The Daily Planet™*:* An Internet-Based Information Server for the Atmospheric Sciences Community and the Public

Robert Wilhelmson, Mohan K. Ramamurthy, David Wojtowicz, John Kemp,
Steve Hall, and Mythili Sridhar

Department of Atmospheric Sciences
University of Illinois at Urbana-Champaign
105 S. Gregory Avenue, Urbana, IL 61801
Tel: (217) 333-8650
Fax: (217) 244-4393
e-mail: bw@ncsa.uiuc.edu

1.0 Introduction

The availability of information about the behavior of our atmosphere and about atmospheric science activities on the Internet is of growing importance in making progress in both our understanding of the atmosphere and in weather forecasting. It is also of substantial value to the general public, particularly in informal and formal educational settings, providing access to current as well as historical information in a timely manner and stimulating curiosity in the behavior of the atmosphere. Beginning in the early 90's, we began to develop Internet-based resources for the atmospheric science community that include the popular University of Illinois (UofI) Weather Machine accessible through Gopher and more recently *The Daily Planet*™ *(TDP)*. The latter is accessed through NCSA (National Center for Supercomputing Applications) Mosaic software, which is based on World Wide Web (WWW) technology, and includes access to the Weather Machine. *The Daily Planet*™ is becoming a full-scale Environmental Information Server that will provide transparent access to meteorological, climatological, hydrological, and Earth Observing System (EOS) databases, multimedia educational modules, distributed archives of data sets (both real-time and retrospective), and other Internet-based resources.

2.0 The Weather Machine
(gopher://wx.atmos.uiuc.edu)

The Department of Atmospheric Science at the University of Illinois has a proven track record of providing earth and space sciences data to the public. In the initial implementation of its weather distribution system, raster image products were created and made available, along with textual information, using the X-window System software to any computer or terminal on the Internet (Ramamurthy, et. al., 1992). This information was expanded and made available via a Gopher server, the *UofI Weather Machine,* in January of 1993 (Ramamurthy and Kemp, 1993). Using Gopher client software, any Unix workstation, Macintosh or an IBM-compatible PC on the Internet has a simple, straight-forward way of accessing the *Weather Machine's* products, most of which are updated on a near-real-time basis. The products include surface, upper-air, and operational model forecasts from the National Weather Service, GOES and AVHRR satellite images, severe weather watches and warnings, some climatological data, and other informational documents of interest to the atmospheric sciences community.

The *Weather Machine* has become one of the most highly visible landmarks on the rapidly expanding information super-highway. Time and time again - in newspapers, magazine articles, and other presentations - it has been pointed to as an excellent example of the potential usefulness of the National Information Infrastructure. We have seen a steady increase in the number of requests made to the server. In the past year-and-a-half, server requests have gone from less than 1,000 to an average of over 80,000 per day (Fig. 1). On active weather days, such as during hurricane Emily, the number of calls to this server increased dramatically, exceeding 100,000 daily. The number of organizations connecting to the *Weather Machine* has also grown to include nearly 5,000 different Internet domains, each containing many individual machines. It is being used in research and education. Private citizens, pilots, sailors, skiers, and community centers who have access to the Internet often request information. In addition, the *Weather Machine* is accessed by high schools, community colleges, universities, businesses engaged in computers, networking and publishing, insurance companies, utilities, museums, organizations providing emergency services, media outlets, and several government organizations.

3.0 *The Daily Planet*™ *from the UofI*
(http://www.atmos.uiuc.edu/)

In March, 1994, the *Weather Machine* Gopher server was extended to a hypermedia environment, which we

call *The Daily Planet*™, using Mosaic - a WWW) browser. NCSA Mosaic provides a unified interface to various protocols, data formats, and information archives accessible over the Internet and there are already millions of copies of Mosaic (both public and commercial versions) in use. Currently, *The Daily Planet*™ features real-time weather information (including all Weather Machine data), a collection of lists of other weather servers and sources of weather data, local information about the Department of Atmospheric Sciences' faculty and research, plus a growing number of on-line level hypermedia/multimedia instructional modules. The real-time weather information includes over 200 current maps and images, over 1,100 archived images, and 52 MPEG (Moving Pictures Experts Group that generates standards for digital video and audio compression) animations, many of which are updated hourly. Further, TDP is used to post important information relevant to the atmospheric sciences community. Examples include a link to the NASA maintained Mosaic page describing the status of the GOES-8 deployment and a feature on the GLOBE Project, recently announced by Vice-President Gore. The latter includes a 3 minute video/audio clip of an interview with the Vice-President that was made available within hours of Gore's appearance on ABC's Good Morning America program.

