The Abstract Data Distribution Environment (ADDE) – A technical overview

Don Murray
Unidata Program Center
May 13, 2004

Outline

- What is the ADDE?
- Who uses ADDE?
- Technical details
  - Datasets and supported data types
  - Protocol details
  - Overview of ADDE requests and returned data
- Java and ADDE
- ADDE use in the IDV
- Summary analysis

What is the ADDE?

- Client/server data access model developed for McIDAS, but not limited to serving McIDAS data
- In use for nearly 10 years
- 4 primary types of data objects
  - GRID, IMAGE, POINT, TEXT
- Primary servers handle McIDAS format files (AREA, GRID, MD, LW)
- Secondary servers read non-McIDAS formats (GINI, NEXRAD Level III, MODIS, netCDF,…)
- Allows browsing and subsetting of datasets

Who uses ADDE?

- Unidata community
  - Cooperating set of servers providing data to McIDAS and IDV users
  - Meteoforum
- McIDAS community
  - International government agencies (e.g., Spain, ESA, Australia)
  - US agencies (e.g., NESDIS, NTSB)
  - Satellite researchers

ADDE Dataset definition

- A dataset is collection of one or more files with a common format. Each dataset has a name that consists of a *group* and a *descriptor* (usually written as *group/descriptor*).
- Each dataset is in a *group:*
  - A group can have one or more data types in it. Examples:
    - **RTIMAGES** – real time image data only
    - **BLIZZARD** – data (images, grids, point data) for a case study or tutorial
- Each group has *descriptors* which define a set of data of the same type (i.e., image, grid, point or text).
  - Examples:
    - **RTIMAGES/GE-IR** – set of GOES-East IR images in the RTIMAGES group
    - **RTIMAGES/MOLL-IR** – Mollweide Composite IR images in RTIMAGES
    - **BLIZZARD/GRIDS** – set of NGM grids for the BLIZZARD data group
- Groups can be interrogated for a list of descriptors and data types it contains.

Currently Supported Data Formats

- Image
  - McIDAS AREA, GINI, NEXRAD Level III radar, netCDF (output only), MODIS, AIRS, GVAR, POES, Level 1B, Meteosat (including MSG), NOWrad®, GMS and FY-n
- Grid
  - McIDAS GRID, netCDF, GRIB (in development)
- Point
  - McIDAS MD, netCDF, HDF4
- Text
  - Plain text, McIDAS-XCD observation text (e.g., raw METAR, RAOB) and bulletins (e.g., watches and warnings)

Protocol Details

- Client connects on a particular port (500, 503 or 112). With next version, only Port 112 will be used. Supports compression using compress or gzip.
- Handshake from client involves sending:
  - Version info (make sure server supports this)
    - ADDE version (1 for now)
  - Pre-request information (for validation)
    - Server IP and compression type
    - Request Type (AGET, ADIR, etc)
  - Request Block
    - Server IP and compression type (again)
    - Client address
    - User, project, password (for authentication)
    - Request Type (again)
    - Actual request for data
- Server sends back data or error message

Validation/Security

- ADDE supports 3 types of validation for a request to a server:
  - IP filtering
  - Username
  - Project number
- Security through obscurity – no mechanism for querying an ADDE server to find out what datasets are available.
- TCP wrappers can be used to limit access to ports
- Subservers can implement their own forms of validation/security (e.g., NEXRAD Level III server)

Request types

Request Type	Java equivalent	Data Type	Description
ADIR	imagedata	image	image header information
AGET	imagedir	image	image header, navigation, calibration and data; data is returned line by line; comments
GDIR	griddirectory	grid	grid header information
GGET	griddata	grid	grid header and entire grid
LWPR	datasetinfo	info	dataset information
MDFH	*not implemented*	point	point file header information
MDHD	*not implemented*	point	point header information
MDKS	pointdata	point	point header and data
OBTG	obtext	text	observational weather text
TXTG	text	text	ASCII text file
WTXG	wxtext	text	textual weather information

Anatomy of a Request

An ADDE request is a text string containing positional parameters and some key=value pairs (just like a McIDAS command).

