UNIDATA Newsletter

Volume 2 Number 2	The Quarterly Newsletter of The UCAR UNIDATA Program	July 1987

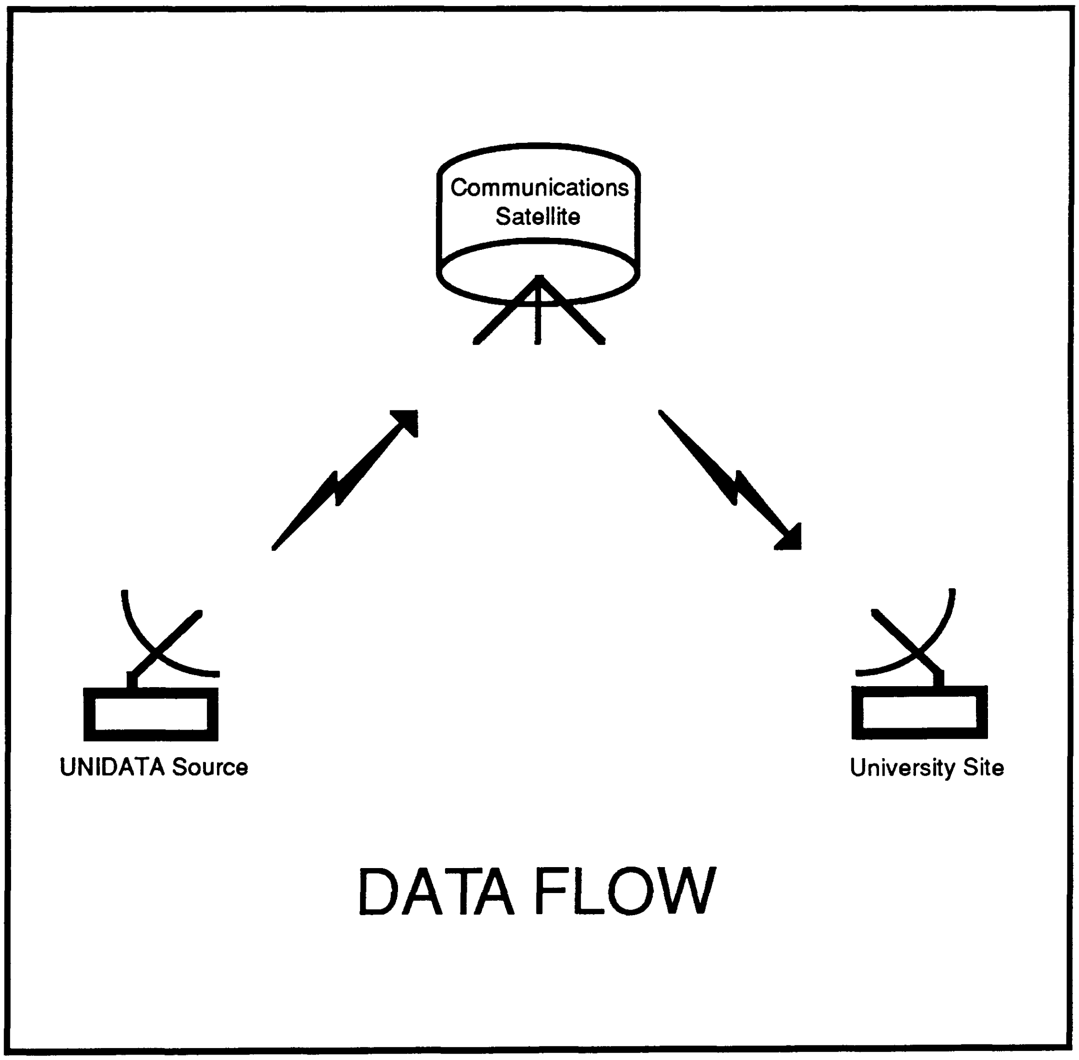


Table of Contents

The UNIDATA Newsletter is published quarterly by the UNIDATA Program Office (UPO). UNIDATA is a cooperative University Project managed by the University Corporation for Atmospheric Research and is funded by the National Science Foundation. Address all correspondence to:

UNIDATA PROGRAM
UCAR
P.O. Box 3000
Boulder, CO 80307

STAFF:

Name	Position	Phone Area 303	OMNET/Tele-Mail Address
Bunting, Martine	Secretary	497-8644	
Davis, Glenn	Systems Administrator	497-8643	
Domenico, Ben	Program Manager	497-8877	B.DOMENICO
Fulker, David	Program Director	497-8696	D.FULKER
Green, Bob	User Support	497-8641	UNIDATA.SUPPORT
Kassinger, Susan	Administrative Asst	497-8695	S.KASSINGER
Rew, Russ	Systems Development	497-8845	R.REW
Vacanti, Frank	Documentation	497-8688	UNIDATA.DOC

Director's Report

The flow and management of data is the focus of this issue of the newsletter. In a guest article, Rolland Hauser of Cal State, Chico, describes experiences of those volunteers who are testing reception of the Wisconsin/Unidata broadcast in PC-McIDAS systems. In this case, data are obtained and organized by the University of Wisconsin, conveyed to Chicago for satellite broadcast by Zephyr Weather Information Service, received on television receive-only satellite dishes at the test sites, and stored on the disks of IBM PC/ATs. These data are received by the users in quasi real time, and they represent the first use of Unidata's capabilities for broadcasting information that has been tailored to the needs of the university community.

