

METEOFORUM – EXTENDING REAL-TIME DATA SHARING THROUGHOUT CENTRAL AND SOUTH AMERICA

Tom Yoksas*

Unidata Program Center/UCAR, Boulder, Colorado, USA

Waldenio Gambi de Almeida

Centro de Previsão de Tempo e Estudos Climático, INPE, Brazil

David Garrana Coelho

Universidade Federal do Rio de Janeiro, Brazil

Vilma Castro Leon

Universidad de Costa Rica, Costa Rica

Timothy Spangler

COMET/UCAR

1. INTRODUCTION

The Unidata Program Center (UPC) of the University Corporation of Atmospheric Research (UCAR) is involved in several international collaborations whose goals are extension of Unidata's proven Internet Data Distribution (IDD) technology to research and educational institutions throughout the Americas.

The international expansion of the IDD is, in part, an outgrowth of the *MeteoForum* pilot project (Spangler and Fulker, 2001, Laing, et. al., 2002, Yoksas, et. al., 2004) being conducted by the Unidata and COMET programs of UCAR. *MeteoForum* is designed to assist WMO Regional Meteorological Training Centers (RMTCs) to improve service in the areas of hydro-meteorology, agriculture, and disaster management. By accessing comprehensive collections of training materials and real-time and historical data, the RMTCs will be able to enrich the education they offer to hydro-meteorological professionals in their region.

A collaboration with Brazil's Centro de Previsão de Tempo e Estudos Climáticos (CPTEC, a division of INPE), the Universidade Federal do Rio de Janeiro (UFRJ), and the Universidade de São Paulo (USP) has resulted in the expansion of the Unidata North American IDD into Brazil through the creation of the *IDD-Brasil*, a data sharing peer to the Unidata IDD that will first provide access to real-time hydro-meteorological data to Brazilian universities and then to universities throughout South America.

A similar collaboration with the Universidad de Costa Rica (UCR) has extended the North American IDD into Central America. At the time this paper was

being developed, new collaborations with the Universidad de Puerto Rico-Mayagüez and the Caribbean Institute for Meteorology and Hydrology (CIMH) in Husbands, St. James, Barbados were being launched to investigate the expansion of the IDD throughout the Caribbean. The combination of these three efforts is aimed at the creation of a peer of the IDD and *IDD-Brasil*, the *IDD-Caribe*.

The *IDD-Brasil* and incipient *IDD-Caribe* are now delivering, in near real-time, the full set of global observations and NCEP and UKMET global model output available in NOAAPORT, and all GOES-East imager channels available in the IDD-conveyed Unidata-Wisconsin data stream to top-level redistribution nodes established at prominent national universities and research institutes.

It is envisioned that the Unidata North American IDD, the *IDD-Brasil*, and the *IDD-Caribe* data sharing efforts will foster new collaborations among universities, WMO RMTCs, and national meteorological agencies throughout North, South, and Central America.

This paper is intended to provide an overview of the UPC efforts to facilitate access to a broad spectrum of scientific data and the establishment of a network of cooperating institutions through which those data may be shared.

2. UNIDATA'S ROLE

The international expansion of the IDD began in earnest as the first phase of the *MeteoForum* pilot project. Unidata's role in *MeteoForum* is particularly well aligned with its primary mission, the provision of:

Data, tools, and community leadership for enhanced Earth-system education and research

*Corresponding author address: Tom Yoksas, Unidata/UCAR, PO Box 3000, Boulder, CO 80307; e-mail <yoksas@unidata.ucar.edu>

Unidata conducts the following activities in *MeteoForum*:

- Facilitate data access to a broad spectrum of observations and forecasts
- Coordinate a data-relay network that collects and distributes data in real-time at no cost to educators and researchers
- Build a community where data, tools, and best practices in education and research are shared
- Support faculty at research and educational institutions in the use of Unidata systems

3. INTERNATIONAL DATA RELAY INAUGURATION

Each of the RMTCs located in WMO Regions III (South America) and IV (North America) are co-located, or closely aligned with prominent national universities:

Argentina	Universidad de Buenos Aires (UBA)
Barbados	University of the West Indies (UWI)
Brazil	Universidade Federal do Pará (UFPA)
Costa Rica	Universidad de Costa Rica (UCR)
Venezuela	Universidad Central de Venezuela (UCV)

Since Unidata's forte is interacting with the university community, we have concentrated our efforts of extending real-time hydro-meteorological data delivery to RMTC associated universities.