A sample request for the latest Mollweide IR image:

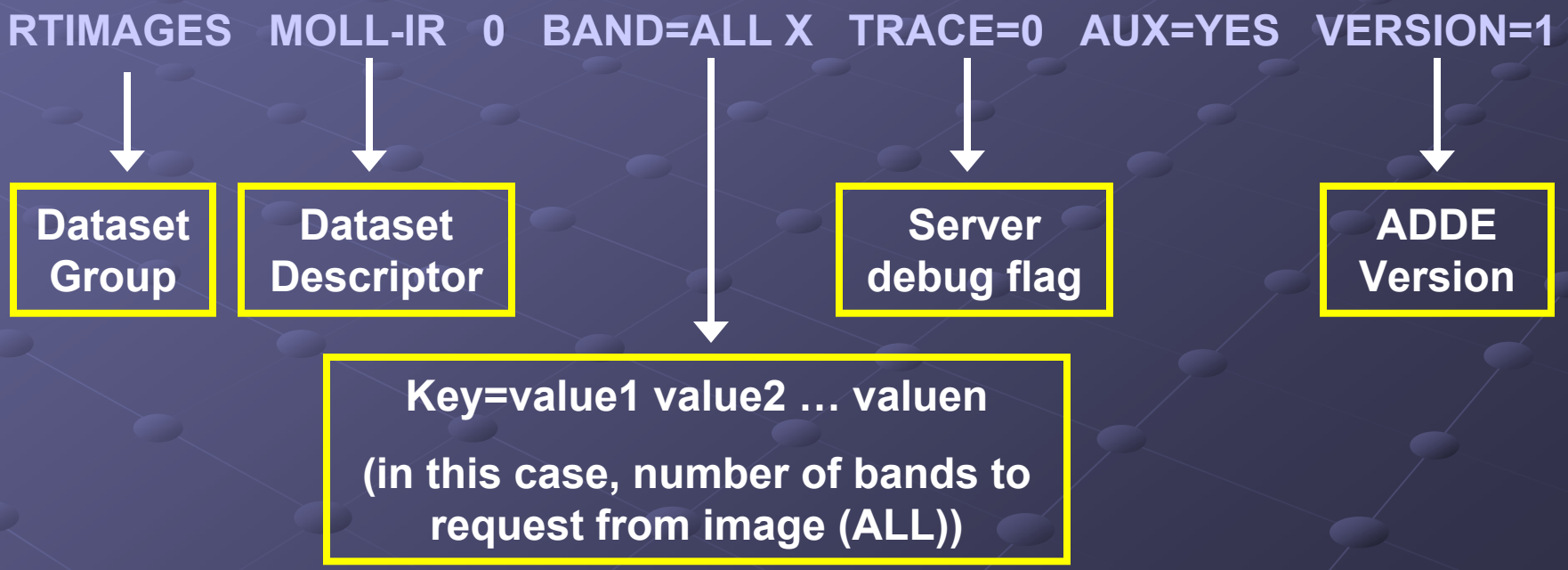
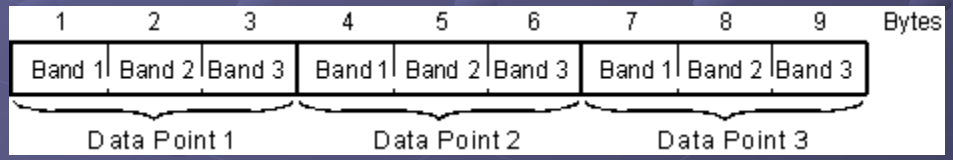


Image data

- ADDE image data model supports multi-banded images
- 2 main types of requests
  - Directory (ADIR) and data (AGET)
- Returned image object models McIDAS AREA format:
  - Directory block
  - Navigation block
  - Calibration block
  - Supplemental (AUX) block
  - Data block
  - Comment block

