U21B MCC: 2002-2004 Tuesday 1020h

Virtual Observatories in Space and Earth Sciences II

Presiding: **P C Cornillon,** University of Rhode Island Graduate School Oceanography**; T Ahern,** IRIS Data Management Center

U21B-01 1020h

The Virtual Heliospheric Observatory: Preliminary Designs

Guan Le[1] (Guan.Le@nasa.gov); Adam Szabo[1] (301 286 5726; Adam.Szabo@nasa.gov); Andrew Davis[2] (ad@srl.caltech.edu); George Ho[3] (George.Ho@jhuapl.edu); Fred Ipavich[4] (ipavich@umtof.umd.edu); Justin C Kasper[5] (jck@space.mit.edu); Davin Larson[6] (davin@ssl.berkeley.edu); Aaron Roberts[1] (Dana.A.Roberts@nasa.gov); Ruth Skoug[7] (rskoug@lanl.gov); John T Steinberg[7] (jsteinberg@lanl.gov)

[1]NASA GSFC, Code 690, Greenbelt, MD 20771, United States

[2]Caltech, MC 220-47, Pasadena, CA 91125, United States

[3]Johns Hopkins University, APL, Laurel, MD 20723, United States

[4]University of Maryland, Dept. Physics, College Park, MD 20742, United States

[5]Massachusetts Institute of Technology, CSR, Cambridge, MA 02139, United States

[6]University of California, Berkeley, Berkeley, CA 94720, United States

[7]Los Alamos National Laboratory, NIS-1, Los Alamos, NM 87545, United States

The ever-increasing volume and diversity of heliospheric data sets along with newly emerging information technology standards make it desirable and possible to develop a distributed data environment for heliospheric data sets that is user-friendly. Preliminary design concepts from the ad-hoc Virtual Heliospheric Observatory (VHO) team will be presented. The current configurations envision the development of only a very thin layer of middleware that will provide a uniform search capability to locate the desired data sets. The data retrieval would take place directly between the data provider and the user. Possible functionalities and requirements on the middleware and on the participating data providers will be discussed.

U21B-02 1035h

Integrated Access to Solar Observations With EGSO

Andre Csillaghy (+41-56-462-4062; csillag@fh-aargau.ch)

University of Applied Sciences Northwestern Switzerland, Dept. of Computer Science, Windisch 5210, Switzerland

Co-Authors: J.Aboudarham (2), E.Antonucci (3), R.D.Bentely (4), L.Ciminiera (5), A.Finkelstein (4), J.B.Gurman(6), F.Hill (7), D.Pike (8), I.Scholl (9), V.Zharkova and the EGSO development team Institutions: (2) Observatoire de Paris-Meudon (France); (3) INAF - Istituto Nazionale di Astrofisica (Italy); (4) University College London (U.K.); (5) Politecnico di Torino (Italy), (6) NASA Goddard Space Flight Center (USA); (7) National Solar Observatory (USA); (8) Rutherford Appleton Lab. (U.K.); (9) Institut d'Astrophysique Spatial, Universite de Paris-Sud (France) ; (10) University of Bradford (U.K) Abstract: The European Grid of Solar Observations is the European contribution to the deployment of a virtual solar observatory. The project is funded under the Information Society Technologies (IST) thematic programme of the European Commission's Fifth Framework. EGSO started in March 2002 and will last until March 2005. The project is categorized as a computer science effort. Evidently, a fair amount of issues it addresses are general to grid projects. Nevertheless, EGSO is also of benefit to the application domains, including solar physics, space weather, climate physics and astrophysics. With EGSO, researchers as well as the general public can access and combine solar data from distributed archives in an integrated virtual solar resource. Users express queries based on various search parameters. The search possibilities of EGSO extend the search possibilities of traditional data access systems. For instance, users can formulate a query to search for simultaneous observations of a specific solar event in a given number of wavelengths. In other words, users can search for observations on the basis of events and phenomena, rather than just time and location. The software architecture consists of three collaborating components: a consumer, a broker and a provider. The first component, the consumer, organizes the end user interaction and controls requests submitted to the grid. The consumer is thus in charge of tasks such as request handling, request composition, data visualization and data caching. The second component, the provider, is dedicated to data providing and processing. It links the grid to individual data providers and data centers. The third component, the broker, collects information about providers and allows consumers to perform the searches on the grid. Each component can exist in multiple instances. This follows a basic grid concept: The failure or unavailability of a single component will not generate a failure of the whole system, as other systems will take over the processing of requests. The architecture relies on a global data model for the semantics. The data model is in some way the brains of the grid. It provides a description of the information entities available within the grid, as well as a description of their relationships. EGSO is now in the development phase. A demonstration (www.egso.org/demo) is provided to get an idea about how the system will function once the project is completed. The demonstration focuses on retrieving data needed to determine the energy released in the solar atmosphere during the impulsive phase of flares. It allows finding simultaneous observations in the visible, UV, Soft X-rays, hard X-rays, gamma-rays, and radio. The types of observations that can be specified are images at high space and time resolutions as well as integrated emission and spectra from a yet limited set of instruments, including the NASA spacecraft TRACE, SOHO, RHESSI, and the ground-based observatories Phoenix-2 in Switzerland and Meudon Observatory in France