Brazilian participation in the IDD was inaugurated in fall 2001 simultaneously at the Universidade Federal do Rio de Janeiro (UFRJ) and the Universidade Federal do Pará (UFPA). At the UFRJ, university personal were searching for ways to access meteorological data that were needed for a variety of research projects. Co-author David Garrana learned about the availability of real-time data through the IDD; installed the Unidata LDM-5 (Davis and Rew, 1994) in the Laboratório de Prognósticos em Mesoescala (LPM) at the UFRJ; and began receiving data.

In the summer of 2002 Unidata installed the LDM-5 at the UFPA to test the feasibility of delivering GTS observational data, model output, and GOES-East satellite data in near real-time to the RMTC co-located on the UFPA campus in Belém. Results reinforced previous observations (Chiswell, 1999) that the data delivery engine behind the IDD, the LDM-5, was inefficient when relaying data between machines that are electronically distant. Counter-intuitively, relaying data to a sequence of intermediate hosts actually improved the end-to-end performance of the IDD.

Lessons learned in the UFPA data relay tests were combined with independent efforts at the UFRJ, the Hong Kong University of Science and Technology, and at the University of Melbourne (Melbourne, Victoria, Australia) in a redesign of the LDM at the UPC. The redesign effort resulted in a next-generation LDM, the LDM-6 (Emmerson, 2003), that is able to relay substantial volumes of data to both local and remote sites with little to no latency (the time difference from when a product is first injected in the IDD and the time the product is received). Stress testing at the UPC has demonstrated that the LDM-6 is capable of relaying over 1.5 terabytes of data per day without introduction of product latencies.

4. NETWORK INFRASTRUCTURE IS CRUTIAL

An element critical to the success of real-time data delivery is the availability of good-quality, high-speed networking at participating institutions. High speed Internet2 interconnects between Brazil and the US have been made possible by connectivity provided by Brazil's Rede Nacional de Ensino e Pesquisa (RNP) and the Academic Network at São Paulo (ANSP) through the NSF-funded America Pathways (AMPATH, Ibarra, 2003) project of Florida International University (FIU) and Global Crossing.

Since one of Brazil's Internet2 connection points is housed in Rio de Janeiro, and since the UFRJ was already ingesting data as an IDD receive-only node, we approached the UFRJ with a proposal that they act as a top-level IDD relay node, initially for the RMTCs at the UFPA and the UBA in Buenos Aires and then throughout Brazil (Yoksas, Garrana, 2002). The UFRJ expressed an eagerness to act in this role on behalf of their fellow institutions by relaying data through improved departmental networking and computing infrastructure. Equipment that they purchased was installed in the UFRJ 24x7 Network Operation Center (NOC).



Figure 1. UFRJ 24x7 Network Operations Center

After the installation, stress testing of the UFRJ IDD ingest capability was conducted jointly by Unidata and UFRJ personnel. Figure 2 shows hourly data volumes received by the UFRJ IDD node over a ten day period at the end of December 2003. The time series plot shows that over 1.5 GB of data was received each hour during the test, and peak ingest rates exceed 2.7 GB per hour.

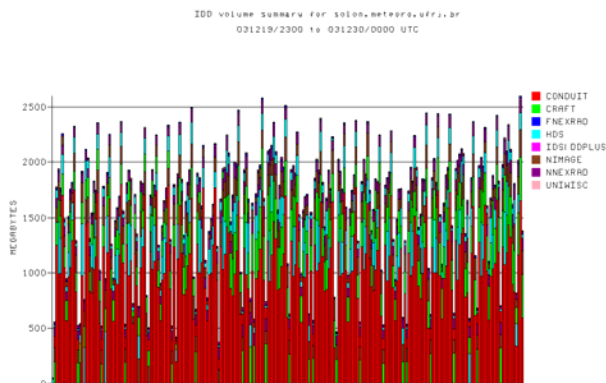


Figure 2. Time series plot of ingest data volume as a function of time and color coded by data type.

