

On Advanced Geocomputations and Related Web Services

J. A. Rod Blais

Department of Geomatics Engineering
Pacific Institute for the Mathematical Sciences
University of Calgary, Calgary, AB, T2N 1N4
blais@ucalgary.ca www.ucalgary.ca/~blais

With scientific data generally increasing faster than computational power, it is imperative to explore new computational avenues in the analysis of legacy and current databases of observations and measurements. The situation is especially critical in areas of geoscience where terabyte (TB) datafiles are becoming very common and new data and metadata are coming in every day. More powerful computers and computer clusters are only part of the solution as more intelligent approaches are obviously needed.

Rapidly changing communication and High Performance Computing (HPC) technologies have brought about new possibilities in the handling of geospatial and related information. Sophisticated web services are opening up new avenues in distributed data processing, interactive visualizations, communication with remote sensors, and critical monitoring systems for better informed decision support. Exciting developments such as virtual globes, observatories and HUBs are discussed with examples to illustrate some of the far reaching implications not only for users in Geomatics and Geoscience but for society in general.

1. Introduction

With the spectacular advances in information technology, the general infrastructure necessary to capitalize on the potentials has been termed Cyberinfrastructure by the U.S. National Science Foundation [NSF, 2003, 2005 and 2007]. In this rapidly evolving environment, cyberinfrastructure integrates hardware for computing, databases and networks, digitally-enabled sensors, observatory and laboratory facilities, and interoperable suites of software and middleware tools and services [Pillai, 2007]. Obviously, the new Cyberinfrastructure greatly changes the geocomputation playing field for Geomatics and Geoscience as well as society in general [e.g. Blais and Esche, 2008]. The ever increasing data volumes and the rapidly changing communication and computer technologies make the exploration of new computational strategies imperative to stay at the forefront in research, education and industrial development.

First in terms of computer hardware performance, there has effectively been an exponential growth over the past decades with the doubling of

- gate densities every 18 months;
- storage capacity every 12 months;
- network capacity every 9 months,

[Pillai, 2007]. This situation has enormous implications for software and middleware developers especially in distributed high performance computing (HPC). In terms of maximum speeds of

computations, given in terms of Floating-Point Operations per second (FLOP/s), the following breakthroughs are quite impressive

- 1.24 MFLOP/s, or 1.24×10^6 FLOP/s, on a CDC 7600 computer in 1971;
- 1.7 GFLOP/s, or 1.7×10^9 FLOP/s, on a Cray 2 computer in 1986;
- 1.068 TFLOP/s, or 1.068×10^{12} FLOP/s, on an Intel ASCI computer in 1996;
- 1.026 PFLOP/s, or 1.026×10^{15} FLOP/s, on an IBM Roadrunner computer in 2008

[Meuer, 2008]. Parallel and grid computing have largely replaced scalar computing as most desktop computers and notebooks have dual and quadruple core processors and web-based transactions becoming commonplace in numerous contexts. New experimental developments in terms of virtual working environments offer some new ways in web interactive visualization and collaboration, especially when coupled with distributed computations.

Advanced geocomputations can greatly benefit from the rapidly evolving cyberinfrastructure that supports distributed high performance computing, virtual environments with advanced visualization capabilities for collaboration in research, development and education. Web services which essentially involve intercomputer communications greatly facilitate the distributed computations and collaborations among the web users. Web services have recently become popular for various kinds of commercial, education and game applications. The technical possibilities for scientific and engineering applications are practically endless and warrant further considerations and developments.

With the new Cyberinfrastructure providing high-bandwidth communications, advanced languages and protocols can be used to network computer clusters, plug-and-play sensors, simulations and processes into seamless contributions to management and decision support systems. Furthermore, recent advances in electronic circuit miniaturization and Micro-Electro-Mechanical Systems (MEMS) have led to the creation of small sensor nodes integrating several sensors, a Central Processing Unit (CPU), a memory and wireless transceiver [Newman, 2006]. Such networking of smart sensors has already been experimentally implemented in the Environment Canada RésEau initiative (<http://map.ns.ec.gc.ca/reseau/>), the Prion project (<http://www.prioninstitute.ca>), the Road Weather Information Systems of the Intelligent Transportation System (<http://www.its-sti.gc.ca>), and numerous others.

Interdisciplinary teams of researchers and systems engineers are necessary to take advantage of the new cyberinfrastructure to advance knowledge and experimentation for the benefit of society in general. However, with the general availability of virtual globes and observatories, the average scientist can use web communications for distributed computing, interactive visualization and collaboration with colleagues around the world. The following sections on web developments are intended to convince the reader that much can be gained from learning about the new cyberinfrastructure and experimenting with the available webtools.

2. Geocomputations and High Performance Computing

Over the past decades, geocomputations have greatly evolved from the simple data processing of laboratory or field observations and measurements to sophisticated distributed computations using computer clusters and computational grids. Geolibraries of data and metadata have

become accessible over the web and communication networks allow the transfer of gigabytes (GBs) of data in relatively short times over broadband networks.