URL: http://www.egso.org

U21B-03 1050h

The Solid Earth Research Virtual Observatory

Andrea Donnellan[1] (818-354-4737; donnellan@jpl.nasa.gov); John B. Rundle[2] (530-752-6416; rundle@physics.ucdavis.edu); Geoffrey Fox[3] (812-856-7977; gcf@indiana.edu); Marlon Pierce[3] (812-856-1212; marpiec@indiana.edu); Dennis McLeod[4] (213-740-7285; mcleod@pollux.usc.edu); Jay W. Parker[1] (818-354-6790; Jay.W.Parker@jpl.nasa.gov); Robert Granat (818-393-5353; Robert.Granat@jpl.nasa.gov); Terry Tullis[5] (401-863-3829; Terry_Tullis@brown.edu); Lisa Grant[6] (949-824-5491; lgrant@uci.edu)

[1]Jet Propulsion Laboratory, 4800 Oak Grove Drive, Pasadena, CA 91109, United States

[2]University of California, Department of Physics One Shields Avenue, Davis, CA 95616, United States

[3]Indiana University, Community Grid Computing Lab 501 N. Morton, Suite 224, Bloomington, IN 47404, United States

[4]University of Southern California, 3651 Trousedale Parkway, Los Angeles, CA 90089

[5]Brown University, Department of Geological Sciences, Providence, RI 02912, United States

[6]University of California, Environmental Analysis and Design, Irvine, CA 92697, United States

Under the auspices of NASA we are developing the Solid Earth Research Virtual Observatory (SERVO). The initial focus of the observatory is spaced-based observational data, ground-based sensor data (GPS, seismicity), simulation data, and published/historical fault measurements, coupled with earthquake fault and tectonic modeling and pattern recognition software. This observatory will enable investigators to seamlessly merge multiple data sets and models, and create new queries. In the SERVO framework, simulation data will be archived together with analysis/animation tools and the original simulation code. Observational data, which is heterogeneous and distributed in nature, will be accessible through cooperative federated databases. SERVO will include tools for visualization, datamining and pattern recognition, with data fusion into a web services (portal) based Problem Solving Environment (PSE). SERVO couples distributed data sources, applications, and hardware resources through an XML-based Web Services framework. Users access the services (and thus distributed resources) through Web browser-based Problem Solving Environment clients. The Web services approach defines standard, programming language-independent application programming interfaces, so non-browser client applications may also be built. Initial application codes in SERVO are GeoFEST, a parallel 3D finite element code, Virtual California, a boundary element code to simulate fault interactions, PARK, a boundary element code for studying unstable slip on faults, and two pattern recognizers RDAHMM, which uses a modification of Hidden Markov Models to analyze time series data, and PDPC, which looks for anomalies in seismicity data.

URL: http://www.servogrid.org

U21B-04 1105h

Evolution of SECAA's Services and Role to Support SEC Virtual Observatories and the Missions of LWS

R E McGuire[1] (301-286-7794; robert.e.mcguire@nasa.gov); R M Candey[1] (robert.m.candey@nasa.gov); R A Chimiak[2] (reine.a.chimiak@nasa.gov); S F Fung[1] (shing.f.fung@nasa.gov); J L Green[1] (james.l.green@nasa.gov); D B Han[2] (david.b.han@nasa.gov); B T Harris[2] (bernard.t.harris@nasa.gov); R C Johnson[3] (johnson@mail630.gsfc.nasa.gov); R L Kessel[1] (ramona.l.kessel@nasa.gov); T J Kovalick[3] (kovalick@mail630.gsfc.nasa.gov); H A Leckner[3] (leckner@mail630.gsfc.nasa.gov)

