P2.6 CONDUIT AND LEVEL II DATA DISTRIBUTION: LEVERAGING THAT WORKS FOR COLLABORATIVE PROJECTS

Linda Miller*, Steve Chiswell, Steve Emmerson, Jeff Weber, Tom Yoksas
UCAR Unidata, Boulder, Colorado

1. INTRODUCTION

Serving a diverse community of users, the Unidata Program Center (UPC) seeks collaborative opportunities to provide data, tools and software for education and research in earth system science. Two projects that began as grassroots efforts with the U. S. Weather Research Program (USWRP), including several collaborating institutions, are highlighted in this paper. Similar to the early days of Unidata, both projects were born out of the need for researchers and educators to access data that were not yet available through NOAA. Nearly eight years ago, Unidata embarked on what sometimes was a rocky road to high resolution model data access from the National Center for Environmental Prediction (NCEP). Discussions with Kelvin Droegemeier, University of Oklahoma, and others regarding the importance of WSR-88D Level II radar data, began about the same time.

Meeting the challenges and the tribulations during the early phases of the collaborations through CONDUIT (Cooperative Opportunity for NCEP Data Using IDD Technology), the Local Data Manager (LDM) and Internet Data Distribution (IDD) system paved the way for distribution of high-resolution model data sets that were not currently available through NOAAPORT (the satellite communication system providing NWS data). Thanks to the cooperation of NOAA's National Weather Service (NWS), NCEP, and several institutions willing to be initial test sites, the data are now flowing to 77 hosts at 47 unique domains.

The Collaborative Radar Acquisition Field Test (CRAFT) project followed a path similar to CONDUIT's. The Level II data were not accessible to the community. Archived data access at the National Climatic Data Center (NCDC) frustrated the community, due to the lag between ordering and receiving the data. Through the cooperation and determination of the University of Oklahoma, University of Washington, Unidata, NOAA, and academic and private sector stakeholders, the project began to move forward. The tornado outbreak of 3 May 1999 underscored the importance of the CRAFT project and its underpinnings which moved the data to NCDC for archiving and subsequent availability to the community. Other mechanisms broke down, but the data were sent to NCDC for archiving during the prototype test phase of the CRAFT project during the horrific tornado outbreak. It's difficult to point to such a devastating event and say that it provided the necessary visibility for the project, but to a degree, it is true.

At the core of both projects is Unidata's Local Data Manager (LDM) technology which delivers over 20.5 TB of data per week in aggregate, via the Internet Data Distribution (IDD) system. Examples of how the community stepped forward to facilitate distribution of CONDUIT and Level II data and how the data are being used, including data distribution volumes and the overall structure that makes it work to serve the user community are discussed below.

2. COMMUNITY, A BROAD REACH TO COLLABORATION

Community building and collaboration at Unidata are everyday activities. The Unidata users community has been involved in collaborations with the UPC, particularly since 1995, the advent of Unidata's Internet Data Distribution (IDD) system. This project combined LDM technology to replace satellite communications at university sites throughout the community and transformed data distribution techniques for research and education. Universities stepped up to the challenge of providing relay nodes within the IDD topology to move the data and cooperate with others to make the project one of the most prominent community efforts in Unidata's history. Since that time, university sites have viewed that collaboration as a *quid-pro-quo*, where if they give, they receive. Sites from as far away as Brazil and Portugal are participating in the IDD and are receiving what is commonly referred to as the CONDUIT data stream. Within the CONDUIT and Level II data

*Corresponding author address: Linda Miller, UCAR Unidata, PO Box 3000, Boulder, CO 80307; e-mail lmiller@ucar.edu.

user community, the reach has extended beyond the traditional meteorological user community to hydrology, ornithology, field projects, spotter operations, and many others.

3. CONDUIT

Early discussions with NMC, now NCEP, the Scientific Steering Committee (SSC) and the Interagency Working Group (IWG) of the USWRP led to the CONDUIT project. The essence of CONDUIT is access to high-resolution model data that are not currently available through NOAAPORT. The data are considered experimental data, and most of the data are currently available on NCEP servers. CONDUIT provides the data through the LDM/IDD "push" technology, providing the data as soon as it becomes available, thus eliminating the constant "pulling" that is required when using FTP protocols to obtain the data.

In the early days of CONDUIT, there were issues with more data than could be handled with the Internet and LDM technology available. When Abilene and Internet2 came along, they provided a valuable testing ground for the IDD and projects like CONDUIT. There are now over 200 entities, including universities, the U.S. government and international partnerships, and corporate and affiliate memberships involved in Internet2. This evolution, along with an enhanced version of the LDM has led to a successful CONDUIT project for high level data users. There are currently 77 hosts and 47 individual sites using CONDUIT high resolution data for the initialization of regional models, such as MM5 and WRF, among other things. There is currently about 28 Gigabytes of CONDUIT data moving to these sites every day.

4. WSR-88D LEVEL II DATA

Access and distribution of Level II data began in a similar way through the USWRP, but it took a more circuitous route to achieve the goal. The collaborative project, coined CRAFT (Collaborative Radar Acquisition Field Test), became a household term as interest grew among the community to find a way to make the data broadly available.

The University of Oklahoma had been receiving some of the radar data, but had the vision of receiving the data from all radars. Unidata played the important role of providing the LDM technology and advocating for access to the data by the university community. The National Climatic Data Center (NCDC) archived the Level II data, but the 8mm tape drives were an inefficient and expensive method of archiving data. Through the CRAFT collaboration and leveraging of technology (i.e., HDSS mass storage system, link to the Abilene network, receiving the data via the IDD, and launching a web-based graphical browsing tool), NCDC was able to save money and provide convenient access to radar data in a user-friendly manner in a matter of minutes, rather than the weeks it took previously.

The full resolution data from all NEXRAD radars are compressed in real-time to about 250 Megabytes per hour using bzip2 and are distributed using the LDM technology with the Internet2 Abilene backbone. As testing continued, more stakeholders from all sectors of the community became interested in participating and gaining access to the data.

The science and testing being performed with the Level II data were convincing enough to gain the interest of NOAA in providing the data to NCEP. NCEP took steps to formulate requirements based on NWS needs for the Level II data related to modeling and forecasting needs. That was an important step in NOAA's interest to take over the management of the data distribution. Previously, NWS regional headquarters received their regional radar data, but distribution was limited to those radars only. Timing was right, due to the recent membership of NOAA in Internet2, making it possible for them to adopt the underlying data distribution techniques, and leverage the testing and work that had been done through the CRAFT project to make Level II data available to the broad community of users via the Internet.

5. CONCLUSION

These two projects have been successful because they involved several partners with common needs and goals. Each project was able to leverage from the other to serve their particular interests and communities. Collaboration is not easy, because each party needs to experience at least some modicum of benefit from a collaborative effort. Within the Unidata community, there are 47 members using CONDUIT data and 56 using Level II data. We view this as a great success.

6. REFERENCES

Droegemeier K., et al., Project CRAFT: A Test Bed for Demonstrating the Real Time Acquisition and Archival of WSR-88D Level II Data, 18th International Conference on IIPS, 1/2000, Long Beach, CA

Dissemination and Utilization of NEXRAD Data in the Unidata Community-Steven R. Chiswell, 21st Conference on Severe Local Storms, 8/2002, San Antonio, TX