In addition to capabilities for capturing and storing the Wisconsin data, the ATs are outfitted with an early release of the McIDAS software, which provides a menu-driven set of functions for processing and displaying the data in 16 colors on the AT's enhanced graphics monitor. The content (data products) of the broadcast will soon be determined by the users themselves, and future versions of the software will provide greatly increased functionality over the current test release.

The PC-McIDAS test is demonstrating a form of data flow in which a single workstation receives information directly from the Zephyr system, and must do so continuously so that nothing is missed. A more general concept is described in the article titled "Basics of Data Flow through the Unidata System." The capture and organization of Zephyr data is accomplished by a local data management subsystem at each Unidata site. In turn, the LDM functions as a "server": it provides data-management services to a number of workstations, all attached to a local area network. (Variations of the LDM/LAN model are described in the article titled "Shared Local Data Management.")

This network model frees workstations to perform analysis and display functions with maximum efficiency, and concentrates the problems of real-time data capture and organization in a system specifically designed to do so.

The prototype UNIDATA LDM system is presently undergoing tests to assess the suitability of its design for meeting long-range needs. We anticipate that the production version of the LDM system will be able to receive all of the Zephyr (digital) data feeds and simultaneously support access to those data by some 20 workstations. Furthermore, the LDM design will permit future reception of high-speed services with data rates up to 56 kilobits per second.

Although some users may choose to operate in the single-workstation mode, we expect most UNIDATA sites to base their systems on the network model. The data-flow article discusses further advantages of the network model, based on a two-way connection between the local network and a new national network, designated NSFnet. This external "long-haul" networking will permit UNIDATA workstations to access national resources such as NCAR and the NSF supercomputer centers. It will also permit communications between users and will bolster UNIDATA's support infrastructure.

The article "Overview of the UNIDATA/Zephyr Broadcast" provides details about the Zephyr broadcast system. Presently, some 60 sites are participating in the UNIDATA/Zephyr program, and they are obtaining data at deeply discounted rates. In general, the addition of a single card in the receiver cage is all that is required to obtain the special UNIDATA/Wisconsin information. (Until August, however, only the designated test sites may obtain the PC-McIDAS software. This was done to assure the quality of both the software and the broadcast data.)

I hope you find these technical expositions interesting and helpful. The UNIDATA system is now taking shape in concrete ways, and those of you who are planning to participate must soon choose a system configuration that will satisfy your requirements and constraints. We intend to provide the information you need for these decisions, and this newsletter is a first step in that direction. Thank you for your interest.

Dave Fulker, Director

Manager's Report

In the last issue, we reported that the staff at the University of Wisconsin had successfully tested the special UNIDATA/Wisconsin data stream by broadcasting data through the Zephyr system back to themselves. Now users at six McIDAS Broadcast Evaluation sites and the UPO are receiving the UNIDATA/Wisconsin data stream on PC-McIDAS systems. The guest article by Rolland Hauser fills you in on their experiences.

GEMPAK/LDM

UNIX/VMS system development work is also moving forward. Two subgroup meetings served to pin down interface specifications, testing plans and decoder priorities that were still somewhat unclear at the March meeting of the Implementation Working Group (IWG). Those plans were reviewed at the IWG meetings of June 11-12. The result is a revised schedule for implementing the GEMPAK/LDM system, which emphasizes more user testing of system components at selected sites early in the integration process, and calls for deployment of the tested systems to begin in the second quarter of 1988.

The portions of the prototype LDM that process domestic data service (DDS) bulletins are being tested in the UPO, using live data from the Zephyr DDS line. The last week in June, DI_Sys, the subcontractor developing the LDM, is scheduled to deliver the corresponding software and hardware for handling gridded data from NMC. Work on adapting GEMPAK to use the new data-access routines continues at NASA Goddard.

Graphics

The UPO is coordinating the UNIDATA graphics efforts of Purdue and Colorado State with the ongoing work at NCAR's Scientific Computing Division (SCD). A draft of the UNIDATA graphics goals serves as the basis for discussion among the groups. A specific plan for achieving the strategic goals will be worked out in a series of meetings between UNIDATA and SCD representatives.

Both Purdue and the UPO are using a new release of the GKS-based NCAR graphics package as the foundation of their work. Early implementations of some utilities in the new release have been done on MSDOS systems at Purdue and on UNIX systems under X Windows in UPO.