Until quite recently in Geomatics, least-squares block adjustments of some 2000 stereomodels (with 7 unknowns per model) and survey networks of nearly 8000 stations (with two unknowns per station) were considered advanced geocomputations in the Canadian context [Beattie et al, 1978; Blais, 1979]. The North American Datum (NAD) horizontal network readjustment of some 450,000 stations in Canada, the U.S., Central America and Greenland was among the largest geodetic network readjustment ever carried out [Schwarz, 1989]. These computations were carried out on mainframe computers of the CDC Cyber and IBM 370 series. Using the same approach of Helmert blocking for geodetic networks, today, such network adjustments could easily be done on standard desktop computers, and even on notebooks. The differences in terms of computational capabilities and costs are clearly of several orders of magnitudes.

Advanced geocomputations are now becoming more complex and involve greater computational efforts. For instance, initial and boundary value problems for climate and other environmental problems, often with multiresolution data and constraints, are generally very complex problems. Furthermore, data assimilation problems and simulations of complex systems in geoscience are often nonlinear and therefore require iterative procedures. Hence the general trend towards distributed computing on computer clusters and computational grids. Furthermore, such high performance computing generally requires computer visualizations to analyze the results and communicate their interpretations to the scientific community and the general public. For collaboration with colleagues at home and abroad, the communication networks of the cyberinfrastructure are essential especially for interactive visualizations.

For global geoscience applications, discrete Spherical Harmonic Transforms (SHTs) of very high degrees and orders are required to achieve the necessary ground resolutions. Recent developments with degrees and orders over 3600 imply a ground resolution of approximately 5 km which is barely sufficient for global environmental modeling [Blais et al, 2006]. The recently released Earth Geopotential Model 2008 (EGM08) is of maximum degree 2190 and order 2160 which implies a ground resolution of approximately 8 km [Pavlis et al, 2008]. For regional applications of even continental extent, Fast Fourier Transforms (FFTs) are the norm with regular grids, and for global applications, discrete SHTs are the analog transforms on the sphere. Notice that such array computations are usually done with equiangular spherical grids and Cartesian planar grids although near equiareal spherical grids are needed in several applications (see examples of graticules in Figure 1) [Blais, 2007].

Environmental research often requires the modeling of spatio-temporal processes as stochastic processes for data assimilation and predictions. Under simplifying assumptions in terms of system identification, evolutionary mean-square spatial representations in terms of Karhunen-Loève and wavelet expansions can be formulated using linear autoregressive or similar interaction schemes [Blais, 2003]. With observations or measurements available only at specific epochs, prediction information can be obtained from the interaction model while observational data enable possible refinements of those predictions. Appropriate matrices of array data sequences need to be analyzed for significant modal patterns in space and time in order to investigate trends and other characteristics for prediction and related applications. Such research

involving all kinds of legacy databases of observations and measurements is especially important for climate change implications [Mutulu et al, 2004].

