

ACRF Data Collection and Processing Infrastructure

December 2004

M.C. Macduff,
Pacific Northwest National Laboratory

R.C. Eagan
Argonne National Laboratory

Work supported by the U.S. Department of Energy,
Office of Science, Office of Biological and Environmental Research

Contents

1. Introduction	1
2. Global ACRF Architecture.....	1
3. Measurement Data Collection	5
4. Data Trafficking	9
5. Data Processing at the DMF.....	14
6. ACRF Data Archive	16
7. Future Challenges.....	18
8. Summary	19
9. Acknowledgements	20
10. References	20

Figures

1. ARM Climate Research Facility Measurement Locations.....	2
2. ARM Climate Research Facility Data Flow	4
3. A View of the Doorstep Data Handling Process.....	10
4. A view of the Doorstep Data Handling Process.....	12
5. Data request and retrieval processing from the Data Archive.	17

1. Introduction

Designated a national user facility, the U.S. Department of Energy's (DOE's) Atmospheric Radiation Measurement (ARM) Climate Research Facility (ACRF) provides a unique asset for the study of global climate change to the broader national and international research community. It has enormous potential to contribute to a wide range of interdisciplinary science in the areas of meteorology, atmospheric aerosols, hydrology, ecology, oceanography, satellite validation, and to provide potential monitoring sites where remote sensing and modeling related to homeland security can be validated. The primary goals for the ACRF are to (1) provide infrastructure to the scientific community for scientific research pertaining to global climate change and the goals of the ARM Program (Ackerman and Stokes 2003), (2) provide data and information to the scientific community for meeting those goals, and (3) provide education and outreach on the activities and scientific findings that result from ongoing research at the ACRF. The foundation of the ACRF infrastructure is based on the scientific infrastructure created for the ARM Program (DOE 1990).

In support of the ARM Program, the ACRF operates three instrumented sites and a mobile facility to provide relevant atmospheric measurements to the ARM Program and to the global scientific community. The goal of the ACRF infrastructure is to deliver these measurement data reliably, quickly, and in a useful format to the scientific community.

The basic focus of the infrastructure is to get the data generated by instruments in the field to a central distribution point. The remoteness of the sites and the diversity of the instruments add to the complexity of the solution. Network access to the sites was often limited and significantly impacted options for data flow and the architecture deployed at each location. Because of several iterations and significant work to establish Internet connections at each site, the ACRF has developed an efficient and integrated data flow. The network and computing infrastructure is able to centrally process the data from all sites on an hourly basis and make daily updates available for general users. This is accomplished through satellite networking, specialized data movement processes, and a tight configuration management process.

We present a description of the data flow from measurement to long-term archive. We also discuss data communications infrastructure. The data handling processes presented include collection, transfer, ingest, quality control, creation of Value-Added Products (VAP), and data archiving.

2. Global ACRF Architecture

The ACRF has established several measurement locations throughout the world, as illustrated in Figure 1. These sites represent three different climatic regimes: (1) the Southern Great Plains (SGP), (2) North Slope of Alaska (NSA), and (3) Tropical Western Pacific (TWP).

The SGP was the first site to come online in 1992. The SGP Site comprises 30 facilities (one central facility, three intermediate facilities, four boundary facilities, and 22 extended facilities) distributed over a 300-km by 300-km area encompassing north-central Oklahoma and south-central Kansas. The SGP site represents relatively homogeneous geography, ease of accessibility,

a wide variability of climate, cloud type, and surface flux properties, and large seasonal variations in temperature and specific humidity.



Figure 1. ARM Climate Research Facility Measurement Locations

The TWP site came online in 1996, and comprises three facilities at Manus Island, Papua New Guinea; the Republic of Nauru; and Darwin, Australia. The TWP measurement facilities are located in the Pacific “warm pool,” which is characterized by deep convective cloud systems that produce high-altitude cirrus clouds that spread out over much of the region. These cloud systems regulate the amount of solar energy reaching the surface of the earth and the amount of the earth’s heat energy that can escape into space.

The NSA site came online in 1997, and comprises two facilities at Barrow and Atkasuk, Alaska. The NSA site provides measurement information related to cloud and radiative processes at high latitudes.

The ACRF plans to deploy the ARM Mobile Facility (AMF) in the spring of 2005. The AMF will deploy to locations of climatic interest, especially where there is little or no information in support of the ARM mission. Typically, these deployments will last about 12 months.

The data collected from the ACRF sites are processed at the ACRF Data Management Facility (DMF) before being transferred to the ACRF Data Archive, where the data are made easily accessible to users. Users can request data collected from other sources. These data are collected and processed at the ACRF External Data Center (XDC).

All ACRF measurement locations are connected to the Internet. The SGP central facility, near Lamont, Oklahoma, has a T1 link to an Energy Sciences Network (ESNet) peering point at Oak Ridge National Laboratory (ORNL). Other measurement facilities in the SGP site's 300-km by 300-km site area use a continuous, low-speed (20-50 Kbps) satellite link or a landline modem dial-up to a local Internet Service Provider (ISP). On Manus Island and Nauru, the network connection is accomplished with satellite ground stations that support 256-Kbps outbound and 64-Kbps inbound channels. The Darwin facility has a 128-Kbps frame relay link with an Australian service provider. At Barrow, the ACRF shares a satellite-based T1 link partially funded by the National Science Foundation (NSF). The Atqasuk facility uses a symmetrical 64-Kbps satellite link through a commercial ISP. The variety in types of infrastructure is a result of identifying the most appropriate, available technology that will support the data transfer requirements of the facility in the most cost-effective manner. In practice, the network infrastructure has, at some locations, changed over time to meet changing needs and to take advantage of current technologies. In the past, limited or no network access dictated that data be transmitted on removable media. This had a significant impact on the efficiencies of the entire data flow process. Simple activities such as upgrading software or debugging problems were much more time consuming.

