying galaxy formation in a cosmological framework. Its transverse and federative nature allow to develop high-level expertise in parallel and distributed (GRID) computing, in database management and virtual observations, in applied mathematics and computer science, and build in the same time a strong theoretical knowledge in astrophysics.

The consortium also studies the influence on the predictions of the resolution and the numerical codes. The main objectives are:

- the numerical study of galaxy formation in a cosmological framework using Grand Challenge applications;
- the development of advanced techniques in parallel computing and in applied mathematics to model galaxy formation and predict their observational signatures, as a function of physical parameters
- the gathering of renowned experts in computational astrophysics to share their software and expertise, and to optimize their access to national and international supercomputing facilities

– the delivery of a friendly access to state-of-the-art numerical simulations to the scientific community of both observers and theoreticians.

Conclusion

The described GRID projects and applications are a small part of the existing ones, but in our opinion they are fully applicable in Bulgarian GRID infrastructure (Dechev et al., 2012, Petrov& Dechev, 2012).

References

Abramson, D., Bethwaite, B., Enticott, C., Garic, S. and Peachey, T., 2011, вЂњParameter Exploration in Science and Engineering using Many-Task ComputingвЂќ, *Special issue of IEEE Transactions on Parallel and Distributed Systems on Many-Task Computing*, June 2011, **22** Issue: 6, 960 вЂ“ 973.
Bertin E., 2009, *Mem.S.A.It., **80**, p.422*
Bertschinger E., 1995, *Astro-ph/9506075. p.20*
Blondin J. & Pope T., 2009, *ApJ, **700**, p.95*
Bode P. & Bertschinger E., 1995, *Astro-ph9504040. 18 p.*
Bode P. et al., 2001, *ApJ, **551**, p.15*
Dechev, M.; Petrov, G.; Atanassov, E., 2012, *PASRB*, **11**, pp. 83-88
Ferrell R. C. & Bertschinger E., 1995, *Astro-ph/9503042, 8 p.*
Governato F. et al., 2010,*Nature, **463**, p.203*
Grid-enabled Astrophysics, 2006, Eds.: Benacchio L. and Fabio Pasian F.
Gualandris A., Portegies Zwart S. & Tirado-Ramos A., 2007, *PARCO, **33**, p.159*
Hayashida M. et al., 2010, *Nature, **463**, p.919*
Kent S. M. et al., 2009, *BAAS, **41**, p.707*
Klypin A. et al., 1999, *ApJ, **516**, p.530*
Klypin A. et al., 2007, *Wurzburg, september workshop*
Lawrence A., 2002, *Proc. SPIE, **4846**, p.6*
Li H., Yu H. & Li X., 2008, *MTAGS08, p.24* and http://dsl.cs.uchicago.edu/MTAGS08/
Mayer L., Governato F. & Kaufmann T., 2008, *Advanced Science Letters, **1**, p.7*
Mighell K. J., 2008, *Proc SPIE, **7019**, p.70191F.1*
Mighell K., 2005, *MNRAS, **316**, p.861*
Parhami B., 1999, *In: Introduction to Parallel Processing: Algorithms and Architectures. Plenuum Press New York, 532 pages*
Passian F., 2004, *Mem.S.A.It. Suppl., **4**, p.86*
Plante R. et al., 2001, *ASP Conf. Ser., **238**. Astronomical Data Analysis Software and Systems X. Eds.: F. R. Harnden, Jr., F. A. Primini, & H. E. Payne (San Francisco: ASP), p.279*
Petrov, G.; Dechev, M.:2012, *PASRB*, **11**, pp. 181-186
Plaza A., 2006, *J. of Parallel and Distrib. Comp., **66**, p.345*
Plaza A., 2006, *Lecture Notes in Computer Science, **3391**, p.888*
Raicu, I., Yong Zhao, Dumitrescu Catalin, Foster, I., Wilde, M.,2007,"Falkon: a Fast and Light-weight tasK executiON framework", in SC '07. Proceedings of the 2007 ACM/IEEE Conference on Supercomputing.
Rixon G., 2009, *Mem.S.A.It., **80**, p.584*
Roberts D. & Crutcher D., 1997, *In: Astronomical Data Analysis Software and Systems. VI ASP Conference Series, **125**, p.120. Eds.: Hunt G. and Payne H. E.*
Schade D., 2001, *In: First Astro-Grid workshop, January 29-30, 2001*
Skoda P., 2009, *Mem.S.A.It., **80**, p.484* Springel V., 2005, MNRAS, 364, 1105
Stanoevska-Slabova K. & Wozniak T., 2010, *In: Cloud basics – an introduction to cloud computing. Grid and Cloud computing. Eds.: Stanoevska-Slabova K., Wozniak T. Ristol S. Shpringer Publ.house, p.47*
Taffoni G., Vuerli C. & Passian F., 2009, *Mem.S.A.It. Suppl., **13**, p.147*
Trujillo-Gomez, Sebastian; Klypin, Anatoly; ColГn, Pedro; Ceverino, Daniel; Arraki, Kenza S.; Primack, Joel, 2015, *MNRAS* **446**, Issue 2, p.1140-1162
Walton N. & Rixon G., 2008, *In: Garching workshop april'2008*
Wozniak H., 2009, *Mem.S.A.It.*, **80**, p.357