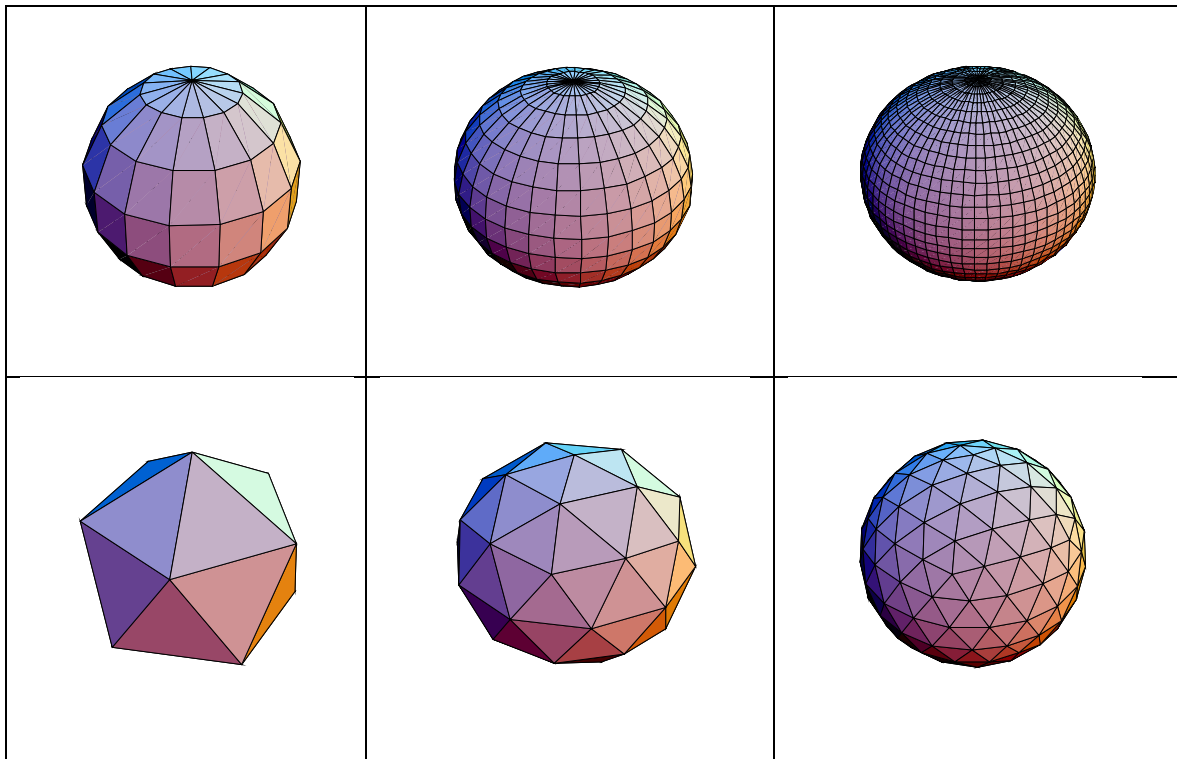


Figure 1. Examples of Spherical Equiangular and Icosahedron-Based Near-Equiareal Graticules

Environmental modeling and climate studies involve all kinds of computer simulations over centuries and millenia and even millions of years in some cases. With very limited observational information in space and time, Monte Carlo simulations are the scientific investigating approach of choice and the involved computations are often distributed with web-based user interfaces for scripting the data processing operations and analyzing the results using appropriate displays of objects and processes. The simulated fields are often vectorial and tensorial which imply complex computer visualizations, often with animations over time sequences.

Other advanced geocomputations such as seismic data processing have always been very computationally demanding. Pre-processing of one and two-dimensional seismic field data, followed by velocity analysis in 1-D, 2-D or even 3-D, and time imaging in 3-D and depth imaging in 3-D require enormous computational efforts that are best carried out on computer clusters and grids such as EGEODE [Thomas, 2007]. Reservoir geophysics using structural models and rock properties then lead to the required geophysical exploration information.

Reservoirs of heavy oil and tar sands (HOTS) consist of porous media and the fluids are heterogeneous on many scales. Hence the challenge of geotailoring, i.e. optimizing well placements and process operations for an existing reservoir. A recovery process often exhibits some geotolerance, i.e. the ability to operate efficiently in the presence of unpredictable geological factors. Computer simulations are highly complex and their nonlinearity require

hundreds and thousands of computer runs to achieve the desired optimal solutions [Gates and Bourque, 2007].

3. Communications and Web Services

Considering the exponential growth of computing capability with the ever expanding coverage and capacity of communication networks, it is becoming imperative to have the distributed data collections seen as an integrated system for research and development. The volumes and complexity of the available information require novel ways of processing, storing, accessing and presenting the information. Terabytes (TBs) and petabytes (PBs) of data also bring about multiple serious problems of communication and maintenance [e.g. Layton, 2008].

Scientists and researchers tend to spend much time identifying and acquiring datasets for their research purposes. Centralized data repositories often serve as a single entry point for acquiring and disseminating scientific data. One such central data repository is the World Data Center System [WDC, 1996], at the NOAA National Geophysical Data Center (NGDC), which supports international exchanges of geophysical observations in accordance with principles set forth by the International Council for Scientific Unions. These Centers collect data and publications for various geoscience disciplines such as Airglow, Astronomy, Atmospheric Trace Gases, Aurora, Biodiversity and Ecology, Climate, Cosmic Rays, Earth Tides, Geology, Geomagnetism, Glaciology, Human Interactions in the Environment, Ionosphere, Land Cover Data, Marine Environmental Sciences, Marine Geology and Geophysics, Meteorology, Nuclear Radiation, Oceanography, Paleoclimatology, Remotely Sensed Data, Remote Sensing of the Atmosphere, Renewable Resources and the Environment, Rockets and Satellites, Satellite Information, Space Science Satellites, Rotation of the Earth, Seismology, Soils, Solar Activity, Solar Radio Emissions, Solar Terrestrial Physics, Solid Earth Geophysics, Space Science and Sunspot Index (<http://www.ngdc.noaa.gov>) .

NGDC's interactive map services provide visual displays of one or more data layers, often with links to download the underlying data. Users can view layers containing different types of data, adding reference information such as continents, rivers, and political boundaries if desired. Zoom, pan, identify, search, find, and hyperlink capabilities are also available. The National Snow and Ice Data Center (NSIDC: <http://nsidc.org>) was established at the University of Colorado by NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) in 1982, to serve as a national information and referral center for polar research. The NOAA team at NSIDC manages about 60 NOAA datasets, and publishes several new datasets each year, with an emphasis on in situ data, data rescue, and datasets from operational communities such as the U.S. Navy. Educational pages for schools and the general public also have a high priority.

The NSIDC also offers some datasets in the form of images, and using Google Earth™, such data-based images and time series can be interactively overlaid on a virtual globe. For a better understanding of the cryosphere, snow, ice, ice shelves, glaciers, permafrost and sea ice can be displayed on a virtual Earth globe. For ocean observations, AVISO (Archiving, Validation and Interpretation of Satellite Observations: <http://www.aviso.oceanobs.com>)

distributes satellite altimetry data from Topex/Poseidon, Jason-1, ERS-1 and ERS-2, and EnviSat, and Doris precise orbit determination and positioning products. Map views of Sea Level Anomalies, Absolute Dynamic Topography, Significant Wave Height and Wind Speed are available.

