Staff Status Reports

March 2015

- ACADIS - Arms, Ramamurthy
- Equipment Awards - Mitchell, Ramamurthy
- AWIPS II/Gempak - James
- Cloud Computing Activities - Fisher, Arms, Caron, Ho, James, Schmidt, Yoksas
- Community Services - Young, Dirks, Weber
- IDV with RAMADDA - Ho, Chastang
- International Activities - Yoksas
- IDD - Schmidt, Weber, Yoksas
- LDM - Emmerson, Schmidt, Yoksas
- McIDAS - Yoksas
- NetCDF - Fisher, Heimbigner, Rew
- Python - Arms, Chastang, Domenico, Fisher, May, Rew
- Rosetta - Arms, Oxelson, Weber
- Support - Yoksas, Oxelson
- THREDDS - Caron, Arms, Davis, May, Ward-Garrison
- Unidata Outreach Progress - Domenico

UCAR COMMUNITY PROGRAMS

The Unidata Program Center is a member of the UCAR Community Programs, is managed by the University Corporation for Atmospheric Research, and is funded by the National Science Foundation.

ACADIS

Status Report: September 2014 - March 2015

Mohan Ramamurthy, Sean Arms, Jeff Weber

Strategic Focus Areas

Community Services supports the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   The ACADIS Data Portal is creating an effective way to access Arctic data
2. **Develop and provide open-source tools for effective use of geoscience data**
   Unidata is creating an ASCII to netCDF translation tool that will allow a large amount of Arctic data to be translated to netCDF CF
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   ACADIS is an exemplar for data portals
4. **Build, support, and advocate for the diverse geoscience community**
   ACADIS continues to champion useful access to data holdings

Background Information

The new Advanced Cooperative Arctic Data and Information Service (ACADIS) is a joint effort by the National Snow and Ice Data Center (NSIDC), the University Corporation for Atmospheric Research (UCAR), UNIDATA, and the National Center for Atmospheric Research (NCAR) to provide data archival, preservation and access for all projects funded by NSF's Arctic Science Program (ARC). ACADIS builds on the CADIS project that supported the Arctic Observing Network (AON). This portal will continue to be a gateway for AON data and is being expanded to include all NSF ARC data.

Activities Since the Last Status Report

- Unidata is moving forward with our contribution to homogenize the data for ease of re-use by the larger scientific comunity. This is being addressed with the Rosetta project, Rosetta's status report can be found here
- The ACADIS project is wrapping up year 4 funding. NSF has decided that the management of Arctic Data will be recompeted, and the solicitation has recently been released.
- The ACADIS project has been asked to submit a budget for a six month extension to ensure that archival of Arctic data still occurs during the gap between the end of year 4 funding (June 2015) and the new NSF award (FY 2016).

Planned Activities

Ongoing Activities

The ACADIS team is currently working on submission of a proposal for the recompete of the archiving and management of Arctic data (LoI: April 17, 2015, Full Proposal: May 18, 2015).

The 2015 Unidata Equipment Awards Program

Sponsored by the National Science Foundation

NSF provides the Unidata Program Center up to $100k in equipment grant funds each year. In alignment with the Unidata 2018 proposal, the Equipment Awards Program is designed to broaden participation and promote the use of Unidata tools and systems (e.g., THREDDS, NetCDF, IDV, GIS connections) to support education and research on various aspects of climate studies (e.g., diagnostics, change and impacts), by providing grants to be used in the procurement of new computers and equipment including upgrades to existing classroom and laboratory equipment.

This year, special consideration will be given to proposals that include one or more of the following:

- Installation of a prototype AWIPS II standalone EDEX server and CAVE client, coupled with the Unidata LDM, to test data ingest and display both locally, and using the CAVE thin client to connect to remote servers
- Implementation of or pilot projects with remotely-accessible storage systems for geoscience data (“cloud-based storage”)
- Implementation of or pilot projects with remote server-based data analysis or visualization systems (“cloud-based analysis”)

A Request for Proposals was sent out on December 22, 2014 with a March 06 submission deadline. A review panel has been formed and will meet on March 25 at the Unidata Program Center to review the proposals that are received.

AWIPS II and GEMPAK

Status Report: October 2014 - March 2015

Michael James

Strategic Focus Areas

AWIPS II activities support the following Unidata funding proposal focus areas:

1. Enable widespread, efficient access to geoscience data
2. Develop and provide open-source tools for effective use of geoscience data
3. Provide cyberinfrastructure leadership in data discovery, access, and use
4. Build, support, and advocate for the diverse geoscience community

Cloud activities

Unidata made available a free and open EDEX data server in the Microsoft Azure cloud for most of 2014 and early 2015. While active, this instance served data to over a dozen universities, including some universities who ran CAVE in classroom computer labs.