The fact that all ACRF sites are Internet-connected has enabled the ACRF to develop the data flow architecture illustrated in Figure 2. The ACRF data flow process has the goals of acquiring the measurement data at the ACRF sites, processing the measurements into a form that is usable by the scientific and educational communities, performing quality analyses of the data set, and archiving the data in a manner that is easily accessible by globally dispersed users.

The data flow process includes the following functionality that is performed at various physical locations:

- collection of measurement data at the ACRF measurement facilities
- transfer of measurement data to the ACRF Data Management Facility (DMF)
- “ingest” of data into a common, standardized format
- perform quality analysis of the data
- perform supplemental processing, which results in VAPs
- reprocess the measurements as necessary to correct errors or otherwise ensure a complete, representative data set
- acquire and process non-ACRF data products of interest to the ARM scientific community
- transfer raw and processed measurements to the Archive
- ensure the long-term availability of the data in the Archive
- make the measurements publicly available to a globally distributed user community in a process that is as transparent as possible.

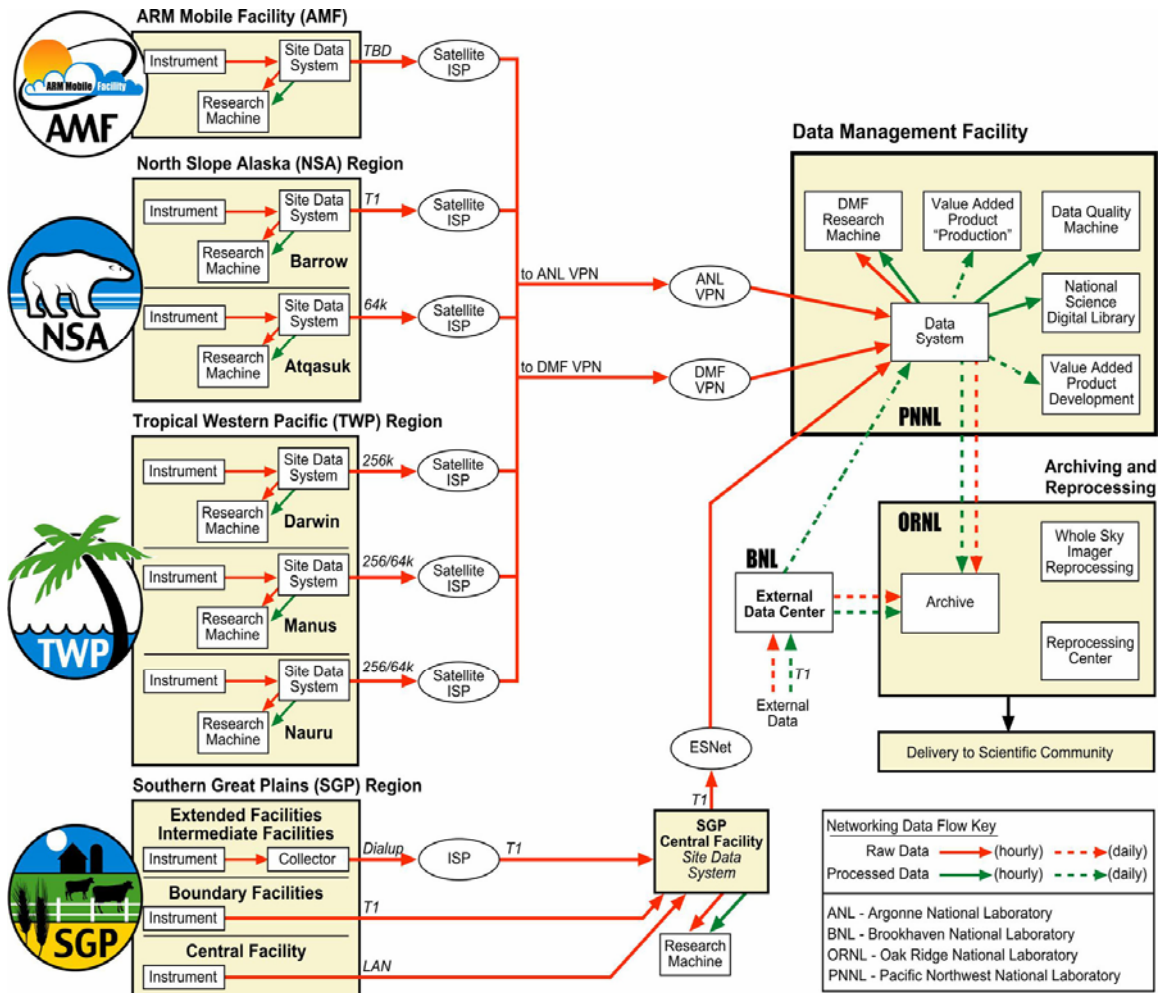


Figure 2. ARM Climate Research Facility Data Flow

Figure 2 illustrates that these ACRF functions are performed at the following various locations:

- The DMF is located at Pacific Northwest National Laboratory (PNNL).
- Data Quality (DQ) processes are located at the DMF and operated by staff at the University of Oklahoma.
- External Data Center (XDC) is located at Brookhaven National laboratory (BNL).
- Data Archive is located at Oak Ridge National Laboratory (ORNL).
- Reprocessing Center is located at the ACRF Data Archive.

Because the sites are Internet-connect, access controls must be in place to ensure the facilities operate continuously and reliably. The ACRF data flow architecture is logically implemented by routing all Internet traffic to the measurement facilities through the ACRF Virtual Private Network (VPN) Server Network (VSN) located at Argonne National Laboratory. The VSN implements VPNs with each of the measurement facilities. This architecture has several interesting features:

- Internet routing effectively remains unchanged when ISPs for local measurement facilities change due to budget or technological reasons for example.

- Firewalls are implemented in a common, central, and easily accessible (by authorized administrators) location.
- Undesirable traffic can be filtered before it traverses (often expensive) satellite infrastructure.
- Traffic monitoring and analysis can be performed at a common, central location.

For reasons of both performance and redundancy, the DMF also implements a set of VPNs with each measurement location. These VPNs are the primary paths of data flow between measurement facilities and the DMF.