Data repositories also exist at the national level such as the Geoscience Data Repository of NRCan (<http://gdr.nrcan.gc.ca>) which includes Maps (MIRAGE – Geoscience map images, digital geoscience maps), Geophysical Surveys (Aeromagnetic and electromagnetic data, gravity data, radioactivity data, seismic and magnetotelluric data), Geochemistry (NGR geochemical data, geochemical survey metadata), Other Geoscience Data (mineral deposits, energy – fossil data discovery tool, canadian geochronology knowledge base, physical rock properties), Marine Geoscience (expedition database, physical archive database, Canadian marine multibeam bathymetric data), and Other Resources (Lexicon of Canadian geological names, GeoPub, GeoGratis, GeoBase, GeoConnections Discovery Portal, geomagnetic data, Canadian landscape photo collection, Canadian diatom database, borehole geophysical logs, BASIN, seismic data, climate change, natural hazards and urban geology). At the local level, libraries such as the McKimmie Library and the North American Institute of North America at the University of Calgary have collections of maps, DEM data, digitized air photos and satellite imagery. Other government and academic centres also have various geospatial data and imagery collections for research and educational purposes. Most of these collections are becoming accessible and searchable over the Internet. A number of these datasets are also becoming available in KML and similar formats for visualization with Google EarthTM and other geobrowsers.

Internet communications between users and data centres would be very limited without inter-computer communications or web services. These web services are pieces of software available over a network with formal descriptions of how they are called and what actions or outputs can be expected in a language that computers understand. Such services are meant to be used by computers and not humans unlike e.g. webpages. It is important to note that entities such as web servers, ftp servers and database servers do not generally qualify as they lack the standardized description of their inputs and outputs.

Different web service models are available depending on the intended applications. Among the most common ones are the following:

- i) Resource-Oriented Models with resources being essentially anything that can have an identifier, e.g. a service.
- ii) Service-Oriented Models with services being something capable of performing a piece of coherent functionality, e.g. actions performed by agents.
- iii) Message-Oriented Models with messages being the basic unit of data exchanged between agents.
- iv) Policy-Oriented Models with policies being sets of assertions expressing capabilities and constraints.

Different frameworks are readily available for designing and developing web services. Table 1 summarizes the packages and libraries available with a number of computer languages.

Table 1. Frameworks for Web Services

LANGUAGE	PACKAGE or LIBRARY	URL
Java	Apache Axis Xfire	http://ws.apache.org/axis http://xfire.codehaus.org
C#	.NET Mono	http://msdn2.microsoft.com/en-us/netframework/ http://www.mono-project.com
Perl	SOAP::Lite	http://www.soaplite.com
Python	SOAPPY/ZSI	http://pywebsvcs.sourceforge.net
C++/C	gSOAP	http://www.cs.fsu.edu/~engelen/soap.html
Ruby	Soap4r	http://dev.ctor.org/soap4r

With remote sensors networked together into a sensorweb, environmental research is fundamentally changing with near real-time monitoring of events under study. For instance, having collared wild animals provide location and other environmental data to researchers and park managers has wide-ranging implications for all concerned (see e.g. Herrero [2005]). Similarly, cattle and other farm animals are being monitored using GPS and other sensors for any anomalous behavior such as in the context of BSE (see e.g. <http://www.prioninstitute.ca>). The possibilities are practically endless with miniature sensors and MEMs for tracking and/or monitoring environmental conditions such as temperatures, water levels, snow conditions, etc..

With appropriate high bandwidth communications, advanced languages and protocols such as the Sensor Model Language (SensorML; <http://vast.uah.edu/joomla>), can be used to network plug-and-play sensors, as well as simulations and processes. Specifically, the Open Geospatial Consortium (OGC) SensorML specifications would allow interoperability across monitoring programs and data extraction tools, such as in the RésEau initiative of Environment Canada to demonstrate the sharing, discovery, access and use of water information over the Internet (see <http://map.ns.ec.ca/reseau/> for more details). RésEau provides the user with modern search tools, interactive mapping and downloadable applications, all accessible in one place. Another illustrative example is the Road Weather Information System (RWIS) of the Intelligent Transportation System (<http://www.its-sti.gc.ca>), which combines the road maintainers' knowledge, meteorological information, other relevant environmental data, traffic and safety data, and multi-jurisdictional station data from passive and active sensors [Pinet, 2003]. The RWIS is part of a larger network of RWIS sites across Canada so that more accurate forecasting of storms can be achieved in real time by following their progress across the network.