[1]Code 630, NASA Goddard Space Flight Center, Greenbelt, MD 20771

[2]Code 580, NASA Goddard Space Flight Center, Greenbelt, MD 20771

[3]QSS, Code 630, NASA Goddard Space Flight Center, Greenbelt, MD 20771

The successful definition and implementation of a Virtual Observatory (VO) framework to be the future Sun-Earth Connection (SEC) / Solar-Terrestrial data environment will enable a distributed and modular implementation but nonetheless increasingly integrated and transparent data view by end users. All important SEC science data must be easily accessible and scientifically usable, both for specialized studies and across the boundaries of missions and traditional disciplines. The systematic long-term preservation, access and usability of data must be ensured. The SEC Active Archive (SECAA) center in the Space Physics Data Facility (SPDF), in close coordination with the National Space Science Data Center (NSSDC), operates uniquely important multi-mission data services and active archiving activities. Services such as CDAWeb, SSCWeb, OMNIWeb, various FTP data access and display capabilities, and new interfaces to translation-software tools supporting the Common Data Format (CDF) must be extended to function within and actively support the VO framework for this SEC thrust to be ultimately successful. In this presentation, we will briefly review the importance of our existing unique capabilities, our understanding of the VO thrust in SEC and the role we see for multi-mission services of this nature within a VO paradigm and in support of future missions and programs such as Living With a Star (LWS). We will discuss the technical evolution of our present services towards full accessability and continuing growth in scope and functionality to the SEC community as webservices. We will lay out our long-term understanding and vision for our services and activities, including the role and importance of acquisition and other information services to the broad needs of future data management. We will define the issues and directions we now see as critical and approaches to meet these challenges.

URL: http://spdf.gsfc.nasa.gov

U21B-05 1120h INVITED

The Open source Project for a Network Data Access Protocol (OPeNDAP): A Discipline Neutral Approach to Interoperability on the Internet

Peter C Cornillon[1] ((401) 874-6283; pcornillon@gso.uri.edu)

James Gallagher[2] (jgallagher@opendap.org)

Daniel Holloway[1] ((401) 874-6831; dholloway@gso.uri.edu)

[1]University of Rhode Island, Graduate School of Oceanography, Narragansett, RI 02882, United States

[2]The Open source Project for a Network Data Access Protocol (OPeNDAP), 165 Dean Knauss Drive, Narragansett, RI 02882, United States

Data access protocols on which many emerging data systems are being based are often tightly coupled with the semantics of the discipline for which the system is being developed. This greatly facilitates interoperability within the data system, but it also generally results in substantial constrains on interoperability with data systems in other disciplines or, for disciplines with a very large range of data types, on interoperability within the discipline itself. In order to build a data access protocol that may be used by a broad range of disciplines, the Open source Project for a Network Data

Access Protocol (OPeNDAP) has approached the problem by focusing on the syntax of the data rather than on their semantics. In particular, the OPeNDAP data access protocol consists of a rigid prescription for the syntax of the data and a mechanism for the communication of the data semantics, but no constraint on semantic content. In this way, the protocol becomes a base on which discipline data systems may be layered - the developers of a discipline data system may specify the form in which the semantic information is to be communicated while using syntactic machinery that is common to other discipline data systems. The advantage is two-fold. First, discipline data systems are interoperable at the syntactic level and, often with little additional effort can be made interoperable at the semantic level. Second, a larger group participates in developing the underlying protocol than would if it were useful by only one discipline. A by-product of this approach is that it is also much easier to make ones data accessible via the system than if rigid constraints were placed on semantic content. The result has been the adoption of the protocol for use in oceanographic, meteorological, and solar/terrestrial data systems with OPeNDAP servers now at over 50 institutions worldwide.

URL: http://opendap.org

U21B-06 1140h INVITED

Data Management in Planetary Exploration and Space Physics

Raymond J Walker[1] (310-825-7685; rwalker@igpp.ucla.edu)

Steven P Joy[1] (310-825-3506; sjoy@igpp.ucla.edu)

Todd A King[2] (310-206-7201; tking@igpp.ucla.edu)

[1]Institute of Geophysics and Planetary Physics and Department of Earth and Space Science, University of California, Los Angeles, Los Angeles, CA 90095-1567, United States

[2]Institue of Geophysics and Planetary Physics, University of California, Los Angeles, Los Angeles, CA 90095-1567, United States