Mosaic is an Internet-based graphical global hypermedia browser that allows the user to discover, retrieve, and display documents and data from all over the Internet. It is part of the WWW project, a distributed hypermedia environment originating at CERN. Global hypermedia means that information located around the world is interconnected in an environment that allows the user to travel through information by clicking on hyperlinks -- terms, icons, or images in documents that point to other related documents. Any hyperlink can point to any document anywhere on the Internet. Mosaic also included forms capability for users to supply information such as that needed in making a database request. Users fill in forms – typing in open fields, clicking on button choices, or choosing a menu item - to build up a complex query, which may then be sent to a database search engine and resolved, with data and other information subsequently sent back to the user. This feature can also be used to supply information to a server collecting data, i.e., a student providing local environmental data to the a Globe Project server. Further information on Mosaic can be found in Schatz and Hardin (1994). Mosaic is licensed software that is provided freely through NCSA. Versions of Mosaic for the Mac, for PC Windows, and for Unix systems can be obtained via anonymous FTP at ftp.ncsa.uiuc.edu under /Web/Mosaic. The Mosaic Demo Page can be accessed from within Mosaic at (http://www.ncsa.uiuc.edu/demoweb/demo.html). Mosaic is also provided commercially by several companies with a variety of enhancements and full support.

Hyperlinked documents in Mosaic are written in HTML, a Hypertext Markup Language. HTML is a subset of SGML (Generalized Markup Language), specialized for simple interactive displays with embedded links. SGML is a specification language for the structure of a document, including headers and references, that has been widely adopted throughout the publishing industry. Information on HTML can be found using Mosaic at the following URL (Universal Resource Locator address): (http://wx.atmos.uiuc.edu/kemp/hotlist.html).

3.1 The Weather World
(http://www.atmos.uiuc.edu/wxworld/html/top.html)

The portion of *The Daily Planet*™ that provides current weather data is called Weather World. The WWW server differs from the Gopher server in that it provides up-do-date animations of a variety of images and for a variety of time periods. Current *animation* products for the United States and vicinity include infrared satellite images, visible satellite images, satellite water vapor images, satellite floater sector images, surface and upper air maps weather maps using WXMAP, 6-panel surface weather maps, and 6-panel ETA and NGM surface and upper air forecast maps out to 48 hours.

This is the largest animation-oriented display of weather information on the Internet to our knowledge, with the updating of images and maps totally automated. To accomplish this we designed an integrated processing system called *upro* (a contraction of "unified product update processor") that handles the processing of all Weather World products. Most of the basic image content in Weather World was already being produced for distribution on our Weather Machine gopher server by a collection of scripts and other processes. The job of *upro* is to gather the output of these processes from the gopher server directories and other places and to reprocess them into full sized images, small images (used for icons and samples in the HTML pages), image archives and MPEG animations. The HTML pages that provide access to these products are also considered products because they always contain new information including new images.

Each product is defined in a product description. The product description contains information such as the product's unique name, the type of product (image, MPEG, HTML page), the location of the input data (such as in one of the Weather Machine gopher directories) and other type specific characteristics. For all image types these characteristics include items like output image size, cropping, labeling and even options to add a raised boarder around the edge of a reduced

image to use as an icon with a three-dimensional appearance. MPEG and archive type characteristics include number of frames or number of images saved, etc.

HTML products also contain a template for the html page to be produced. This template contains normal HTML plus special layout macros that help to keep the style consistent and to keep references to other products such as images. These references are replaced by the actual URL of the latest instance of a given product. This keeps the HTML menus on the server in synchronization with the products available through them.