4. Virtual Globes and Observatories

Earth Virtual Globes allow the draping of geographical information over a three-dimensional Earth globe with interactive orientation, scaling, lighting and other operations. Virtual Globes have already been used in numerous geographical and geoscience applications, including some interfacing with existing HPC projects. A number of presentations at the Fall Meetings of the American Geophysical Union in San Francisco, CA, in 2007 can be found at <http://conferences.images.alaska.edu/agu/2007> . Three examples of such geobrowsers are going to be briefly described with their capabilities and advantages.

i) **GOOGLE EARTH™** (<http://earth.google.com>)

With software acquired from Keyhole Inc. (2004), Google developed a multilingual and multi OS (Windows, Mac and Linux) virtual Earth globe for web applications. Different versions are available from Google i.e. a free version, a superior version for \$20 USD/year, and developer's versions (with features like movie making and GIS data importing) for \$ 400-1030 USD/year. Satellite imagery from Landsat 7 and DEM coverage from NASA's Shuttle Radar Topography Mission (SRTM) provide global coverage with resolutions from 15 cm in urban areas to over 15 m in remote areas of the globe. Three-dimensional datasets in KML language can be read in for visualization purposes. Google Street View has also become available with Version 4.3 (May 2007). Note that terrain heights below sea level have been found in error such as locations in Death Valley, CA, which were displayed at some -86m instead of approximately -420m. Google has resolved many inaccuracies in the vector mapping since the original public release of the software without having to modify the software and of course, the geospatial data are been upgraded all the time.

ii) **Microsoft Virtual Earth™** (<http://www.microsoft.com/virtualearth/>)

With business objectives, Microsoft developed three-dimensional city views using immersing imagery obtained at street levels to supplement global satellite imagery. Improved location and direction functionality are aimed at traveling requirements to points of interest. Online searching and mapping are available in a web environment at <http://maps.live.com> and a Virtual Earth (VE) interactive Software Development Kit (SDK) can be found at <http://dev.live.com/virtualearth/sdk/> This interactive VE provides hands-on, task-based demos of various Application Programming Interfaces (APIs), including designer tools and technologies for developing one's own website.

iii) **Poly9 FreeEarth™** (<http://freeearth.poly9.com>)

Using Adobe Flash Player 9, Poly9 FreeEarth requires minimal downloading from the computer server, making this virtual globe approach most appropriate for numerous mobile applications. With cross-browser and multi-OS (Windows, Mac and Linux) capabilities, the software is very light with a main application requiring only ~300 KB. Using global satellite imagery with high resolutions for U.S. cities, Web Map Service (WMS) is supported for GIS applications and any data encoded in KML (GeorSS specifications) can be displayed easily. All software features are exhibited using simple JavaScript interfaces. Imagery for other planets is also included for visualization purposes.

Somewhat more generally, Virtual Observatories (VOs) provide web environments for

- browsing, searching and downloading data and tools,
- integrating, synthesizing and analyzing new and legacy datasets,
- correlating, modeling and assimilating datasets,
- prototyping and developing new tools and procedures,

along with interactive visualizations for web-based research, development and collaboration. Notice that distributed computing is considered an integral component of VOs as very large datasets and complex computational tasks are often encountered.

The best known example of VOs is the National VO (NVO) in Astronomy (<http://us-vo.org>), with an overview shown in Figure 2. Among the NVO applications are the following:

- i) Web-based applications through a web browser such as with Virtual Globes;
- ii) Downloaded applications run locally on desktop, notebook, local workstation, cluster or grid;
- iii) Computationally intensive applications initiated through a web browser, carried out on some cluster or grid, with status information available through a web browser, and with appropriate visualization of the results;
- iv) Toolkits for building other applications in terms of scripting languages such as IDL and Python, and programming languages such as Java and C++.

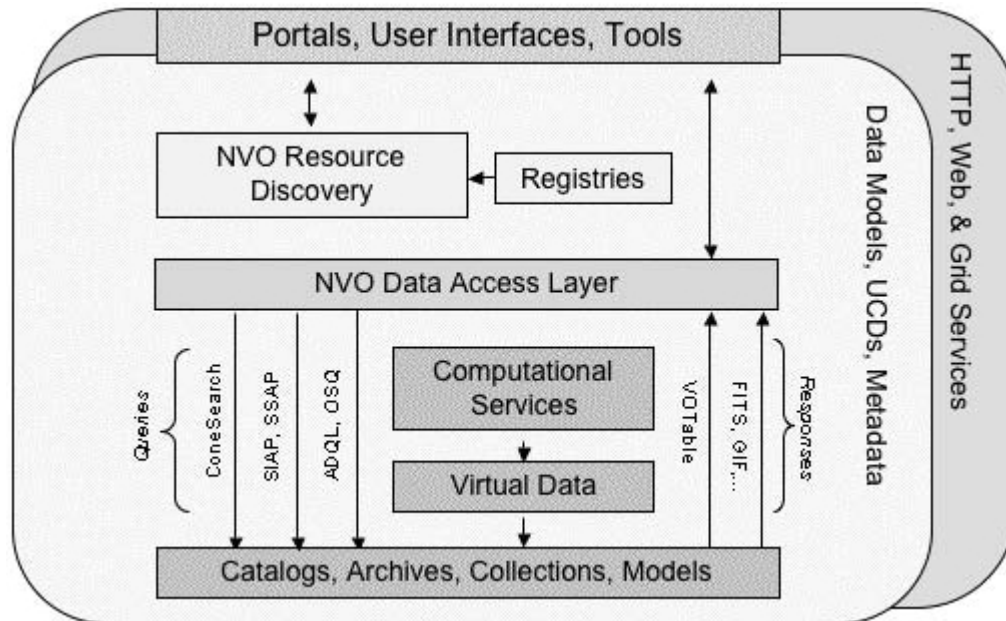


Figure 2. Overview of National Virtual Observatory. Source: [Graham et al, 2008].