Planetary exploration and space physics approach data management in very different ways. In this talk we will compare the approaches in these two disciplines with emphasis on how each has dealt with the problems of locating and accessing distributed data. We also will outline the data management challenges each will face in the next decade. Sixteen years ago the NASA Solar System Exploration Division founded the Planetary Data System (PDS) to coordinate the data activities of planetary missions, provide the scientific community with access to planetary data and preserve the data from planetary missions for future analysis. PDS is organized into “nodes” by scientific sub-disciplines (Atmospheres, Geoscience, Plasma Interactions, Rings and Small Bodies) and experimental technique (Imaging and Radio Science). In addition the Navigation and Ancillary Information Facility (NAIF) addresses data issues involving navigation and instrument pointing. All planetary data from NASA missions are prepared to the same metadata standards that include a common data dictionary. Initially access to the data was by sub-discipline although within a sub-discipline the access was to all missions and instrument types. More recently planetary science has become more interdisciplinary and now PDS is moving toward a system that supports cross discipline access. Data are available either on online or on hard media (CDROM or DVD). In recent years space physics data access has been organized by missions. Some missions support data systems through which all of the data from the mission can be accessed while for others the data are available from individual principal investigator sites. In general the space physics missions support an open data policy and much of the data is available online. There are no discipline wide metadata standards. Different missions support different data dictionaries, schemas and interfaces. The data come in a variety of formats. In the near future both planetary science and space physics will be challenged with massive volumes of data from new missions (approximately 10^{15} bytes). In both disciplines the need for comparative research across missions or sub-disciplines is becoming more common. Both will have to support data structures that allow users to readily locate, access, and use data from distributed and diverse sources. Both disciplines must address the issue of how to distribute these massive data sets to the science community. Distributable media (DVD) are too expensive due to their limited storage capacity and the network bandwidth is unlikely to be able to support online distribution.

U21B-07 1200h INVITED

A Virtual Geophysical Network: Using Industry Standard Technology to Link Geographically Distributed Sensors and Data Centers

Tim K. Ahern[1] (206 547-0393; tim@iris.washington.edu)

Rick B. Benson[1] (206 547-0393; rick@iris.washington.edu)

Henry Philip Crotwell[2] (803-777-0955; crotwell@seis.sc.edu)

[1]IRIS DMC, 1408 NE 45th Street #201, Seattle, WA 98105, United States

[2]University of South Carolina, Dept. of Geological Sciences 701 Sumter Street, Columbia, SC 29208, United States

The IRIS Data Management System has long supported distributed data centers as a method of providing scientific researchers access to data from seismological networks around the world. For nearly a decade, the NetDC system used email as the method through which users could access data centers located around the globe in a seamless fashion. More recently the IRIS DMC has partnered with the University of South Carolina to develop a new method through which a virtual data center can be created. The Common Object Request Broker Architecture (CORBA) technology is an industry standard distributed computing architecture. Traditionally used by major corporations, IRIS has developed a Data Handling Interface (DHI) system that is capable of connecting services at participating data centers (servers) to applications running on end-users computing platforms (clients). For seismology we have identified three services. 1) A network service that provides information about geophysical observatories around the world such as where the sensors exist, what types of information are recorded on the sensors, and calibration information that allows proper use of the data, 2) an event service that allows applications to access information about earthquakes and seismological events and 3) waveform services that allow users to gain access to seismograms or time series data from other geophysical sensors. Seismological Data Centers operate the servers thereby allowing a variety of client applications to directly access the information at these data centers. Currently IRIS, the U. of South Carolina, UC Berkeley, and a European Data Center (ORFEUS) have been involved in the DHI project. This talk will highlight some of the DHI enabled clients that allow geophysical information to be directly transferred to the clients. Since the data center servers appear with the same interface specification (Interface Definition Language) a client that can talk to one DHI server can talk to any DHI enabled server. This ability allows one to construct a virtual network of data centers giving one the illusion of a single virtual observatory. Some of the features that will be discussed include direct connections to MatLab and the ability to access globally distributed sensor data in real time.

URL: http://www.iris.washington.edu/DHI

U22A MCC: Level 1 Tuesday 1330h

Virtual Observatories in Space and Earth Sciences III Posters

Presiding: **L Paxton,** Applied Physics Laboratory, Johns Hopkins University**;** **M B Weiss,** Applied Physics Laboratory, Johns Hopkins University

U22A-0001 1330h POSTER

The SuperDARN Data System: An established, Internet Based, Data Access Tool for Ionospheric Physics

Robin J Barnes[1] (240 228 8134; robin.barnes@jhuapl.edu)

Raymond A Greenwald[1] (raymond.greenwald@jhuapl.edu)

Michael Ruohoniemi[1] (mike.ruohoniemi@jhuapl.edu)