3. Measurement Data Collection

The ACRF measurement facilities include a wide range of instrumentation. The instrumentation ranges in complexity from relatively simple RS-232-based data loggers to network-based cloud radars, lidars, and sky imagers. The instrument systems may be commercially available or specialized systems developed by universities, laboratories, or government agencies.

Many instruments include a computer-based controller that may be embedded, a laptop, or a desk-side system. The instrument computer uses an operating system often dictated by the availability of drivers for the instrument's digital signal processors or other data collection hardware. Consequently, instrument operating systems may be Microsoft (Windows 98 through Windows XP), OS2, Linux, MacOS, or Solaris based.

The ACRF data system components at each facility include:

- collector system
- dual (redundant) data processing systems
- file server
- research computer.

Currently, the collector, file server, and research systems are Linux x86-based. The processing systems are Solaris-based SPARC systems. The collector and file server systems can be easily configured to perform the tasks of the other (or both systems).

The collector system is responsible for collecting measurement data from the local instruments.

The data processing system performs the local facility data ingest function and hosts a website, DSView, which provides a graphical view of the status of each instrument data stream and of various facility operational processes. The ingest process operates on each instrument data stream to:

- read the raw data
- validate that data values are within a predefined, reasonable, range of values, if so configured

- convert the raw data to netCDF (network Common Data Form)
<http://my.unidata.ucar.edu/content/software/netcdf/index.html>
- update a local database that archives various metadata about each instrument data stream.

While the official raw data ingest is performed at the DMF, a local ingest provides data stream status information to local operators and visiting scientists. The ingest and DSView processes are discussed in more detail in Section 5, Data Processing at the DMF.

The file server is a RAID-V disk array that archives several weeks of locally collected measurement data, local databases, and the ACRF data system software.

The research system allows visitors, instrument developers, and other selected users local, read-only access to local data streams.

A critical characteristic of the measurements collected at each of the facilities is the accurate time stamp of the data streams. Each facility includes a network time protocol (NTP) server that uses the Global Positioning System (GPS) as a time reference. It is critical to ensure, through automated tools, that all data collection and processing systems are properly synchronized with an accurate time reference. Each critical system at the site, with an operating system that can support it, runs a version of the network time protocol daemon (ntpd: <http://ntp.org>). Windows 98 systems use an open source program, "AboutTime," which supports the NTP service. A regularly scheduled process in the ACRF data system automatically polls all time-critical systems at a facility and, using features of the NTP service, determines the time offset each has from the local reference source. If any system is determined to be not running the NTP service or is in error by more than a specified amount, local operators and systems administrators are informed of the problem, so it can be resolved quickly.

Collection of measurement data from the instruments is performed in various ways depending largely upon what kinds of data-transfer processes the instrument and its associated processing system can support. Collections fall into the following categories:

- serial (RS-232, RS-422) device collections
- FTP transfer
- network file system transfer
- removable media transfer.

3.1 Serial Device Collections

The collector system runs an instrument collection module that periodically polls the configured RS-232 instruments through a network-based serial hub. For each configured serial instrument, the collection process:

- acquires the current reference time from the local time server
- acquires the instrument time and, if configured to do so at this polling time, resets the instrument time if the difference is outside a configured tolerance
- logs instrument time adjustments (both magnitude and time of change)
- acquires all data not yet collected from the instrument.

Data-collection events are limited to a time limit or to a maximum number of records. This ensures the collection process does not consume too many instrument or data logger resources. Some systems have limited resources, and continuous polling can have an adverse effect on measurement data acquisition. In addition, the collection process should be begin and end within the same clock hour, due to constraints in the ACRF processing system.

The ACRF developed POSIX (<http://www.pasc.org>) compliant data collection software to support various data logging platforms. However, the current serial collection software includes only that supporting the Vaisala DigiCora MW11 and MW15 models.

The ACRF adapted the xtty program by Jim Schlemmer (<http://hog.asrc.cestm.albany.edu/~rsr/bits/softdocs/xtty/xtty.html>), which supports the multi-filter Rotating Shadowband Radiometer (RSR) for ACRF use.

3.2 Instrument Time Correction Considerations

Serial-based instruments, by definition, do not have network access and do not support the NTP service. In the ACRF, the collection software ensures that the logger clock is appropriately synchronized with real time. Typically, the logger clock is corrected if the apparent time drift is in error by more than 2 seconds or two times the resolution of the logger clock. It is important to recognize that data logger logic often will cause one or more data records to be skipped when the clock is adjusted. In most cases, the ACRF finds that data logger clocks are sufficiently stable, and it is adequate to check the clock accuracy only once per day to avoid an excessive number of missed data records. Typically, time will be checked and reset, if necessary, near local midnight, which is an appropriate time for instruments measuring characteristics of solar radiation.

3.3 FTP Collections

Most instruments operating in the ACRF are implemented under a computer operating system that supports the FTP network service. The ACRF has been able to implement a common protocol for data collections from virtually all such instruments. The protocol can be summarized as follows:

- The instrument computer generates hourly files of data and metadata (i.e., logs, hardware, and metadata). Instrument vendor software does not always support this notion, but generally software can be developed for the instrument computer to support this requirement.
- The FTP server for the instrument PCs has a data system username and a password.
- Files delivered from the instrument PC to the ACRF data system will be placed in a directory readable and writable by the data system user.
- The data system, through the FTP process, acquires only files with specified filename extensions from this directory. This process allows the instrument computer to place “growing” files in the directory and rename them to an agreed convention when the file is complete.
- The data system ensures the files acquired from the instrument PC are the same size as the original files.

- The data system removes files it has collected from the directory of the instrument PC data system collection. This process prevents multiple collections of the same data. The instrument may maintain an historical archive of its data in a different directory.
- The data system, in the process of collection, may need to translate the source filename to one with an ACRF conventional naming structure to ensure uniqueness.
- The data system logs the names and sizes of files collected and any errors encountered.
- The data system updates the local DSView database to reflect collection status.