We're still unsure how many CAVE clients can simultaneously connect to a single EDEX instance, but with the Azure cloud instance we noticed that bandwidth was of critical importance for usability. Loading the 0.5 degree GFS in the NCP, for example, and the delay in rendering within the client is significant (10-20 sec for 10 frames). This was the reason I turned off the 0.5 and ingested 1.0 degree GFS on Azure and default on 14.2.1 EDEX installs at Universities.

Azure download volume ranged between 500-1000MB daily.

Connecting a dozen clients to an internal EDEX server shows much better performance, which is clear. So far we have reach bandwidth limits but not Qpid message broker pool size limits. So if bandwidth is good it appears that AWIPS II can scale to classroom sizes.

Building RPMs

Branches of the "AWIPS2 Builds" project version 14.4.1 and later have been reorganized into multiple code repositories instead of a single one as before. With this reorganization came changes in the build structure. Unidata has built custom RPMs for AWIPS II 14.2.1 to install at university sites, and would like to continue to build RPM sets for our user base. After updating the Unidata development branch to 14.4.1 the build structure is no longer functional.

- Raytheon has not to my knowledge made available 14.4.1 RPM build instructions and working build scripts to development organizations.
- The 14.4.1 branches (master, omaha, ncep, etc.) contain build files and scripts which have not been updated since version 14.2.1, and contain build targets for projects which have been removed from the baseline (such as datadelivery and dataaccess plugins), so the build fails.
- Manually editing the 14.2.1 build scripts to adapt to 14.4.1 is a trial and error process and with enough edits we can force a build but the RPMs fail on install since they are missing critical executable files. Requests for help from NCEP have still not solved this problem as they themselves also find the build process confusing.

- 14.4.1 saw the introduction of Groovy feature build files in awips2-builds/build/build.core (Feature.groovy and FeatureParser.groovy added December 2014 by Bryan Kowal), which suggest that Raytheon has moved from an Ant build to using a Groovy plugin in Eclipse. There is no documentation as to how these scripts work.

AWIPS II updates since last status report

- Moved from java 1.6 to 1.7
- Removed autostart CAVE and alertviz functionality, so no longer will clients start on login as happens at WFOs.
- Created a Boulder, Colorado "localization" RPM to investigate what needs to be done in order to change an EDEX server site localization (answer: a lot, more than should be necessary).
- Updated profile.d scripts to only source AWIPS II python for 'root' and 'awips' users, allowing other software packages like IDV to use system python without conflict.
- Increased default LDM product queue size from 500M to 1500M after noticing data holes on Azure instance.
- New grids supported: NAM 5km, CMC, NAM 32km, NAM243, HRRR (very slow rendering), FNMOC (some still in development)
- Reorganization of satellite postgres database tables to support UNIWISC imagery.
- Separated the build of shapefile RPMs from other EDEX RPMs, no need to re-build very large shapefile executable file when minor changes are made to EDEX.
- Enabled NCP-specific data decoders ncpafm, ncscd, nctext, ntrans on EDEX. Noticed some large ingest latency due to dual decoding (by Raytheon 'text' and NCEP 'nctext' plugins).
- Restarted UPC development branch as tracking remote ncep_14.4.1, folding in previous updates to UPC 14.2.1 orphan branch.

AWIPS II Training Workshop

As of 2014, the AWIPS II training workshop consists entirely of new material, adapted from the Raytheon document and NWS training workshop course material. Topics include

- Understanding Linux Tools in the AWIPS II Environment
- AWIPS II System Architecture
- Install the EDEX Data Server
- Regular Expressions in AWIPS II
- The Qpid Message Broker
- Data and Process Logging
- Data URIs
- Data Purging
- Manual Data Ingest
- Grid Configuration in EDEX
- Adding a New Grid to EDEX
- Site Localization Guide
- CAVE Localization Perspective
- CAVE National Centers Perspective
- Adding a new Shapefile to CAVE

Updated NCP "Overlay" menu to indicate what overlays are currently loaded on screen and allow toggle of overlays on-click with new checkboxes.

Continued support of UNIWISC McIDAS imagery support (1km/4km GOES, GOES E/W composites, Arctic POES composites, Global Rectilinear and Mollweide composites)

Further grid support: CONDUIT, FNMOC and CMC feed partially supported (though CMC GEM model is known to seg fault 14.2.1 grib decoder, uh oh)

Working with a HS student on a college credit project to deploy LDM/EDEX in the Softlayer.com cloud.