There is also the capability to use template products. These template products form a class-like hierarchy that simplifies product definition. All products need not redefine every characteristic. A parent class for that product type can hold default characteristics while the specific product description holds only information unique to that particular product. As many levels can be added to the hierarchy as desired.

Upro keeps all of this information in its internal database and is launched automatically about once an hour. It scans its database and looks for products that need updating. An MPEG satellite loop may need to have a new frame added to it, for example, when a new image has appeared in its input directory (in this case, one of the image directories on the gopher server). The image is processed and added to the animation. If this particular animation is set to hold only the last 24 frames, the oldest frame is removed to make room for the new one. The new MPEG file is then placed in one of the TDP server's directories. The HTML page that references it is also updated to reflect the newly updated animation. This sequence is repeated in a similar manner for all other products.

Nearly all of our weather data files (before and after processing) have the time and date encoded into the filename. *Upro* can interpret this (via a filename format specification in the product description) and use this information to better track the files and organize them properly.

Because it maintains information in its database about the contents of both input and output directories as well as the products, *upro* can detect changes in any of these places and respond accordingly. It can take note of new files in input directories and sense the removal of files in output directories. For example, if one were to start randomly deleting files from the WWW server directories, these files would automatically be replaced during the next *upro* run. With the product definitions safely backed up, the system is fully self-recoverable from major problems. In fact, we've purposely deleted the entire server directory structure in rare instances to force a complete rebuild.

Efficiency is of major concern. If there are products that need to be updated every hour and it takes more than an hour to process them all, the server would certainly not be able to keep up with incoming data. One of our approaches has been to distribute the load (by task) across multiple machines. Input data is generated and stored on two machines, while output data is served to the WWW on a third machine. A fourth machine sits in the middle of the chain running *upro*. The other approach has been to develop special software to increase the efficiency of certain computationally intensive tasks such as MPEG production in particular. At present it takes approximately 45 minutes to process each hour's worth of data on the machine running *upro*.

A future version will be more comprehensive and use a more sophisticated database, possibly even to store the image and animation data itself.

3.2 Multimedia Modules Available in *The Daily Planet*™ (http://www.atmos.uiuc.edu/covis/modules/html/module.html)

Internet-accessible multimedia instructional modules that introduce and explain a variety of important concepts in atmospheric sciences are available in *The Daily Planet*™. They consist of text, colorful diagrams, animations and movies, audio, and scanned images, that introduce and explain a variety of important concepts in atmospheric sciences. These multimedia instructional modules are being developed for use at the high school level, but are also useful for general undergraduate education (Ramamurthy and Wilhelmson, 1993; Ramamurthy et al., 1994). The modules are being tested at the two current CoVis schools in the Chicago area, and they are being revised and refined based on the feedback from them (Ramamurthy et al., 1995). Such multimedia-based instruction provides an alternative approach to learning, one in which the student, through interaction with the computer, becomes actively involved in the learning process that includes current weather data.

The Pressure and the Forces and Wind modules include descriptions of high and low pressure centers and the balance of forces that generate winds. These are enhanced through the use of colorful diagrams and animations, video clips, and audio narration. A module entitled Guide to Weather Maps and Images provides important information on understanding many of the weather displays available in TDP. We have also developed a hypermedia Glossary for the modules that have been developed thus far. Other modules currently under development include: (1) Cloud Catalog, (2) Guide to Atmospheric Optics, (3) Tornado Spotters Guide, and (4) Severe Storms Guide. The ultimate goal

is to deliver an entire multimedia textbook over the Internet for use by students and the general public.

4.0 Future Development

The growth in data available over the Internet has been astronomical and with the availability of data through such programs as EOS will continue to grow. It is vital that appropriate information be locatable by an interested researcher, educator, or the general public. For the most part, data archives and digital libraries in earth sciences have been generally established to aid scientists in carrying out research. Typically, scientists know a lot about the type of data they are studying or have the ability to find what they need to know. Further, they generally have the skills to deal with different data formats, user interfaces, and query requirements, and they have considerable computer resources available to handle the massive volumes of data which might have to be filtered in order to obtain the desired data. However, even they will have difficulty locating useful information within the growing number of Internet data servers. Mosaic development and digital library research is currently underway at the University of Illinois to address these needs.