The FTP process is also used for transferring configuration or other information from the data system to the instrument. The protocol is similar to that described for data collection.

3.4 Network File System Collections

Certain instrument computer operating systems, while supporting network connectivity, do not have sufficient operating system resources to support the FTP or similar protocols. In such cases, the data system file server exports a data collection directory to the instrument computer using the NFS or SMB protocol. The file transfer protocol is similar to that for an FTP-based transfer:

- The instrument computer generates hourly files of data and metadata (i.e., logs, hardware, and metadata). Instrument vendor software does not always support this notion, but software can be developed for the instrument computer to support this requirement.
- The instrument PC has a data system username on the NFS server, and the user shall have read and write permission on the exported data-transfer directory.
- Files delivered by the instrument PC to the ACRF data system will be placed in the directory exported by the NFS server.
- The data system acquires only files with specified filename extensions from this directory. This process allows the instrument computer to place “growing” files in the directory and rename them to an agreed convention when the file is complete.
- The data system ensures that the files acquired from the instrument PC are the same size as the original files.
- The data system removes files it has collected from the directory of the instrument PC data system collection. This process prevents multiple collections of the same data. The instrument may maintain an historical archive of its data in a different directory.
- The data system, in the process of collection, may need to translate the source filename to one with an ACRF conventional naming structure to ensure uniqueness.
- The data system logs the names and sizes of files collected and any errors encountered.
- The data system updates the local DSView database to reflect collection status.

Although few instruments require it, this transfer mechanism is used only when no practical alternatives exist. Under this mechanism, the risk of causing the instrument computer to “hang” and stop collecting data exists if there is a network problem or if, for some other reason, the instrument PC fails to communicate with the file server.

The network file system process is also used for transferring configuration or other information from the data system to the instrument. The protocol is similar to that described for data collection.

3.5 Removable Media

A few ACRF instruments generate data sets too large to transfer over network connections through the normal data-flow process. The millimeter cloud radar spectral data and the whole sky imager data sets are representative examples. Such data are typically shipped from the measurement facility to the DMF on removable media. When data are transferred using removable media, the procedure involves creating two identical removable-media devices containing identical files. One device is shipped to the DMF and the other remains onsite. The on-site copy is available for data retrieval, should there be problems in the shipping process.

4. Data Trafficking

The collection process, logically, receives the data from the various instruments. Its focus is interfacing with the unique characteristics of each instrument. Once the data are collected, it is ready for a greater level of generic automation in moving the data through the system. To move the data through the system, three processes are used: (1) doorstep, (2) Site Transfer Suite (STS), and (3) transfer suite. All three processes focus on moving files around and do not know or care about the uniqueness of the data or the instrument from which it came. These processes each have a particular focus and allow customization to discriminate different data sets, but in a more general sense (i.e., high vs. low priority) than the instrument-specific collections.

Generally, moving data through the systems in an orderly way involves using different directories at each step. In each directory, there is the possibility that a process is only partially finished writing a file before the next process tries to read it. To avoid this problem, a locking mechanism is necessary, such as writing a file with a special extension (e.g., .lck) and, when the file is complete, simply renaming it to the intended filename. The reading process is then programmed to ignore files ending with the lock extension. All data trafficking processes use this technique to avoid contention.

4.1 Doorstep

The doorstep process manages the incoming area of the data-processing system (see Figure 3). Many types of files may need to enter the data-processing system, including logs, status files, configuration files, and data. The doorstep process can handle each of these files uniquely and, based on the system configuration, regulate which files may enter. On first inspection, much of the doorstep function could be done by extending the collection process. However, separating this logic provides several benefits. The doorstep process is an intentional bottleneck that allows tracking of every file going through the system. It can perform specialized parsing or renaming of files for different types of inputs (logs versus configuration files). This single process runs as a daemon, ensuring optimal flow of data and immediate detection of problems. It also updates the DSView database with collection-status information, providing a central point of coordination for this important information.