Broadcast Protocols

The UNIDATA Broadcast Protocol Committee has written and formally adopted a protocol that will ultimately be used to handle the high-speed UNIDATA broadcast (56 kbits/s). While there are no specific plans for implementing a broadcast using this protocol in the immediate future, we anticipate that such a broadcast will be needed as the demand for more high-resolution satellite images increases.

Documentation

A draft of the data-flow specification for the overall UNIDATA system has been written. It served as the basis for many of the articles in this issue of the newsletter.

Staff Changes

Martine Bunting started work as the UPO secretary on April 29. Russ Rew, head of systems development in UPO, successfully defended his Ph.D. dissertation in the Computer Science Department at the University of Colorado.

Ben Domenico

The First McBE Report

It's still in its early days, but the UNIDATA McIDAS Broadcast Evaluation (McBE) has begun. On April 30 and May 1, representatives of California State University at Chico, Northern Illinois University, State University of New York at Albany, University of Alabama at Huntsville, University of Arizona, University of Kansas, and UNIDATA User Support met at the Space Science and Engineering Center (SSEC) in Madison, Wisconsin.

There we were briefed on the IBM PC/AT McIDAS hardware and software system. Each institution was provided with UNIDATA Level 1 workstation software, version 2.1, in the form of five high-density floppy diskettes of executable code.

The SSEC also copied an assortment of demonstration files for our use during system checkout and while configuring our ATs for the receipt of live McBE data. The demonstration files included:

1. Dithered visible images for the nation prepared from merged east and west GOES data and remapped in polar stereographic projection.

2. Similar images of pseudo-colored temperature-sliced infrared data.

3. Similar images prepared from manually digitized radar data.

4. Twenty-four national-scale grids of nested grid model output from NMC, which is contoured at the local workstation site.

5. Fourteen regional-scale graphics of hourly airways surface observations.

6. The text of hourly airways observations in coded form for display by state.

The first update of the workstation system software was received by mail in mid-May.

Each of us arrived home to excited audiences who were prepared, of course, for instant gratification. From early reports each institution has a different tale of installation experiences.

At the University of Alabama Huntsville, Mike Goodman reports that installation proceeded smoothly, with routine operation beginning about May 12. Initially, there was a problem with the buffer box between the receiver and the AT. It kept overflowing, for reasons which Mike still doesn't understand; the version 2.1.1 software appears to have cured the problem. The system is being used by students, particularly in conjunction with lectures and in a lab for freshmen on the subject of weather forecasting.

Arizona became active about May 28, after Bob Gall cannibalized three other machines for chips and put a fire under his purchasing office to bring the AT up to full UNIDATA hardware specification. For a few hours, the McBE was connected to the wrong serial input port. The system has been used by several faculty, staff, and students. There is a "lot of interest and enthusiasm" for the concept. Unfortunately, spring semester ended before the system could be used in classes.

Joe Eagleman at Kansas has used only the demonstration files so far. His students are "anxious to get at the live data." His attempt to operate under "double DOS" didn't work.

At Northern Illinois Tim Spangler installed his AST card on June 1, at which time his system was up and running as advertised. His attempts to use Intel memory expansion boards didn't work. He has installed Ethernet and finds it to be "complicated to administer." It's not clear (my fault) whether his UNIDATA McBE AT is a node on the Ethernet local area network. He and his colleagues have met and discussed further McBE products they would like to see and evaluate this summer.

Rich Pyle has returned to Albany, having battled spiders and snakes while maintaining the SUNY (Albany) lightning network. His AT is not yet configured properly for either the demonstration files or live data. He is now trying to encourage his purchasing office to throw off their cobwebs and help him out.

(Continued)

Our efforts at Chico State brought us live data on May 29. Our labors included adding a final 0.5 Mbytes to our AST board, discovering that IBM delivered only 128 Kbytes on their EGA board instead of the ordered 256, and replacing our two-port serial card setup in the Zephyr network (which had been received with just one four-port serial card). When these hardware problems were resolved, everything functioned immediately and well. The students in an upper-division laboratory course were enthusiastic about the demonstration package. The receipt of live data began only after finals week was over, unfortunately. The transition from version 2.1 to version 2.1.1 was painless and virtually effort free. The updated software, which displays the new-style graphics files and which McBE started broadcasting on June 1, is noticeably faster than the first version.

All in all, the initial McBE efforts are moving ahead well. The satellite transmissions are received in a remarkably error-free condition. Upwards of 60 Mbytes of test-pattern transmission have been received at Chico State without a single bit error. The highest total has been received from Zephyr at Madison, where more than 100 Mbytes of test-pattern data were received with fewer than 30 errors.

Congratulations are in order, in our opinion, to the University of Wisconsin and to the UNIDATA Project Office. The job so far has been well done. Onward, upward, and forward!