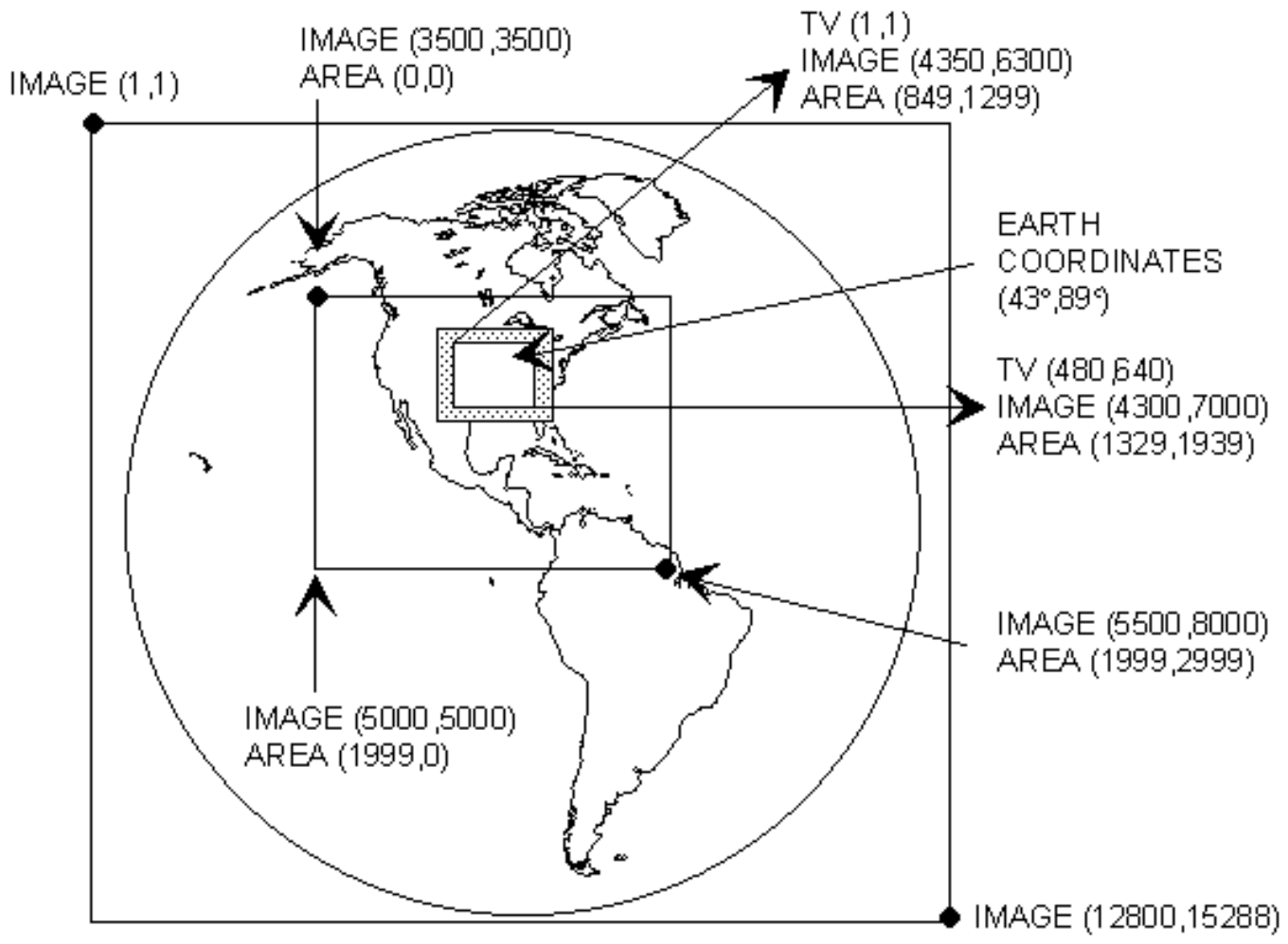
Image Object Details

- Image Directory
  - contains a list of ancillary information about the image, such as the day and time, number of lines and data points, the satellite ID, and the number of spectral bands.
- Data Block
  - contains the matrix of image data values. Multibanded image have values interleaved:



- Navigation block
  - contains information for determining the location of data points in physical space. Client must have navigation module which uses this block to convert (line,element)<->(lat,lon) for geolocation
- Calibration block
  - contains the information for converting image data from its internal (stored) units to more meaningful physical units, such as radiance or albedo
- AUX block
  - contains additional information that is specific to a data type. For example, information specific to radar data is stored in this block. Also, the latitude/longitude grid for the LALO navigation is stored in this block.