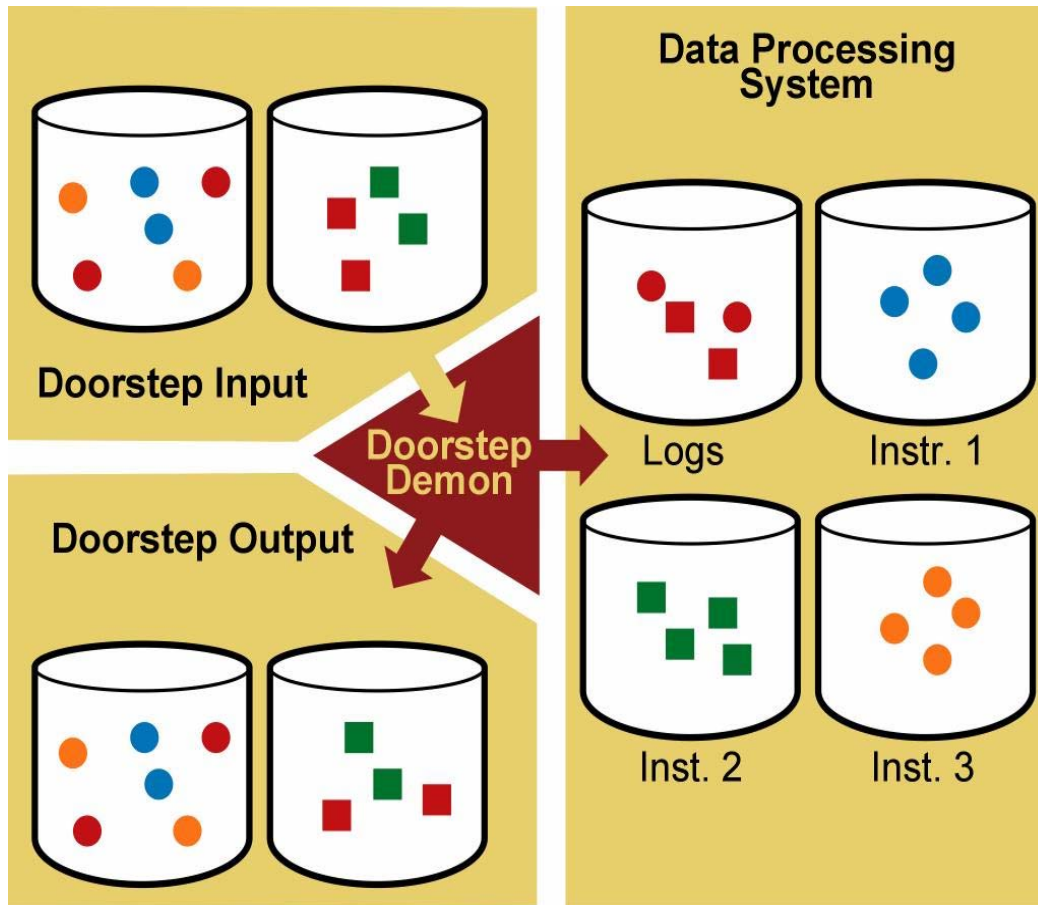


Figure 3. A View of the Doorstep Data Handling Process

Because all data flows through the doorstep fresh from the instrument and before any processing, it is a convenient point to split the data down multiple paths. This function is essential to the overall data flow. The data are needed on each site for local processing and inspection by operations staff and visiting scientists. It is even more critical that all of the same data are transmitted to the DMF and ultimately to the ACRF Data Archive. The doorstep is able to split reliably the data through the following process:

- copy file to a lock filename at multiple destinations
- validate success of all copies
- if any of them failed, remove all new files and quit
- rename files to correct names
- if any renames fail, remove all new files and quit
- remove original file.

The process is not bulletproof, since it is possible that the data file will be unlocked (renamed) in a new destination, but still removed because another unlock (rename) failed. The impact of this worst case would be that the same data are put into a destination multiple times. To not impact the data flow that ends up in the ACRF Data Archive, the critical path is always listed last to avoid duplicating data to it.

4.2 Site Transfer Suite

While the doorstep process is responsible for directing traffic, the Site Transfer Suite (STS) is responsible for inter-site deliveries. It makes reliable, long-distance file moves through the following algorithm, which are illustrated in Figure 4:

On Host A:

- create an MD5 checksum companion file for a data file
- copy both the file and checksum file to Host B with a lock extension
- rename files on Host B with original names
- log the filename and checksum to a PACKLIST file
- move the data file to a sent files directory and delete checksum file
- copy PACKLIST file to Host B after all files are sent.

On Host B:

- verify file matches checksum file
- safely move file to a received files directory and delete checksum
- log receipt of file with checksum to incoming logs
- validate each line in PACKLIST file with incoming logs
- create MANIFEST file from PACKLIST indicating OK and MISSING files
- copy MANIFEST file to HOST A.

On Host A:

- process MANIFEST file
- delete from sent files directory, if the file is OK
- move it back into the outgoing directory, if the file is MISSING.

While this algorithm exhibits some paranoia through redundant validations of correct transmission, it still has potential holes. (Of course, any solution will have some probability of failure.) The reality is that bad things rarely happen to good files and this algorithm catches the few bad cases. The mechanism used to copy files is anonymous FTP, although alternative mechanisms are being considered.

The STS typically works on the output from the doorstep process. The doorstep faithfully writes the files in the order it receives them, big or small. However, across low-bandwidth satellite links, a poorly timed backlog of “big” files could easily clog the pipe. Maintaining the in-order delivery of data without starving the small files creates a dilemma. The doorstep design allows for a simple solution by segregating files into many directories. The in-order delivery requirement is limited to a particular directory. Therefore, the STS can be configured to prioritize directories.