Rolland Hauser
California State College at Chico

Basics of Data Flow through the UNIDATA System

Any description of how data flows through our system involves understanding several specific components: communications, LANS, NSFnet considerations, and LDM configurations. Let us look at each of these in turn.

Intersite Communications

The most global aspect of data flow in the UNIDATA system involves what are called "long-haul connections." Figure 1 shows the two on-line methods for moving data in the UNIDATA system: one-way broadcast and two-way communication. A third approach (not shown here) involves off-line portable media, such as magnetic tape and optical disks. (The April 1987 issue of the UNIDATA Newsletter contained a sizable article on state-of-the-art optical-disk technology as it applies to storing large meteorological data sets.)

The one-way broadcast of current weather data includes:

- NWS domestic and international data services;
- NMC product circuit-containing gridded analyses and forecasts;
- DIFAX and NAFAX (digital and analog facsimile maps);
- Specially tailored UNIDATA information, includ–ing satellite images and radar data, prepared by the University of Wisconsin.

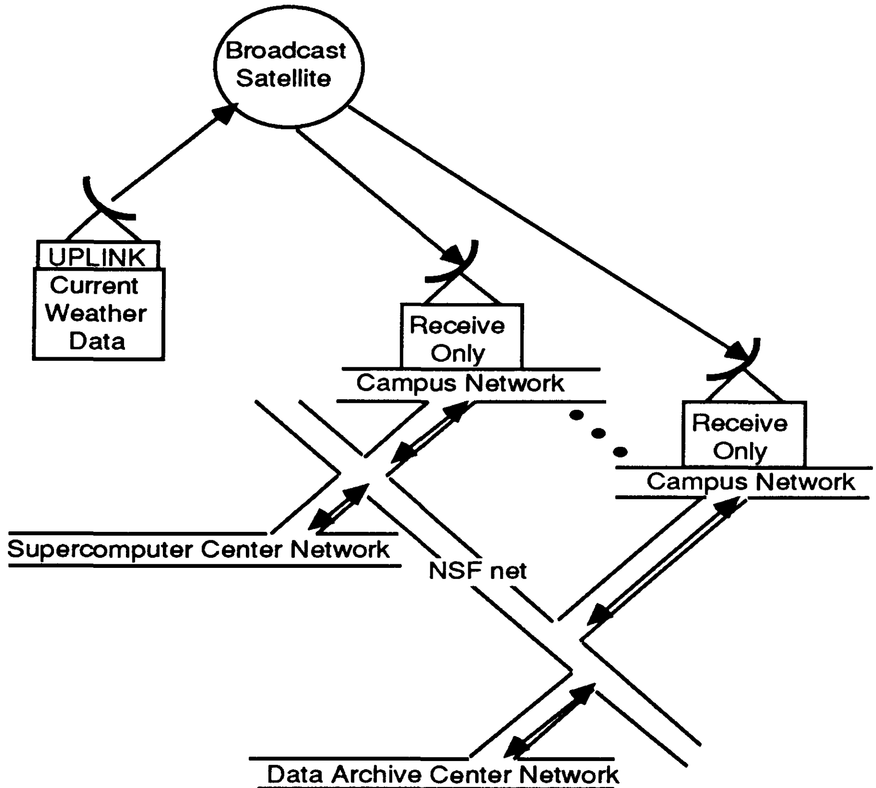
Most of these services are already being delivered (via satellite broadcast) to UNIDATA sites under a contract with Zephyr Weather Services.

As Figure 1 shows, the planned two-way communications between campuses will rely on NSFnet, the national "network of networks." This two-way communication system will allow for computer-to-computer communica-tion, and it will provide:

- access to supercomputer centers;
- access to major data archives such as those at NCAR;
- electronic mail communications with the UNIDATA Program Office and with other UNIDATA sites;
- interactive traffic (remote login) and file transport among computers at different UNIDATA sites.

(Continued)

LONG HAUL CONNECTIONS



(Figure 1)

For more information on UNIDATA two-way communication, see the UNIDATA Brochure, which can be obtained from the UNIDATA Program Office for a copy.

Local Area Networks

Once data have been received at any given site, they need to be made available to each workstation. Individual UNIDATA workstations are interconnected via a departmental LAN at each site. Also on the LAN is the UNIDATA LDM system which decodes the incoming data stream. Campus networks and department LANs may offer facilities for reading magnetic tape and for printing and graphics. Thus, individual workstations potentially have access to a wide variety of capabilities, including:

- data exchange with other systems on the departmental LAN;
- communication with other systems on the campus network, and, through NSFnet, systems on other campuses;
- broadcast weather data;
- access to centers of archived data;
- access to supercomputing centers;
- access to shared facilities for producing graphics and printed output, or for reading and writing magnetic tapes.