Recently, support from NASA has been obtained to test applications and digital library technologies in Support of Public Access to Earth and Space Science Data. This joint work involves the Department of Atmospheric Sciences, NCSA, and the Computer Science Department faculty and staff at the University of Illinois. Data from the earth and space science community (including supplementary information and education modules) will be utilized to test server technologies needed to support effective access to the data and information. These technologies will address the issue of scalability needed to deal with the growth in available data. In the data management area, the focus is on integrating data from different sources without undergoing costly data conversion and the need for rapid access to parts of very large data sets. For information technologies, work is being undertaken to provide the users with the server-side tools needed to find the information they desire, to interact with it, and to analyze it. The scalable server technologies merges the other technology areas, addressing problems of dealing with large and numerous files on web servers along with tertiary storage issues related to these files. In addition, client software development, the only component directly seen by the user, will include Mosaic enhancements and associated software development needed to improve the use of images in providing hyperlinks and hypermedia and in overlaying and subsetting of data and images.

The Daily Planet™ will serve as the major initial testbed of the new software developed. A prototype interface, designed in Mosaic, will allow users to browse the available metadata and select subsets of this data to be delivered in either HDF or netCDF formats for downloading. The available data would initially include GOES and AVHRR processed and value-added data and images. The amount of data available from on-line will be significantly increased in order to assess scalability and tertiary storage technology developments. This will be accomplished using the above data together with additional datasets (specifically DMSP or SMM/I data and UARS).

The Daily Planet™ will also incorporate software developed to allow the comparison or overlaying of data in order to examine relationships. An example would be to overlay AVHRR derived vegetation data with SSM/I derived precipitation data to note the relationship between rainfall amount and vegetation cover or to overlay GOES water vapor data with precipitation to note their relationship. A Mosaic-based interface that extends the data browse and subsetting features would allow selections from different datasets to be compared. New Mosaic features such as the extended GIS (Geographical Information System) hypermedia interface would also be incorporated in *The Daily Planet*™.

Finally, through other funding and collaborations, new multimedia modules, new weather products, and additional climate data will be added to *The Daily Planet*™ . This will include data from the Midwest Climate Center and other midwest hydrologic data, as well as flood and water quality information. *The Daily Planet*™ will be adapted to include environmental data collected in the Globe Project and adaptations will be made to maximize its usefulness in K-12 education in both the urban and rural settings and to improve scientific literacy both nationally and internationally.

5.0 Acknowledgments

6.0 References

Ramamurthy, M. K., K. P. Bowman, B. F. Jewett, J. G. Kemp, and C. Kline, 1992: A Networked desktop synoptic laboratory. *Bull. Amer. Meteor. Soc.*, **73**, Cover and 944-950.

Ramamurthy, M. K., and J. Kemp, 1993: The Weather Machine: A Gopher server at the University of Illinois. *STORM*, 1(3), 34-39.

Ramamurthy, M. K., and R. B. Wilhelmson, 1993: A networked multimedia meteorology laboratory. *Proceedings of the Second Symposium on Education*, Anaheim, California, American Meteorological Society.

Ramamurthy, M. K., R. B. Wilhelmson, S. Hall, M. Sridhar and J. G. Kemp, 1994: Networked multimedia systems and collaborative visualization,

Proceedings of the Third Symposium on Education, Nashville, Tennessee, American Meteorological Society.

Ramamurthy, M. K., R. B. Wilhelmson, R. D. Pea, L. M. Gomez, and D. C. Edelson, 1995: CoVis: A national science education collaboratory. *Proceedings of the Fourth Symposium on Education*, Dallas, Texas, American Meteorological Society.

Schatz, B. R. and J. B. Hardin, 1994: NCSA Mosaic and the World Wide Web: Global hypermedia protocols for the Internet. *Science*, Vol. 265, 12 August 1994, 895-901.