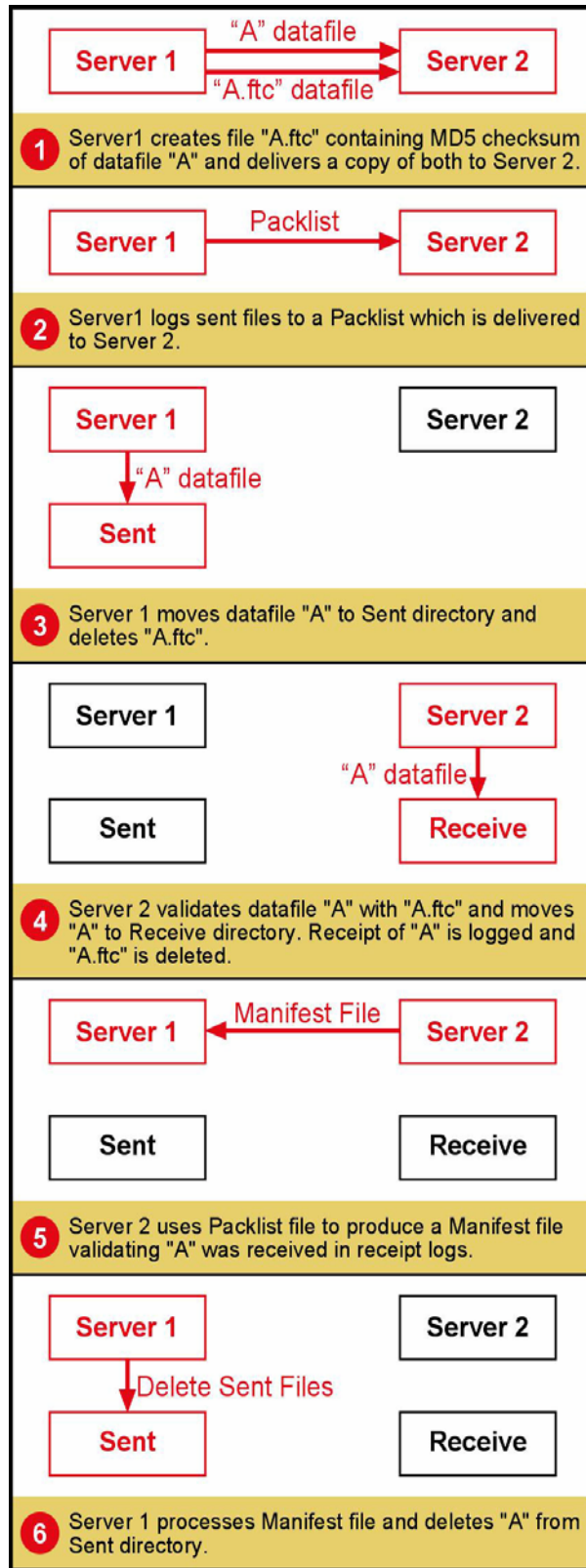


Figure 4. A view of the Doorstep Data Handling Process

Thus, it can transmit the small-file directories before working on the larger ones. This problem is solved by putting limits on the time or number of files for each session. Thus, the STS can ensure the needed in-order delivery AND avoid data flow starvation.

The requirement for in-order data delivery causes more problems than just the potential data starvation described above. Understanding that the real in-order requirement is limited to similar types of files in different directories provides a solution to the initial problem. However, the reality is that a few directories contain many of the files (which are mostly very small). The process of serially copying files, using FTP, across a long-latency satellite link creates an obvious problem. With a latency of about 1 second, this limits the maximum rate of file transfer to about one file per second. Adjusting the TCP window parameters improves the performance of larger files, but there is still a throttle on the maximum rate for small files. This results in a complete waste of bandwidth and starvations of the big files falling farther and farther behind on data flow.

Overcoming latency problems may include using a streaming protocol, adjusting TCP window parameters, bundling files, and using multiple streams. STS does the latter by creating multiple threads that each control an FTP session. The list of files to send is centrally maintained and passed out to each thread one at a time. This parallelization allows for many files to be in flight simultaneously, using up more of the bandwidth and removing the throttled limit. Using this, in combination with adjusting the TCP window parameters, results in a good overall performance solution.

Moving data in parallel violates the requirement of in-order data delivery. To resolve this required further insight; the in-order requirement only applies to files from a specific instrument in a directory. Each file is tagged with a unique instrument string at the beginning of its name. Using this tag, STS can ensure that it only sends one file from an instrument at a time, which means that the data for an instrument will be sent serially but in parallel with other instruments. Note that some instruments do not have an in-order requirement. In these cases, the parallelism can be applied to one instrument.

The STS, through the use of a slightly paranoid algorithm, has reliably delivered data throughout the ACRF for many years. The addition of prioritization and instrument-aware parallelism has allowed for efficient use of bandwidth even for the many small files created in an hourly data flow process.

4.3 Transfer Suite

The data processing system does its work to produce and manipulate data on periods from an hour to several months. Unfortunately, no mechanism exists that indicates a file has been created or is complete. In the case of the DMF, the files produced by the processing system need to be delivered to the ACRF Data Archive. The STS process is used to do the actual delivery but needs to be given the files to send. To get the files, a process called transfer is used. This is a customized “find” command. Based on a configuration file, it searches through directories for specific patterns and modification times and copies the files to a target directory. Because many files are updated hourly, their simple existence does not presume completeness. Therefore, we assume that if the file has not been modified for one day, then it is probably complete and safe to

send to the Data Archive. This scenario may present problems with either missed files or sending multiple copies of the same file. To mitigate some of these issues, a process parses the file list for unusual data patterns and notifies operations of unexpected files.

While the techniques used in the final stages to move files to the Data Archive may seem rudimentary, it is partially because the problem is more complex than has been indicated. To state it simply, how can one know that a file is ready for archival? When an instrument goes off-line, the processing system has a file with only part of the day's data in it. It is possible that there is just a delay in shipping the data or some other outage preventing receipt of the files. When the instrument is reactivated and the data begins flowing, it may include all of the data during the outage, or it may only include the current files. It has also been the case that a removable disk was used to ship the data. This data almost certainly ends in a partial day. The only option is to wait for the next disk, which, given shipping times from remote facilities, could be a month or more. Some instruments have produced gaps in data that are retrievable by manual intervention with the instrument. Finding and filling gaps is often a multi-week process. Other anomalies can cause difficulties in identifying whether a file is complete. The current solution relies on operators to identify problems and intercept them appropriately. In the event that a partial file is archived and the file is later completed, the latest file is simply re-archived. This puts a burden on the Data Archive to manage multiple copies of the same filename. This is handled through versioning of files. The automated identification of completed files is an area for improvement in the ACRF data flow. However, it is difficult to imagine rules that would completely prevent archiving multiple copies.

5. Data Processing at the DMF

Every hour, the DMF receives raw data files from each measurement facility. These files are delivered fresh to the DMF complete with their original names in their original format. Ingesting the data into a common format and with consistent naming is accomplished during initial processing, making the data accessible for further processing and evaluating. While much of the work is automated, this is the point at which operators and scientists begin validating the quality of the data. Having the data processed in near real time enables the rapid resolution of problems.

5.1 File Naming

The raw filenames produced by the instrument have a variety of cryptic naming conventions. The ARM Program established a file naming convention that provides consistency and enables automation. The structure is centered on the idea of a data stream. A data stream is a time series of files with the same structure. The data, over time, are segmented into files to facilitate continuous storage and incremental distribution of the data. Each file is associated with a data stream name to which the date and time are appended. The data stream name includes shorthand information about the data's location and type. The naming convention can produce files that are cumbersome to the uninitiated, but provides tremendous value.

5.2 NetCDF and Zebra

The data formats of each instrument have at least as much variety as the file names. To deal with the variety, the ARM Program established that all data would be stored in a common netCDF file format (<http://my.unidata.ucar.edu/content/software/netcdf/index.html>). This provides a format that includes metadata and standard tools for accessing the data. In addition to the netCDF libraries from the University Corporation for Atmospheric Research (UCAR), the data system also uses the Zebra software developed at UCAR. Although UCAR no longer supports it, the Zebra software is central to the processing of ARM data. It provides a data management interface between the processing software and the data files. Among other features, it ensures files are split on a day boundary.