(Continued)

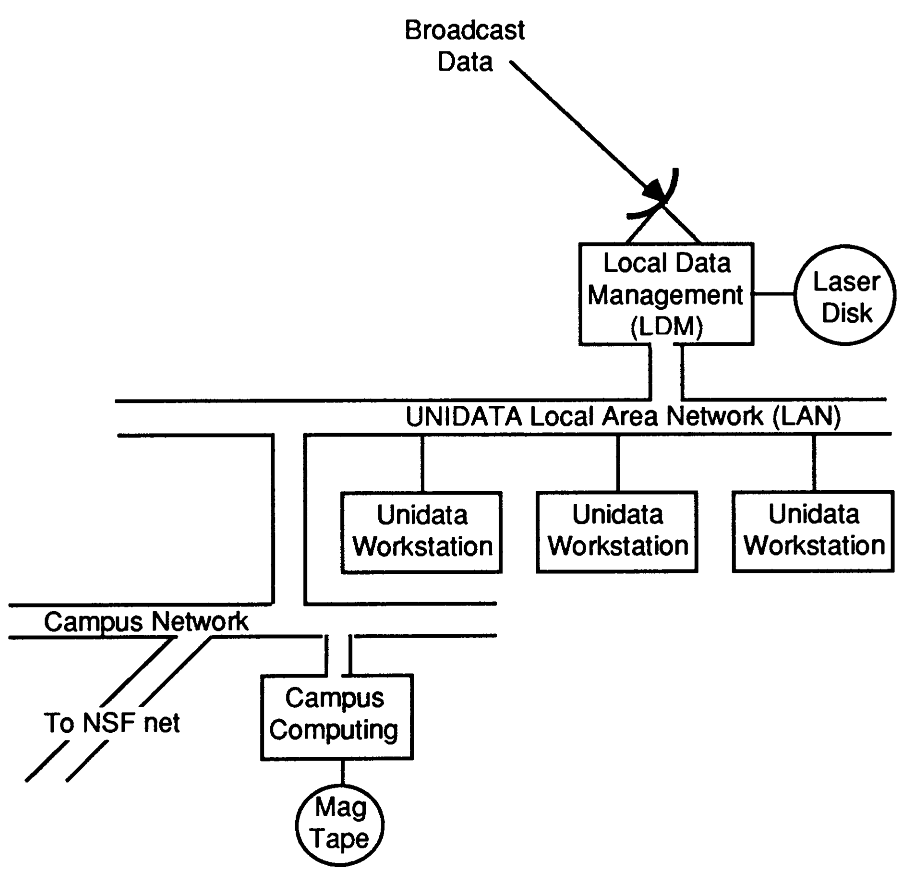
Local Data Management

The LDM system, under development here at UPO, builds a local archive of data assembled from a variety of sources (see Figure 2; below for the LDM configurations.) The data typically stored by LDM will include:

- current weather data ingested from the satellite broadcast system;
- local data (possibly from experimental systems at the University);
- data collected from other NSFnet sites, e.g., data centers, supercomputer model runs, field experiments, archives at other universities;
- data brought in on portable media, e.g., magnetic tape, diskette, optical disk.

With the above overview in mind, let us now move on to see how data flow through the UNIDATA system. ■

CAMPUS NETWORK CONNECTIONS



(Figure 2)

Shared Local Data Management: Two Approaches

LDM is central to the UNIDATA system design and we approach it in two ways. One serves only IBM PC/AT computers running under the MSDOS operating system, while the other approach serves a variety of computers running under the UNIX and VMS operating systems.

Before we begin, we need to make one important clarification concerning the PC-McIDAS software. All references in this article to PC-McIDAS describe the ideal configuration: a PC-McIDAS system operating on a local area network. This ideal has not yet been realized, as there are technical challenges yet to be overcome (e.g., fitting all the LAN software and PC-McIDAS software in the MSDOS's 640K program memory space). In fact, we do not anticipate offering this LAN configuration until UNIDATA Phase IV, which commences in 1988. Until then, PC-McIDAS systems will operate in a "stand-alone" mode, not attached to any network.

The concept of local data management addresses the question, "How do I manage (i.e., capture, organize, and access) incoming data in my own local computer environment?" Many UNIDATA sites will have several computer workstations accessing the same data. Rather than performing data-management functions on each workstation, we have chosen to create a separate LDM system that captures data from the broadcast system, decodes and organizes these data, and stores them in such a fashion that they can be accessed by all workstations on the LAN.

In this configuration, the data are held on the disks of the LDM computer, which is called the "server" system. The individual workstations on the network are called the "client" systems. Users of any client workstation may obtain current data from the LDM server.

Using the PC-McIDAS program is relatively simple because PC-McIDAS will be running only on MSDOS systems. This means that only one data format will be in use. A program like GEMPAK, however, will run on a variety of computers, under both UNIX and VMS. Different brands of computers store data in different formats, and therefore data stored on the server in any given format may be incompatible with certain computers.