Figure 3 shows a time series plot of GTS global observations data product latencies for products received during the UFRJ stress testing.

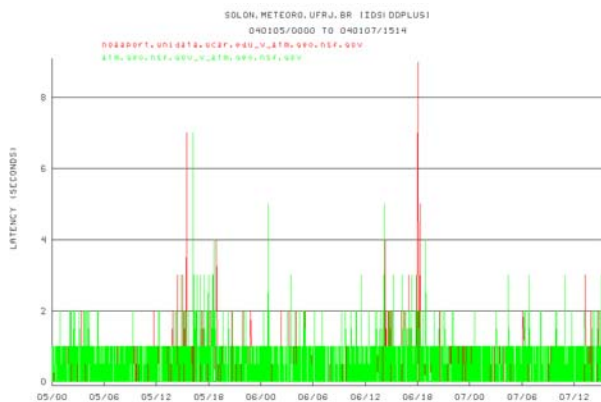


Figure 3. GTS global observation latencies measured during UFRJ IDD stress testing.

Figure 3 shows that the great majority of products were received with latencies of one second or less.

The ability to relay virtually all of the data that is available in the IDD (no restricted data were used in the testing) convinced us that the UFRJ could assume a leading role in real time data dissemination in Brazil.

In the fall of 2003, an additional connection was made that has strengthened the data relay capabilities in Brazil. Co-author Waldenio Gambi de Almeida, a member of the data section of the CPTEC (<http://www.cptec.inpe.br/>), attended the fall series of Unidata training workshops in Boulder, Colorado.



Figure 4. CPTEC facility located at Cachoeira Paulista, Rio de Janeiro, Brazil.

While in Boulder, Mr. Gambi was made aware of the data relay experiment being jointly conducted by the UPC and the UFRJ. Since the IDD offers a means by which CPTEC can receive a reliable stream of real-time hydro-meteorological data, and since it also provides a mechanism that allows sharing of locally held datasets with other IDD participants, CPTEC joined in the data relay experimentation. Like the UFRJ, CPTEC is connected to Internet2 by a high speed connection through AMPATH.

The ability and willingness of the UFRJ and CPTEC to actively participate in the **IDD-Brasil** peer expansion of the Unidata IDD are direct results of their excellent network connections to Internet2. As more institutions get better connected to the high-speed Internet, more will be able to participate in data sharing efforts like the **IDD-Brasil** and communities like Unidata.

5. INAUGURATION OF THE IDD-BRASIL

Three distinct efforts made it possible to move substantial volumes of data from the US to electronically distant IDD sites in Brazil or back:

- Brazil's connection to Internet2 through the AMPATH project
- the establishment of a working relationship among Unidata, the UFRJ, and CPTEC

- the modernization of the Unidata LDM to Version 6

The combination of these three advances is allowing Unidata North American IDD participants to relay large volumes of data to the UFRJ and CPTEC with very low latencies (typically one to a small number of seconds) and will allow the South American participants to share locally-held datasets with their North American counterparts. The datasets being moved routinely to Brazil include high resolution NCEP model output (the IDD CONDUIT and HRS streams), high resolution GOES-12 satellite imagery (the IDD UNIWISC stream), and GTS global observation data (the IDD IDS|DDPLUS stream).

Part of the establishment of the **IDD-Brasil** was the drafting of a set of *principles of participation*:

- For the most part, participants must be educational institutions
- Participants must acquire and maintain appropriate computer hardware and Internet access
- Real-time data must be relayed free-of-charge
- Cost of participation is sharing of locally-held datasets with fellow participants
- Top-level relays must take ownership of the expansion and support processes

6. IDD-BRASIL EXPANSION

The first institution to receive **IDD-Brasil**-relayed data was the UFPA and its associated RMTTC. Soon thereafter, data relays were established between the UFRJ and CPTEC so they could act as each other's real-time data ingest backups and share data-relay duties.