5.3 Ingest, Quality Assurance, and Raw Data

An hourly ingest process picks up the hourly raw data files. Each raw data stream has a separate process run against it. The function of the ingest varies with each raw data stream, but the data are typically parsed and structured into a netCDF format, converted to engineering units, and have calibration factors applied. Most of the raw data streams also have basic quality analysis done in the form of min/max/delta checks, which are also included in the output file. The output of each ingest is a new data stream. Often, the names of the raw data stream and netCDF data stream will be similar, for obvious reasons. The ingest process is also responsible for correctly naming the raw file. This cannot be done earlier because some information necessary for naming is often included inside the data of the raw file. While there is typically one netCDF file per day, the raw data files can accumulate quickly. The raw files are therefore bundled into tar files at the end of each day for ease of management and long-term storage. Meanwhile, as soon as the netCDF file is written, it is available to many other processes that are interested in understanding the data.

5.4 Operational Review

The most basic review of the data is accomplished through the DSView Web interface. This provides operators at the DMF and in the field with a view of the collection and its processing status and data completeness. This helps ensure data flow continues, as well as initiate same-day troubleshooting of significant data problems. Processing errors provide an early warning of instrument problems.

5.5 Value Added Product (VAP)

Once the netCDF data are complete for a period, it may be used by a VAP. VAPs often incorporate multiple data streams and complex algorithms to produce an entirely new data stream. Other VAPs can also use these data. Because of the variety of data input requirements, it has been difficult to automate the running of VAPs. Each VAP has its own periodicity, as well as nuanced input requirements. A VAP may require three data streams to run, but has three other optional input data streams. Deciding when a VAP is ready to run is currently a human activity. Some tools have been and are being developed to assist the operators, but this is a non-trivial

problem. The data streams produced by VAPs are often the actual data many scientists want to see. Therefore, it is a priority of the DMF to keep VAP data flowing.

5.6 Data Quality Review

Many ingests automatically include quality analyses that are put into the netCDF files. Other data streams require more complex algorithms to discover quality problems. The ARM Data Quality Office (DQO) collates the quality information available in near real time and provides an interface, Data Quality Health and Status (DQHandS), which enables assessments of data quality. The ACRF scientists use this interface to evaluate potential problems with data quality as early as possible. This interface allows for a rapid response to potential problems, as well as documentation of known problems.

5.7 Clean Script

The DMF needs to keep only a few months of data for VAP processing. The data are sent to the ACRF Data Archive a day after they are produced. Therefore, it is reasonable to delete the data from the DMF after a few months. However, it is possible that the data to be deleted never made it successfully into the Data Archive. (See discussion above in Transfer Suite.) It would be extremely unfortunate to process a piece of data through all of the hoops but then delete it permanently.

Instead of deleting a file, a script queries the Data Archive's file database. The MD5 checksum of the file is compared with the Data Archive's records. If there is a problem, the file can be resent. This final check ensures that collected data are archived.

6. ACRF Data Archive

Raw and ingested data are transferred from the XDC and DMF to a buffering system at the ACRF Data Archive using the site_transfer process. The data are archived by saving two redundant copies of files in a robotic tape archiving library. The High-Performance Storage System (HPSS) manages the tape libraries, which are part of a Mass Storage System (<http://www.hpss-collaboration.org/prodinfo.html>). A third copy of the data files is kept in a location physically separate from the tape robot library. This backup copy is managed under a simplified inventory system to minimize replicating inventory problems occasionally encountered in HPSS. For security reasons, no mechanism for interactive processes exists to interface with the tape archive system. The Archive staff update the Data Archive using script-controlled or batch processes.

The user interface to archive data is available at <http://www.archive.arm.gov/>. ACRF data are freely available to any requestor through the user interface. The process is summarized in Figure 5.

The ACRF Data Archive currently supports about 1,800 registered users from 7 U.S. agencies, 110 universities, and 40 countries. The Archive services about 350–500 requests per month consisting of about 100 K–200 K files or 600 GB–1,200 GB per month.

Data request and retrieval processing from Archive

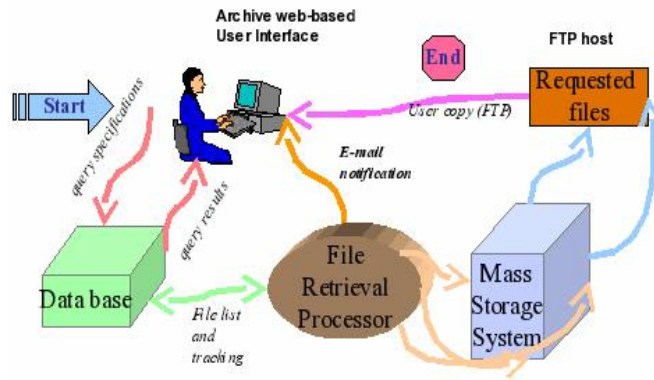


Figure 5. Data request and retrieval processing from the Data Archive.

The Data Archive currently supports about 37 TB of data in approximately 6.5 M files. About 2.4 M files are user accessible. The Data Archive is currently growing at an approximate rate of 45 K–80 K files (500 GB–800 GB) per month.

The ACRF Data Archive developed a Web-based interface to enable both new and experienced users to navigate the available data and select sets of data files to download. The details of the interface are beyond the scope of this paper; however, we present a subset of the basic functionality.

The ARM data browser provides a query interface based upon specific selection criteria, including measurement location, data stream name, or data type and time period. This interface provides user access to data quality information and graphical representations of the selected data.

In general, acquiring data involves the following steps, as illustrated in Figure 5:

- create a new account (new users only)
- login to interface
- formulate data selection criteria
- view results
- optionally review data quality reports for selected data
- optionally review thumbnails of data displayed in graphical format
- accept or reject selection
- logoff or perform an additional request
- wait for email notification
- copy files using armguest account for anonymous FTP on ftp.archive.arm.gov
- search <http://www.arm.gov> website for ancillary information.