In the MSDOS case, local data management uses "transparent file server" technology. This technology permits a collection of MSDOS workstations on a LAN to access a single common disk system. For the user of the workstation, files on the common system are nearly indistinguishable from those stored directly on the workstation; that is, the network seems transparent. Our LDM scheme will use this technology to permit a number of MSDOS (PC-McIDAS) workstations to read data simultaneously from common files in which broadcast data have been stored.

Unfortunately, an equivalent form of transparency is not available for networks of dissimilar computers with UNIX, VMS, and MSDOS operating systems. The principal reasons for this are (1) the computers may vary with regard to their internal representations of numbers and character strings and (2) the rules and concepts of file handling differ among the operating systems. Therefore, we have chosen to implement the LDM for UNIX and VMS systems (i.e., for use with GEMPAK) based on a primitive mechanism for transferring blocks of arbitrary binary data across the network.

This mechanism is the remote procedure call (RPC): one computer on a network requests actions by another computer, which returns the results as a string of bytes. For the UNIDATA LDM, the results will generally represent selected blocks of information from common storage, where the data have been stored in a well-defined structure. This circumvents problems related to variations in the syntax of files, but there remain difficulties related to the internal representation of numbers and character strings.

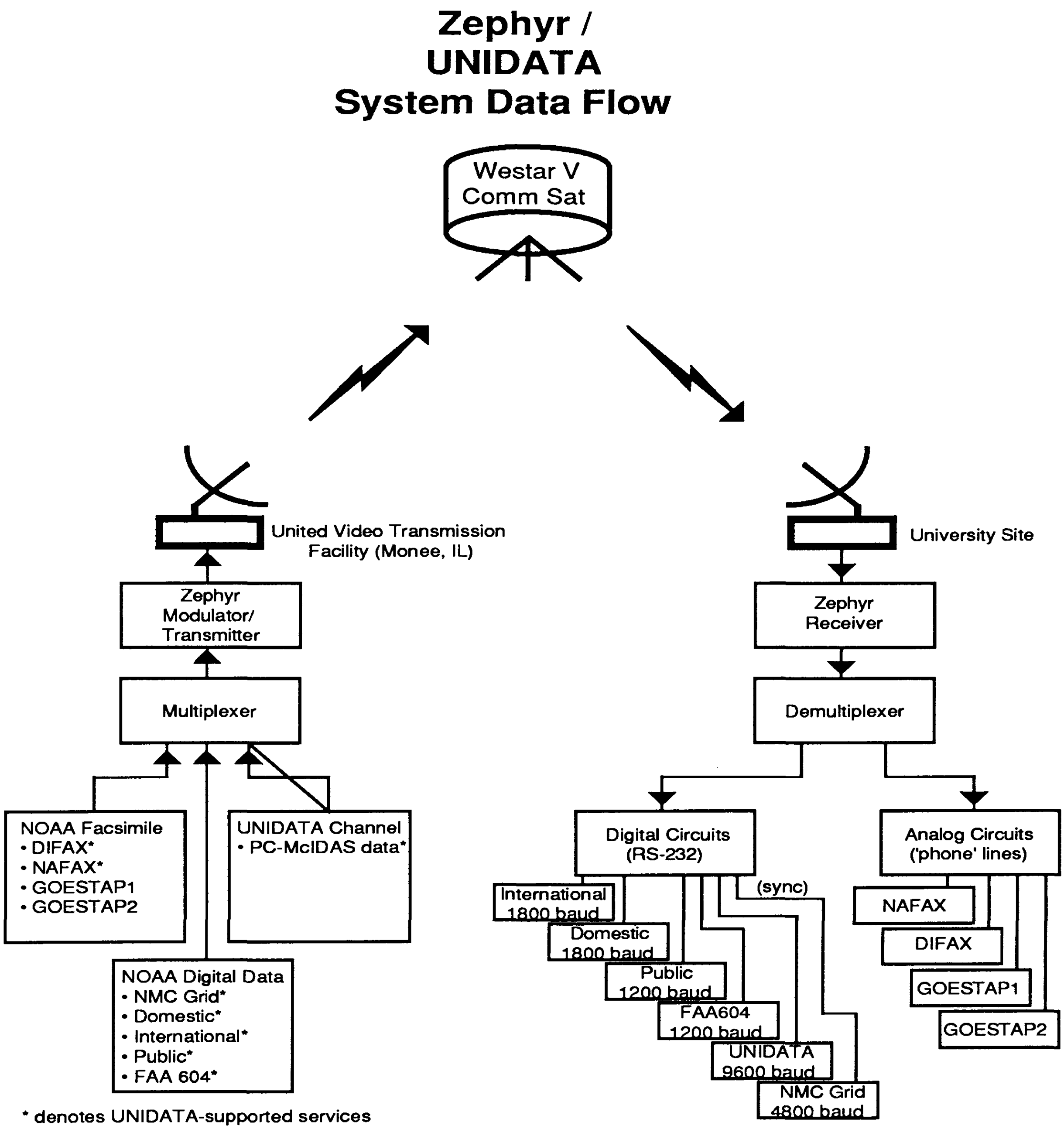
To cope with this problem, we will use the concept of external data representation (XDR). With XDR, computers can transmit and store numbers and character strings in a well-defined (external) form that can be translated easily to and from the computer's internal representations for these values. The combination of RPCs and XDR will permit the LDM concept of shared central storage to be implemented in a heterogenous network of computers. ■

Overview of UNIDATA/Zephyr Broadcast

Meteorological data are broadcast to UNIDATA participants by Zephyr Weather Information Services, Inc. With funding from NSF, UCAR has contracted Zephyr to provide certain services to educational institutions at reduced end-user rates.