Image Coordinates



AREA=FILE coordinates, ignore TV coordinates for now

Image Data (con't)

- An image directory (ADIR) request returns all images matching the request
- An image data (AGET) request returns only one image at a time.
- Request refinements
  - Location can be specified by image, file or lat/lon coordinates.
  - Single or multi-banded images
  - Day and time
  - Calibration units
  - Size (number of lines/elements)
  - Resolution (line/element magnification)
  - Relative position number (i.e. last N images)

Grid Data

- Grid data is two-dimensional data representing a parameter along an regularly spaced matrix (e.g., model output, objective analysis).
- 2 main types of requests
  - Directory (GDIR) and data (GGET)
- Grid object consists of:
  - Grid Header
  - Data block

Grid Object Details

- Grid Directory
  - contains a list of ancillary information about the grid, such as the parameters and units of the data in the grid, the level in the atmosphere or ocean the data represents, the grid navigation information, and the time.
  - Client must have navigation module which uses the navigation info to convert (row, column)<->(lat,lon) for geolocation
- Data block
  - contains the matrix of gridded data values.

Grid Data (con't)

- A grid directory (GDIR) request returns stream of grid directories matching request
- A grid data (GGET) request returns stream of grid data objects matching request
- Request refinements:
  - Parameters or Derived quantities
  - Levels
  - Model run and/or valid day/time
  - Originating center
  - Maximum number of grids to return

Point Data

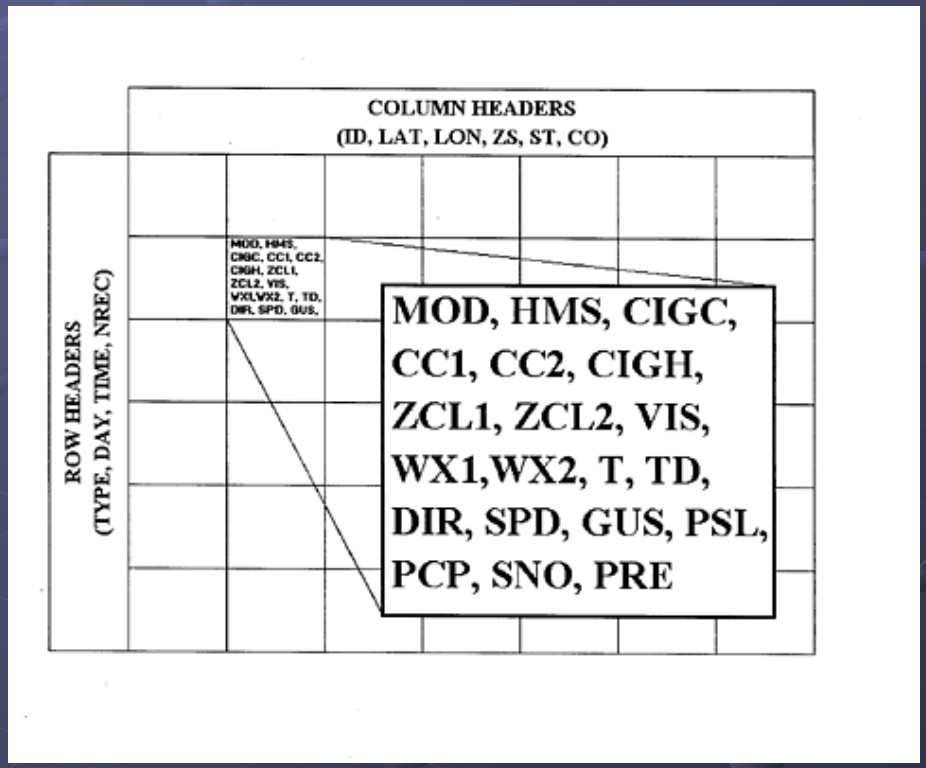
- Point data typically represents data occurring at irregularly spaced locations around the Earth (ex. Surface observations, upper air reports, lightning flashes)
- Main type of request is for data (MDKS)
- Point data object consists of 5 blocks:
  - Parameter block
  - Unit block
  - Scale block
  - Form block
  - Data block

Point data object details

- The *parameter block* contains a list of the parameter names in the point object returned by the server.
- The *unit block* contains a list of units for the parameters returned by the server.
- The *scale block* contains a list of scaling factors for the floating- point values returned by the server.
- The *form block* contains a list of the return forms for each of the parameters.
- The *data block* contains the actual data values returned by the server.