The NVO has a number of stand-alone applications for (astronomical) image processing, viewing and visualization, and some web-based tools such as

- i) Datascope for accessing data holdings and related services,
- ii) Registry for descriptions of data archives and services.

The NVO started as a self-formed alliance of 17 data centres, observatories and computer science organizations going back to 2001 that has led to funding from the US NSF's Information Technology Research Program. The group is now known as the International Virtual Observatory Alliance (IVOA: <http://ivoa.net>) [Graham et al, 2008].

A slightly more general concept is that of a 'HUB' which is intended at general users including researchers, educators and students to access tools and share information. A HUB is a web site built with many familiar open source packages such as a Linux operating system running Apache web server with

- LDAP (Lightweight Directory Access Protocol) for user logins,
- PHP (personal Home Page) for web scripting,
- Joomla (Open Source Content Management System),

- MySQL database for storing content and usage statistics, (<http://www.hubzero.org>). The HUBzero software builds upon that infrastructure to create a web-based collaborative environment with features such as interactive simulation tools, online presentations, mechanisms for uploading new resources, tool development areas, ratings and citations, content tagging, user groups for private collaboration, user support areas, usage statistics along with news and events, and feedback mechanisms.

The signature service of a HUB is its ability to deliver interactive, graphical simulation tools through an ordinary web browser, unlike most other portal-based cyber-environments. The HUBzero infrastructure includes a tool execution and delivery mechanism based on Virtual Network Computing (VNC). Just about any tool with a graphical user interface can be installed on the hub and deployed within a few hours. For legacy tools and other codes without a graphical interface, an interface can be quickly created using HUBzero's associated Rappture toolkit which can be used to set up jobs and visualize results. The jobs themselves can be dispatched to the TeraGrid, the Open Science Grid, and other participating cluster resources. A demo using nanoHUB can be found at <http://www.hubzero.org/demo.html>.

To summarize, some of the key web-based features of VOs and HUBs are

- i) Interactive simulation tools and online presentations;
 - ii) Mechanisms for uploading new resources;
 - iii) Tool development area and usage statistics;
 - iv) User groups for private collaboration;
 - v) Ratings and citations and user support area;
 - vi) News and events, and feedback mechanisms;
- which obviously promote collaborative research and development in ways never possible in the past when electronic communications were very limited.

5. Web Collaboration

The NVO has become the enabler of integrative, comparative research based on astronomical surveys, archives and simulations. By providing a suite of standard protocols for discovering, accessing and correlating these diverse data resources, the NVO facilitates a new era of discovery. Similar possibilities are achievable in numerous other domains.

The environmental applications such as weather predictions and climate studies are often mentioned as the outstanding ones to benefit from web communications. Multiple layers of satellite imagery and field sensor data from large, and often remote, regions can be combined to provide more complete pictures of events of interest. This is especially useful in arctic science where accessibility is often problematic.

Furthermore, natural hazards such as earthquakes, tornadoes, volcanoes and tsunamis and others can now be studied as never possible before with extensive data assimilation schemes and computer simulations. These computations are very complex and normally carried out on computer clusters and computational grids. The results need proper visualization and animation tools to help the scientists interpret the results. The web communication capabilities enable

groups of scientists around the world to collaborate in designing experiments and analyzing the graphical renderings of the results. For instance, strain rate monitoring in earthquake research and ash tracking in volcano eruptions are becoming feasible with sensor webs and the analysis can involve all the interested scientists.

Using NOAA's Live Access Server (LAS) (<http://ferret.pmel.noaa.gov/Ferret/LAS/>) and the Distributed Ocean Data System (DODS), (<http://www.unidata.ucar.edu>), interactive data fusion and mapping tools are available for visualization and subsetting multidimensional scientific data for Web users in support of collaborative research [Hankin et al, 2001]. On-the-fly visualization makes it possible to explore the dataset entirely within a web browser environment. Any further analysis on data subsets can be carried out locally by extracting the desired data subset and downloading them over the Internet. The reformatting capabilities make it simple to ingest the data into the scientist's choice of desktop environments.

OceanDIVA (Ocean Data Inter-comparison and Visualization Application: http://lovejoy.nerc-essc.ac.uk:8080/MOD-OBS/OceanDIVA_info.html) was developed as a model-data comparison and validation tool for oceanography that can be used with standards-compliant datasets across the Internet. OceanDIVA is a Java web application developed by and hosted at the Reading e-Science Centre, UK, with the Environmental Systems Science Centre at the University of Reading, UK. In-situ ocean observation profiles in NetCDF format are analyzed and compared with gridded model output, providing discrepancy maps either in KMZ (i.e. zipped KML) format for viewing in Google Earth™ or series of digital images of probability density functions (PDFs) within selected regions. More information on OceanDIVA is available from the website.