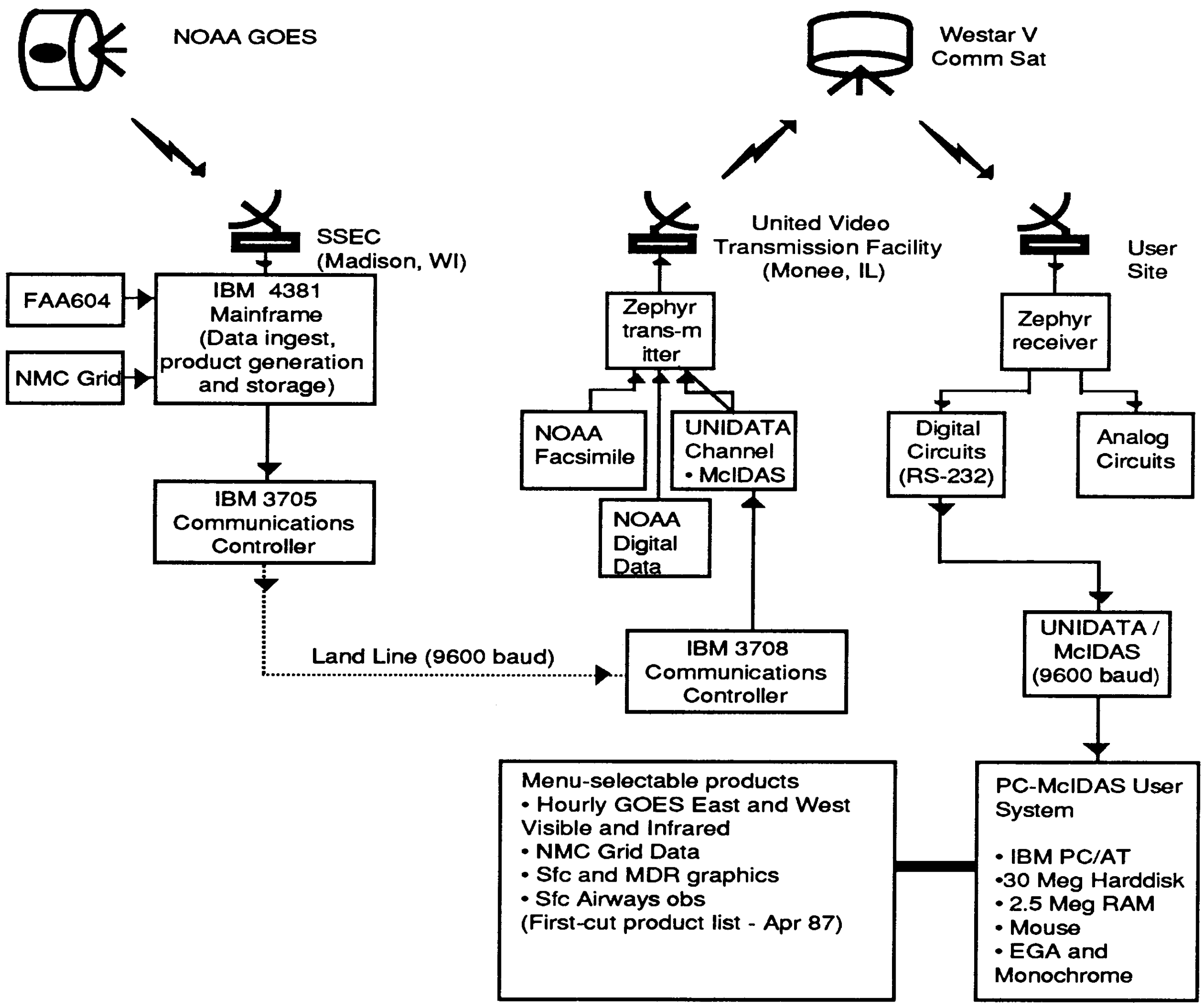
As shown in Figure 3, Zephyr gathers its broadcast data from several sources. These sources fall into three classes:

(Continued)



(Figure 3)

McIDAS / UNIDATA System Data Flow



(Figure 4)

(1) NOAA facsimile data, including DIFAX, NAFAX, and GOESTAP. (The GOESTAP information is not presently available at UNIDATA pricing.)