Efforts aimed at broadening participation in the **IDD-Brasil** are currently underway. Co-authors Waldenio Gambi de Almeida (CPTEC) and David Garrana Coelho (UFRJ) have been promoting the benefits of participating in the **IDD-Brasil** and in the usefulness of Unidata display and analysis systems through discussions with a variety of Brazilian universities.

In the summer of 2003 Dr. Elen Cutrim, a professor at Western Michigan University, member of the Unidata User's Committee, and former department chair and WMO RMTTC head at UFPA, visited Brazilian universities to promote the use of various Unidata offerings like real-time data delivered by the IDD and

data analysis and display using the Unidata-developed Integrated Data Viewer (IDV).

In the late summer of 2004, Dr. Cutrim and lead author Tom Yoksas attended the Congresso Brasileiro de Meteorologia (CBMET, the semi-annual meeting of the Brazilian Meteorological Society) in Fortaleza, Ceará, Brazil to make university attendees aware of the availability of real-time data through the **IDD-Brasil** and of the various display and analysis packages that Unidata makes freely available to research and educational institutions. CPTEC demonstrated the availability of real-time data being conveyed in the **IDD-Brasil** in their CBMET booth, and the system attracted remarkable attention. **IDD-Brasil**-received data was visualized by Unidata GEMPAK, and visitors were made aware of useful documentation like the CPTEC-hosted **IDD-Brasil** web page (<http://solon.meteoro.ufrj.br>) that contains Portuguese versions of LDM-6 and GEMPAK installation instructions. The workstation running the LDM-6 and GEMPAK was the most visited point in the entire congress, and free access to hydro-meteorological data for education and research through the **IDD-Brasil** was a major topic of discussion.

The first university outside of Brazil to connect to the **IDD-Brasil** is the Universidad de Avarez in Portugal. The next site will likely be the Universidad de Buenos Aires. UBA's participation has been enabled by recently gained access to Internet2 and improvements in departmental networking infrastructure made possible through a Unidata Equipment Grants program award.

Additional information on the expansion of the **IDD-Brasil** is presented in a companion paper by co-author Waldenio Gambi de Almeida (Almeida, et. al., 2005).

7. CENTRAL AMERICA DATA SHARING EXPANSION

In parallel with the establishment of the **IDD-Brasil**, the UPC approached the University of Puerto Rico at Mayaguez (UPRM) to enlist their help in relaying real-time data to Caribbean and Central American universities. Like the UFRJ and CPTEC, the UPRM was enthusiastic about playing an expanded role in the Unidata IDD.

Where delivery of real-time data over the Internet is not practical, and when a university site is within the NOAAPORT broadcast footprint, the UPC has recommended installation of satellite-based data reception systems.



Figure 5. 7 meter satellite dish portion of Unidata NOAAPORT reception system installed at the UCR.

In February 2004, the UPC worked with the Universidad de Costa Rica to install a UPC-designed and built NOAAPORT satellite ingestion system on the UCR campus in San Jose, Costa Rica. Since the installation, the UCR has been able to ingest real-time global observations and NCEP model output for use in education and research. The UCR has agreed in principle that, as its Internet connectivity improves, it will assume a leading role in extending access to its real-time meteorology data to Central American universities that also have sufficient Internet connections.

Very recently, the UPC has begun working with the Caribbean Institute of Meteorology and Hydrology (CIMH), a WMO RMTTC, to test IDD-delivered real-time data into Barbados. While it is too early to comment in detail on this effort, we have observed that IDD delivery of data is possible, but not spectacular given the limited network connection (a dedicated 256 Kbps link) that the CIMH currently has to the Internet.

The success of the incipient data distribution/sharing efforts among the UPRM, UCR, CIMH, and Unidata university community, named the *IDD-Caribe*, will depend entirely on the quality of network connections available at participating sites.