Coastal ocean observations and hydrological resources are very important at the all levels in the environmental sciences ranging from regional watershed and freshwater conservation planning to global hydrological, climate, biogeochemical, and land surface modeling. HydroSHEDS (<http://hydrosheds.cr.usgs.gov>) which was developed by the Conservation Science Program of the World Wildlife Fund, is based primarily on elevation data obtained during NASA's SRTM. At the most basic level, HydroSHEDS information will allow scientists and managers to create digital river maps and delineate watershed boundaries. These maps can then be coupled with a variety of other datasets or applied in computer simulations for hydrological research. The HydroSHEDS database contains both raster and vector layers describing the topography, drainage networks, and watersheds of the Earth's surface [Lehrer et al, 2006].

At the more engineering levels, streetviews and traffic monitorings are very important in everyday life. Web tools using Google Earth™, Microsoft Virtual Earth™ and others have already contributed real-time navigation tools for military operations overseas, automobile and ambulance drivers in congested urban zones and in natural disaster regions of recent tornadoes, hurricanes and floodings. In some contexts, ship traffic in restricted channels and sensitive areas can be monitored using appropriate remote sensors with the results being accessible interactively on graphical monitors to support decision management systems.

A necessary element to convince the potential user community for space products is effective data visualization: facilitating the display, animation and layering of multiple satellite imaging and sounding sensors (providing complementary information) in a user-friendly and intuitive

fashion. These data are also making their way into the classroom owing to efficient and timely data archival systems and dissemination over the Internet, thus introducing environmental science to much wider spectrum of users. An interesting example in a near real-time setting from the 2007 hurricane season was developed in the Google Earth™ framework and presented at the AGU 2007 Fall Meeting by [Turk and Miller, 2007]. A menu of imagery based sequential satellite overpasses (GOES and other geostationary satellites, TRMM, CloudSat, Terra, Aqua, DMSP, NOAA, QuikScat) during the storm lifecycle were presented in a structured folder format. Web tools enable the user to view, animate, zoom, overlay and combine visible, infrared and passive microwave imagery and combine with other data such as surface reports, forecasts, surface winds, ground and spaceborne radars, at various stages of the hurricane lifecycle. A combination of such geo-navigable datasets provides a convenient framework for efficiently demonstrating meteorological, oceanographic and weather and climate concepts to students, planners, and the public at large.

6. Concluding Remarks

Advanced geocomputations often involve HPC for complex applications such as multiresolution data assimilation problems and simulation studies for climate and related environmental research. Such computations are often distributed over computer clusters and computational grids with web-based user interfaces. The new cyberinfrastructure enables the web users to collaborate using virtual environments with interactive visualization and related communication tools and facilities. All the interested community of scientists, managers and students can fully participate in the experimentation and interpretation of the results.

Virtual globes are especially useful for displaying imagery at various scales and perspective views. With the enormous amounts of satellite imagery available from Low Earth Orbiting Environmental satellites and legacy databases, such tools are really advantageous to exploit the imagery of various spatial and spectral resolutions. Such Virtual Globe approaches can be combined with web applications and interfaces for analyzing and sharing data and information with colleagues around the world.

Virtual observatories are new developments which include web-based applications and distributed computing as well as sensor networks to collect new observations and measurements. The National Virtual Observatory in astronomical research is perhaps the best known example of such web development which offers full access to all astronomical databases with a full range of data processing and computational facilities, including extensive HPC applications on the TeraGrid. The NVO initiative has already been considered for applications in geomagnetism and polar research associated with the International Polar Year of 2007-2009. Geoscience can greatly benefit from this new technology.

Investments in interdisciplinary teams of researchers and systems development are necessary to take advantage of the full power of cyberinfrastructure to create, disseminate, and preserve scientific data, information, and knowledge [Pillai, 2007]. The challenge for the average scientist and manager is to start experimenting with unfamiliar web technologies to access data and metadata bases along with simple processing and visualization tools. There is no substitute

for hands-on experimentation and the future looks most promising for those who will take the challenge!

Acknowledgements

The author would like to acknowledge the sponsorship of the Natural Science and Engineering Research Council in the form of a Research Grant on Computational Tools for the Geosciences. Comments and suggestions from colleagues are also gratefully acknowledged.