7. Future Challenges

7.1 Importance of Ensuring Raw File Integrity

The most important component of any of the ACRF data streams is the raw data component. From this stream, the remaining data streams associated with an instrument can be reproduced. It is critical that the integrity of this stream be ensured from the instrument until it is stored in the Data Archive. In the next release of the ACRF data system, additional safeguards will be implemented. This could be in the form of a checksum companion file or other reliable mechanism.

7.2 Timing Issues of Computers with Unstable Clocks

While the main facilities at the TWP, NSA, and SGP ACRF sites have stable GPS-based reference clocks, the SGP site includes over 20 extended facilities (EFs), which do not have access to the time reference because they do not have continuous network communications with the time reference. To provide a local time reference, the local EF Linux laptop-based collector system is equipped with a GPS receiver. The GPS receiver is configured under ntpd as a stratum 1 server. Because there is minimal thermal shielding in the EF communications enclosure, problems arise because the laptop is exposed to wide variations in operating temperature over a diurnal cycle. For the equipment in use at the EFs, the frequency of the laptop's system clock is highly variable with temperature, to the extent that ntp cannot maintain a stable time reference. The ACRF Engineering staff are evaluating replacement systems that have clocks with temperature coefficients more appropriate for the application.

Others who fielded data collection systems in similar environments are encouraged to investigate the temperature stability of their system clocks. It is important to note that we have observed this problem only with certain laptop computers, not with any commonly used data logging systems.

7.3 Data Processing Automation

To automate VAP processing, it is necessary to have a tool that understands the varied data inputs and can make suitable recommendations. This tool would need to understand even the command-line switches of some VAPs that have multiple operational modes. While inspecting data files is an option, it would be preferable to know the state of each file regarding its completeness and availability as an input. This capability would also provide options for faster delivery to the Data Archive.

7.4 Inter-site Data Delivery

The current STS software uses anonymous FTP as a transport mechanism. Alternative mechanisms include SFTP, rsync, and LDM. Using standard software is preferred from maintenance and security perspectives. Comparing the alternatives in light of the ACRF requirements should be considered.

7.5 Tracking the Inventory of Processed Data through Various Steps

At each stage in processing there needs to be a better understanding of what needs to be done next for each piece of data. While it is currently possible to recover (or remove) incompleteness or replication of processing inconsistencies, much time is spent managing and fixing these problems.

7.6 Management of Data Stream Changes with Time

The program currently manages changes in data streams and distributes information about those changes to data users. Change is a fundamental characteristic of a research program, and there are procedures in place to manage change. However, it is necessary to research methods to make this task easier. Specifically, tracking and effectively communicating the state of a data stream throughout its history is not a clean process.

8. Summary

The ACRF infrastructure collects data hourly from its measurement sites and delivers it to the DMF for processing within 20 minutes. The underlying network infrastructure is critical to this ability. The architecture of the entire data flow provides reliability and performance through a division of labor from data collection to the doorstep, to STS, and back to the doorstep at the DMF. The ingests can process data hourly, making them available for many uses that lead to overall improved data quality. The ACRF Data Archive provides this high-quality data to a global audience of data users.

The current ACRF infrastructure represents a significant milestone in the history of the program. With all data centrally processed on an hourly basis, operational and scientific users gain near real-time feedback. This, in turn, reduces downtime and instrument outages. Centralizing the monitoring functions of the ACRF site scientist and instrument mentors provides for efficiencies in tools and labor. The DMF operations can centrally control data flow and processing, reducing incidents of duplicate data being sent to the ACRF Data Archive. The centralized processing at the DMF provides ease of managing processing software updates and coordination with engineering staff. The real benefits of the streamlined architecture are yet to be realized. Most significantly, the structured approach to data flow makes it possible to implement changes simply. This positions the ACRF infrastructure for the challenges ahead.

As a national user facility, the ACRF infrastructure will need to continue to improve. Addressing the challenges addressed above (Section 7) is part of this. There are also the issues that come with the territory of a large distributed computer installation. New technology often helps with efficiency, but requires careful integration to be effective. Cyber security is of top concern to any Internet-based operational enterprise. This often implies limiting network access, but the needs of a national user facility imply a certain amount of openness and accessibility. The resolution will likely be found in new technologies and innovative solutions. As always, changes that can reduce engineering or operational costs translate into fuller services for ACRF users. Along these lines, improved management, maintenance, and monitoring of systems will provide benefits. Additionally, continued improvements in software robustness through software

engineering design, review, and testing will have long-term dividends. As a user facility, there is a need for more user services involving access to instruments, data, and especially useful metadata. To assist forecasting worldwide and specific field campaigns, the ACRF needs to improve the timeliness of data available for use. While the ACRF infrastructure has developed an effective data-flow architecture, embracing the challenges of a national user facility provides for an exciting future.

9. Acknowledgements

The ACRF infrastructure in its present form is the result of the hard work from a team of worldwide staff. The authors participated with this team, but much of the credit belongs to many unnamed contributors. Thanks also go to the following for their review and comments in preparation of this paper: Raymond McCord, Randy Pepler, Doug Sisterson, Jimmy Voyles, and Craig Swietlik.

The Climate Change Research Division, U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research, as part of the Atmospheric Radiation Measurement Program, supported this work. Battelle Northwest operates Pacific Northwest National Laboratory for the U.S. Department of Energy. The University of Chicago operates Argonne National Laboratory for the U.S. Department of Energy.

10. References

Ackerman, T.P., and G.M. Stokes, 2003: The Radiation Measurement Program. *Physics Today*, **56**, 3842.

U.S. Department of Energy, 1990: Atmospheric Radiation Measurement Program Plan. U.S. Department of Energy, Office of Energy Research, Office of Health and Environmental Research, Atmospheric and Climate Research Division, Washington, D.C., Report DOE/ER-0441. 116pps.