7. COMMUNITY BUILDING BY DATA SHARING

The existing North American Unidata community has grown into the vital entity that it is through a variety of efforts spearheaded by the NSF-sponsored UCAR Office of Programs Unidata project:

- Establishment of an easy-to-use data distribution network that provides access to a

wide variety of real-time, atmospheric science-related datasets

- Provision of data analysis tools and high quality support services for use of those tools
- Encouraging of active participation through community-based governance and guidance

The same community building approach is being extended to research and education institutions first in Brazil and eventually throughout WMO Regions III and IV. The establishment of the *IDD-Brasil* represents a first step in this community building process, and the creation of the *IDD-Caribe* represents a second.

The second phase of building a community is encouraging the sharing of locally-held datasets by all participating sites with colleagues throughout the extended Unidata community. The need for full and open international exchange of environmental data has been articulated in many documents. In order to understand the myriad complexities of the earth system and the way the different elements interact, it is crucial to collect, share and analyze environmental data from all parts of the world.

The US Unidata community recognizes the importance of access to environmental data from other countries, particularly those in the southern hemisphere, to enhance education and research capabilities toward studying global problems like climate change, ozone depletion, and ENSO. A critical requirement for such research studies is the acquisition and assimilation of a complete spectrum of global meteorological, oceanographic and hydrological observations. We view the data relay infrastructure that has been installed in Brazil, Costa Rica, Puerto Rico and Barbados as the beginnings of a hemispheric-wide network that can act as conduit for multi-way sharing of locally held data sets among peers in an expanded Unidata university community.

8. NEXT STEPS

The UPC plans to continue to play an active role in establishing a robust IDD relay node for South America in collaboration with the UFRJ, CPTEC and the USP. We intend to continue to work to establish a data sharing network for Central America and the Caribbean. In addition, we will be informing university representatives attending the Red de Universidades del Pacifico Sur (RUPSUR, the South Pacific University Network) 2004 meeting in Santiago, Chile of the availability of Internet-delivered, real-time hydro-meteorological data through the Unidata North

American IDD and Brazilian-based **IDD-Brazil** networks.

We envisioned that the Unidata North American IDD, the **IDD-Brazil**, and the **IDD-Caribe** data sharing efforts will foster new collaborations among universities, WMO RMTCs, and national meteorological agencies throughout North, South, and Central America.

9. ACKNOWLEDGEMENTS

The Unidata Program Center is funded by the US National Science Foundation (NSF). COMET is funded by the US National Weather Service (NWS).

10. REFERENCES

Spangler, T. C., D Fulker: **MeteoForum** – An International Network of Meteorological Training Centers for the 21st Century, A Pilot Proposal developed jointly by the COMET® and Unidata Programs of the University Corporation for Atmospheric Research, Boulder, CO, U.S.A., March 2001.

Laing, A. G., L. Miller, T. C. Spangler, D. Fulker, 2002: **MeteoForum**: A Pilot Project for Advanced International Meteorological Training Centers, Preprints, 18th International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, 326-327.

Yoksas, T. et al: **MeteoForum –Initial Successes in Data Sharing Leading to the Creation of the IDD-Brazil**. 20th International Conference on IIPS for Meteorology, Oceanography, and Hydrology, Seattle, WA, 2004.

Davis, G. P., and R. K. Rew, 1994: The Unidata LDM: Programs and Protocols for Flexible Processing of Data Products, Proceedings 10th International Conference on IIPS for Meteorology, Oceanography, and Hydrology. 131-136.

Chiswell, S. R. and L. Miller, 1999: Toward Expanding Model Data Access Through Unidata's Internet Data Distribution System, Preprints, 15th International Conference on IIPS for Meteorology, Oceanography, and Hydrology, 441.

Emmerson, S: Differences between LDM-5 and LDM-6, Unidata seminar series presentation, Unidata Program Center, UCAR Office of Programs, July, 2003, PPT.

Ibarra, J. – Florida International University/AMPATH: Personal Communication, April, 2003.

Yoksas, T, and D. Garrana Coelho, Personal Communication, October, 2002.

de Almeida, W., L. A. de Carvalho, S. H. S. Ferreira, D. Garrana Coelho, M. G. Justi, and T. Yoksas. Perspectives on Internet Data Distribution Expansion and Use in Brazil, 21st International Conference on IIPS for Meteorology, Oceanography, and Hydrology, San Diego, CA, 2005.