References

- Beattie, D.S., J.A.R. Blais and M.C. Pinch [1978]: Test Adjustments of the Canadian Primary Horizontal Network. Second International Symposium on Problems Related to the Redefinition of North American Geodetic Networks, pp. 41-54.
- Blais, J.A.R. [2007]: Optimal Spherical Triangulation for Global Multiresolution Analysis and Synthesis. Poster presentation at the Fall Meeting of the American Geophysical Union in San Francisco, CA.
- Blais, J.A.R. [2003]: Modeling Spatio-Temporal Environmental Processes for Data Assimilation and Prediction. Oral presentation at the Annual Meeting of the Canadian Geophysical Union in Banff, AB.
- Blais, J.A.R. [1979]: Least-Squares Block Adjustment of Stereoscopic Models and Error Analysis. Technical Report, no.30001, Division of Surveying Engineering, University of Calgary. 208 pages.
- Blais, J.A.R. and H. Esche [2008]: Geomatics and the New Cyberinfrastructure. *Geomatica*, Vol.62, No.1, pp.431-443.
- Blais, J.A.R., D.A. Provins and M.A. Soofi [2006]: Spherical Harmonic Transforms for Discrete Multiresolution Applications. *The Journal of Supercomputing*, vol.38, pp. 173-187.
- Gates, I.D. and M. Bourque [2007]: CI in the Oil and Gas Sector. Cyberinfrastructure Summit, Banff, AB.
- Graham, M.J., M.J. Fitzpatrick and T.A. McGlynn (Editors) [2008]: *The National Virtual Observatory: Tools and Techniques for Astronomical Research*. Proceedings of Summer Schools held at Aspen, CO, in 2004, 2005 and 2006. Astronomical Society of the Pacific, Conference Series, Volume 382, San Francisco, CA.
- Hankin, S., J. Callahan and J. Sirott [2001]: LAS Overview The Live Access Server and DODS: Web Visualization and Data Fusion for Distributed Holdings.
<http://ferret.pmel.noaa.gov/Ferret/LAS/home/las-overview>
- Herrero, S. (Editor) [2005]: *Biology, Demography, Ecology and Management of Grizzly Bears in and around Banff National Park and Kananaskis Country*. Final Report of the Eastern Slopes Grizzly Bear Project, Environmental Sciences Program, University of Calgary, Calgary, AB, 276 pages. Available from www.canadianrockies.net/Grizzly

Layton, J.B. [2008]: Meeting the Petabyte Challenge - How Are We Going to Petabytes? Linux Magazine, April 2008, pp. 42-47, 52-53.

Lehrer, B., K. Verdin and A. Jarvis [2006]: HydroSHEDS Technical Documentation, World Wildlife Fund, Washington, D.C. Available at <http://hydrosheds.cr.usgs.gov> .

Meuer, H.W. [2008]: The TOP500 Project: Looking back over 15 Years of Supercomputing Experience. <http://www.top500.org>

Mutulu, P.M. and J.A.R. Blais [2005]: Space-Time Modeling of Climate Change Effects on the Water Balance of Southwestern Alberta. Poster presentation at the Fall Meeting of the American Geophysical Union in San Francisco, CA.

Mutulu, P.M., J.A.R. Blais, I. Muzik and L. Nkemdirim [2004]: Spatio-Temporal Modeling of Vegetational Transboundary Propagation Due to Climate Change. Poster presentation at the Fall Meeting of the American Geophysical Union in San Francisco, CA.

Newman, R. [2006]: Smart MEMs & Sensor Systems. Imperial College Press, U.K.

NFS [2007]: Cyberinfrastructure Vision for 21st Century Discovery. National Science Foundation Cyberinfrastructure Council.

NSF [2005]: GEON – Cyberinfrastructure for the Geosciences. 2005 Annual Report.

NSF [2003]: Revolutionizing Science and Engineering Through Cyberinfrastructure. Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure.

Papitashvili, V.O., V.G. Petrov, A.B. Saxena, C.R. Clauer and N.E. Papitashvili [2006]: A Virtual Global Magnetic Observatory Network: VGMO:NET. Earth Planets Space, vol.58, pp.765-774.

Pavlis, N.K., S.A. Holmes, S.C. Kenyon and J.K. Factor [2008]: An Earth Gravitational Model to Degree 2160: EGM2008. Presentation at the European Geophysical Union General Assembly 2008, Vienna, Austria.

Pillai, B. [2007]: Cyberinfrastructure Essential to 21st Century Advances in Science and Engineering Education & Research. International Conference on Control, Automation and Systems, Oct. 17-20, 2007, in COEX, Seoul, Korea.

Pinet, M.F. [2003]: Development of a Road Weather Information System (RWIS) Network for Alberta's National Highway System. Intelligent Transportation Systems, 2003 Annual Conference of the Transportation Association of Canada, St. John's, NL. Available at www.tac-atc.ca/English/pdf/conf2003/pinet.pdf

Schwarz, C.R. (ed.) [1989]: North American Datum of 1983. NOAA Professional Paper NOS 2, National Geodetic Survey, Rockville, MD.

Thomas, D. [2007]: GRID: New Business Opportunities? Cyberinfrastructure Summit, Banff, AB.

Top500 [2007]: Top 500 List June 2007. <http://www.top500.org/lists/2007/06>.

Turk, J. and S. Miller [2007]: Google Earth as a Vehicle to Integrating Multiple Layers of Environmental Satellite Data for Weather and Science Applications. Presentation at Fall Meeting of the American Geophysical Union in San Francisco, CA.

Walker, C.K. [2006]: The National Virtual Observatory: A Potential Template for Collecting and Disseminating IPY Data. Invited Presentation in Session IN44A at the American Geophysical Union Meeting in San Francisco, CA.

WDC [1996]: Guide to the World Data Center System. Issued by the Secretariat of the ICSU Panel on World Data Centres. World Data Center A, NOAA/NGDC, Boulder, CO.