MD File structure

- Meteorological Data (MD) file schema determines data layout
- An MD file is like a spreadsheet with each cell containing a predefined number of data values. Each cell contains data for a specific location at a given instant in time.



Point data (con't)

- Request refinements
  - List of parameters
  - Maximum number of obs to return (default 1)
  - The SELECT clause gives the user the ability to subset on any parameter in the dataset.
- Examples:
  - **SELECT='T[F] 40 50; ST WI, MI; TIME 12 13' MAX=ALL**
    - Selects all parameters for all observations between 40 and 50 degrees Fahrenheit for stations in Wisconsin and Michigan between 12 and 13 UTC
  - **SELECT='ID KDEN' PARM=T TD PRE MAX=ALL**
    - Selects all temperature, dewpoint and pressure values for all times in the dataset for Denver

Text Data

- There are 3 types of text data that are served up by ADDE:
  - Flat files, ancillary data files
  - Weather bulletins such as forecasts, warnings, watches
  - observational data, such as METAR or RAOB reports
- 3 types of requests:
  - TXTG ASCII text file
  - WTXG textual weather information
  - OBTG observational weather text
- Returned data has a header and the text

Java and ADDE

- Client interface developed and refined collaboratively by Unidata, SSEC and Australian Bureau of Meteorology (BoM)
- Provides ADDE data access for Java-based data analysis and display tools (e.g., IDV, Matlab)
- Uses specialized URL
  - adde://server/request?keyword_1=value_1&keyword_2=value_2 …keyword_n=value_n
    - “adde” specifies protocol
    - “request” specifies data type/action
    - keyword/value pairs refine request
- Bundled with VisAD component library
  - Package edu.wisc.ssec.mcidas.adde contains core package
  - Package visad.data.mcidas has classes for converting ADDE data objects into VisAD objects.

Java ADDE Use in Applications

- Unidata
  - Integrated Data Viewer (IDV)
    - Access to satellite, Level III radar, METAR, synoptic, upper air and profiler data
- SSEC
  - McIDAS AREA to netCDF converter for AWIPS use
  - Java client for browsing and copying real-time and archive imagery
  - MODIS data exploration
- BoM
  - Development of subservers (Oracle/NEONS, radar)
  - Java-based clients
    - Australian Marine Forecast System
    - Tropical Cyclone Forecast

ADDE Use in IDV

- The IDV uses ADDE to access:
  - Satellite and Level III radar imagery
  - Surface (METAR, synop) data
  - Upper Air (RAOB) data
  - Profiler data
  - Text data
- ADDE data objects are converted to VisAD data objects
  - Navigation of images done through AreaCoordinateSystem

IDV/ADDE Demo

- Satellite image (GINI East – 1km VIS)
- Level III Radar selection (Denver)
- RAOB
  - Sounding
  - Plan view
- Profiler (Platteville)
  - Time/Height
  - 3D view
- Text
  - File (PUBLIC.SRV)
  - Denver METAR
  - Denver RAOB

Summary Analysis

Strengths

- Supports many data formats, especially for imagery
- Subsetting capabilities
- Datasets can be queried for metadata
- Server freely available to research and education institutions through Unidata
- Java/C/FORTRAN client APIs
- Java client freely available

Weaknesses

- Point data limitations
  - MD file limitations (4 character names/unit names, 400 parameter limit, scaled integers)
- Grid data limitations
    - 2D grids only, 4 character param/unit names
- Server configuration (best done from McIDAS)
- Currently not separate from McIDAS

Future Development Ideas

- Middleware to support metadata queries for available times and parameters (not just catalogs)
- Enhanced data choosers which take into account the semantics of the datasets
- Support more navigation modules in Java
- Enhance netCDF point server
- Serve up GEMPAK grids
- Java server

ADDE Resources

- McIDAS Programmer's Reference
  - http://www.ssec.wisc.edu/mug/prog_man/2003/prog_man.html
- McIDAS User's Guide (ADDE section)
  - http://my.unidata.ucar.edu/content/software/mcidas/2003/users_guide/index.html
- Javadoc for AddeURLConnection class:
  - http://www.ssec.wisc.edu/~dglo/visad/edu/wisc/ssec/mcidas/adde/AddeURLConnection.html