[1]Johns Hopkins Applied Physics Laboratory, 111000 Johns Hopkins Road, Laurel, MD 21044, United States

The Super Dual Auroral Radar Network (SuperDARN) is an international network of High Frequency Radars that operate in the Northern and Southern Polar regions in the study of Ionospheric Plasma Convection. Through the SuperDARN web-site hosted at APL (http://superdarn.jhuapl.edu), the network has developed a set of sophisticated, Internet based, data access and analysis tools. The network has also successfully developed real-time data streaming from the Northern Hemisphere radar chain and acts as a space-weather monitor for Ionospheric Convection. Recent work has involved incorporating measurements from a variety of other instruments into the system. The techniques developed for the SuperDARN project can easily be extended to the wider Space Physics Community.

URL: http://superdarn.jhuapl.edu

U22A-0002 1330h POSTER

The Mission Independent Data Layer MIDL Approach to Space Physics Data Interoperability

Jon D Vandegriff[1] (240 228-8558; jon.vandegriff@jhuapl.edu)

Tech Choo[1] (240 228-8874; teck.choo@jhuapl.edu)

[1]Johns Hopkins University Applied Physics Lab, 11000 Johns Hopkins Rd., Laurel, MD 20723, United States

U

We report progress in the development of the Mission Independent Data Layer MIDL, a mechanism for accessing multi-mission magnetospheric and heliospheric datasets. The desire of the space physics community and the LWS program to do Sun-to-Earth integrated science has initiated renewed interest in more capable data systems which can provide simplified access to multiple datasets which are in different formats and from different missions. The first step towards integrated data access is the discovery of data from different datasets, and this aspect of a possible space physics virtual observatory is being addressed by several other projects. The second step in data interoperability is the merging of actual data as discovered by a search engine, and because MIDL provides this capability, it complements other search engine mechanisms. Currently, the MIDL web site (http://sd-www.jhuapl.edu/MIDL) provides visual browsing and direct access to energetic particle data, magnetic field data, and plasma data for 6 missions: ISEE-1, ISEE-2, AMPTE-CCE, Geotail-EPIC, IMP-8, and Cassini. MIDL uses the concept of a data interface to achieve interoperability among datasets with different formats. Details on implementing a data interface for different missions will be discussed, as will the relationship of MIDL to other virtual observatory efforts in space physics.

URL: http://sd-www.jhuapl.edu/MIDL

U22A-0003 1330h POSTER

A Data Grid Framework for Sharing Data Across Distributed Heterogeneous Repositories

J. Steven Hughes[1] ((818)354-9338; Steve.Hughes@jpl.nasa.gov)

Daniel J. Crichton[1] ((818)354-9155; Daniel.J.Crichton@jpl.nasa.gov)

Sean C. Kelly[1] ((972)235-1361; Sean.C.Kelly@jpl.nasa.gov)

Susan K. Lavoie[2] ((818)354-5677; Susan.K.Lavoie@jpl.nasa.gov)

Raymond J. Walker[3] ((310)825-7685; rwalker@igpp.ucla.edu)

[1]Jet Propulsion Laboratory, MS 306-438 4800 Oak Grove Dr., Pasadena, CA 91109, United States

[2]Jet Propulsion Laboratory, MS 168-527 4800 Oak Grove Dr., Pasadena, CA 91109, United States

[3]University of California - LA, Inst of Geophysics & Planetary Physics, Los Angeles, CA 90095-1567, United States

As the volume and diversity of science data sets continues to expand, the ability to share and correlate data across distributed heterogeneous repositories remains a serious challenge. An architectural data grid framework has been developed by the Object Oriented Data Technology (OODT) project to address this challenge and has been successfully deployed in several diverse scientific domains. Characterized by separate technology and data architectures, this framework was used by NASA's Planetary Data System (PDS) to develop the PDS-D distribution infrastructure to provide timely delivery of 2001 Mars Odyssey data to the science community as soon as the data were released from the Odyssey project. PDS-D is now being extended to address the huge data volume increases expected for the Mars Reconnaissance Orbiter (MRO). Based on the success of PDS-D, the Space Physics Archive Search and Extract (SPASE) effort is considering the framework to support a data search and retrieval system for the Space Physics science community. Further demonstrating its flexibility, the framework is also currently being used by the National Cancer Institute's Early Detection Research Network to establish a virtual specimen database by connecting together data from multiple cancer research centers across the country. This paper will highlight achievements and lessons learned from several of the deployments, identify new scientific, technical, and programmatic challenges, and describe how the OODT architectural framework is evolving to meet those challenges.