(2) NOAA digital data services (NMC gridded, domestic, international, public), and the FAA604 circuit. Most of these services provide text in the form of coded bulletins about temperature, pressure, etc. The NMC gridded-data service provides meteorological parameter fields from the numerical prediction models.

(3) A special UNIDATA channel, the contents of which are prepared by the University of Wisconsin and contain satellite imagery, radar data, and other products tailored for use on UNIDATA's PC-McIDAS system.

Zephyr's transmitter uplinks the data to the Westar V Communications Satellite that serves all the user sites. The university receiver stations split the signal into digital and analog data. With the exception of the NMC Grid 4800-baud line (which is synchronous), all the digital circuits are RS-232 asynchronous.

(Continued)

As Figure 4 shows, the satellite images are taken by the NOAA GOES satellite, processed through Wisconsin's Space Science and Engineering Center, and sent via land line to the Zephyr transmitter in Monee, IL. In fact, only a subset of these image data is broadcast because of the bandwidth limitations (9600 bits per second) of the UNIDATA channel. (This bandwidth may possibly be increased in the future.) Other data are also transmitted by Wisconsin as part of the UNIDATA channel; these include information from the FAA604 and NMC grid services, merged with the satellite data and organized for ease of processing on the PC-McIDAS system. This UNIDATA information stream is being evaluated by the McIDAS Broadcast Evaluation Committee.

The PC-McIDAS system accesses weather bulletins and NMC grids via the McIDAS data stream. Other computers (under VMS and UNIX) access them through the LDM system.

Obtaining Broadcast Services

Each university requests its data services directly from Zephyr. Note that certain services are not covered under the UNIDATA agreement and are available at the commercial rates from Zephyr. Configurations of Zephyr services and equipment requirements (with associated costs) can vary quite a bit. Three major factors will help you derive ballpark cost estimates.

The first factor is equipment, i.e., a satellite dish, receiver, cables, etc. The cost for equipment is the most flexible because it depends on the desired number of services and on whether you want to lease or buy the equipment. Leased equipment from Zephyr costs about $100-$200 per month. If you opt for purchasing your equipment, the cost will be determined by the number of services to be received and will range from $2,500 to $6,000.

The second cost consideration is equipment installation, which of course involves just a one-time fee. Based on our experience, those people who have had Zephyr install their dishes (as well as all other equipment) have had fewer data-reception problems. Any future system-debugging is also greatly expedited when one single vendor has been responsible for all equipment installation. Zephyr's installation fee may vary depending on the region of the country and on the specific equipment being installed; as a point of reference, the UPO paid Zephyr $575 in early 1987 for complete installation of all equipment.

The third factor to consider is the subscription costs of actual data services. This can vary a bit, too, depending on the data you want. For ballpark estimates, you may subscribe to any three low-speed text services for a combined price of $75 per month. Bear in mind, however, that exceptions exist. For example, if you require the NMC product service, a $100 monthly premium will be added to that $75.

Configuration Problems

Some sites may not be able to have the satellite antenna and receiver located near the equipment and computer area due to zoning or site-beautification regulations. In such cases, local circuits or phone lines and associated modem equipment may be required between the satellite receiver and the user processing systems.

Another special case involves users who want to receive the NMC numerical-grid product service. The products are transmitted in a synchronous protocol that may require special equipment on the user's computer system for proper processing. Also, the special UCAR data channel operates at a higher data rate (9600 baud) than the other services and may require extra cards in the Zephyr mainframe card cage.

The best way to calculate details regarding configuration options and dollar amounts is to talk directly to Zephyr: telephone (617) 898-3511. If there are any questions that fall outside Zephyr's domain, call UPO.

If you encounter problems in the use of the Zephyr system, please report these, in as much detail as possible, to Zephyr's user-support staff, at (617) 898-3511. In accordance with the UNIDATA contract, all such problems are recorded and tracked until resolved. Among the capabilities of the UNIDATA LDM system will be means for systematically measuring the quality of the data received via

(Continued)

Zephyr; this will provide quantifiable assessment of contractual compliance for use in reporting problems. Until this resource is available, accurate measurement of performance will be difficult; however, we recommend that problems be reported to Zephyr even if their determination is somewhat subjective. ■

Initialisms

GEMPAK	General Meteorological package (NASA)
GKS	The international standard Graphical Kernal System
GOES	Geostationary Operational Environmental Satellite (NOAA)
LAN	Local Area Network
LDM	Local Data Management
McBE	McIDAS Broadcast Evaluation Committee
McIDAS	Man-Computer Interactive Data Access System
NASA	National Aeronautics and Space Admnistration
NCAR	National Center for Atmospheric Research
NOAA	National Oceanic and Atmospheric Administration
NMC	National Meteorological Center
NWS	National Weather Service
UPO	Unidata Program Office