Source Code

Moved large files outside of source code repo to allow for upload to Unidata github account (max file/commit pack size of 50M/100M firm)

Size

- awips2-builds: 6 GB
- awips2-core 55M
- awips2-core-foss 360M
- awips2-foss 212M
- awips2-build-server 475M
- external from repos: awipscm 6.2 G
- Total: 13.3 GB of source code and files to build and deliver AWIPS II to universities

Files and lines of code

- num of files: 49,262
- num lines of code: 16,214,405
- num commits to baseline: 7,842

Ongoing and Planned Activities

- Deliver Unidata AWIPS II 14.4.1 to universities in spring/summer 2015. Git reorganization of tracking branch and merging will allow Unidata to more easily update our development branch and deliver updates along with remote NCEP/WFO branches.

GEMPAK/NAWIPS

- GEMPAK now includes the capability to request and display data from an EDEX data server, both on internal networks and remotely (assuming remote tcp access is granted by the EDEX server). Data is returned from the EDEX request JVM via Thrift serialization as a Python object, then converted to a Fortran object for processing and display by GEMPAK programs. The functionality is limited (no NMAP2 data loading, for example), and only certain data types are supported (sfc and uair obs and gridded data), but the EDEX Data Access Framework Python interface allows for all data types to be returned.
- Created a full GEMPAK grid inventory to display projection, available times and grid parameters.
- NOAAport grib tables updates December 2014
- NCEP NAWIPS maps/bounds updates
- Internal memory increased to handle processing and display of 5km NAM
- CONDUIT GFS pattern action updates

Updated: March 4, 2015

Unidata Cloud Computing

Status Report: October 2014 - March 2015

Fisher, Arms, Caron, Ho, James, Schmidt, Weber, Yoksas, Chastang

Strategic Focus Areas

Unidata's *Cloud Computing* activities support the following Unidata funding proposal focus areas:

1. **Enable widespread, efficient access to geoscience data**

   Making Unidata data streams available via various commercial and private cloud services will allow subscribers to those services to access data quickly and at low cost.

2. **Develop and provide open-source tools for effective use of geoscience data**

   Running existing Unidata-developed and supported tools and processes (e.g. IDV, RAMADDA, generation of composite imagery) in a range of cloud environments makes these tools and data streams available to cloud service subscribers at low cost. It also gives us insight into how best to configure existing and new tools for most efficient use in these environments.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**

   Unidata is uniquely positioned in our community to experiment with provision of both data and services in the cloud environment. Our efforts to determine the most efficient ways to make use of cloud resources will allow community members to forego at least some of the early, exploratory steps toward full use of cloud environments.

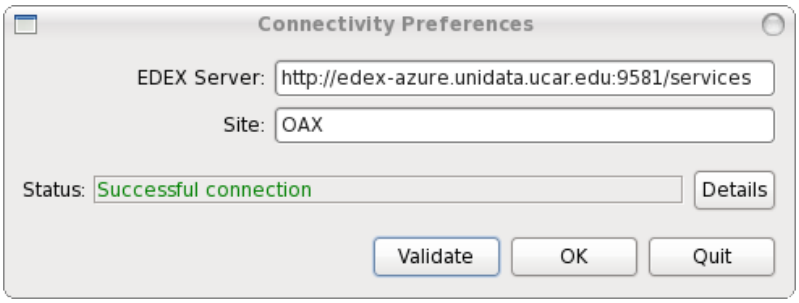
AWIPS II Cloud Servers

- Unidata is testing small footprint EDEX servers (no NEXRAD Level 2 or 3 or high-resolution CONDUIT models) on both **Microsoft Azure** and **Amazon EC2** cloud server environments.

- An EC2 instance was created cooperatively by Unidata and Embry Riddle Aeronautical University (ERAU) as part of ERAU's equipment grant award. This instance, which is configured to run AWIPS II EDEX, has the following characteristics:

  - AWIPS II Size on Disk: **220 GB**
  - Grids: **97 GB/day raw, 51 GB processed**
  - NGRID: GFS 201, 212, 213, GFS/LAMPTstorm, MOSGuide, NAM 12km, NamDNG5, RTMA 5km and 2.5km, SREF 40km, GEFS. HiResW-NMM and HiResW-ARW, RAP 13km
  - CONDUIT: GFS global 1.0/2.5, NDFD, NAM 40km and 90km, RAP 20km and 40km, GFS 0.5 turned off
  - FSL2: HRRR (72 GB raw)
  - CMC: Regional GEM Model breaks grib2 decoder (turned off)
  - UNIWISC: 5 GB/day
  - FNEXRAD: DHR, DVL, EET, HHC, N0R, N1P, NTP
  - NEXRAD3, FNMOC: turned off

  This Azure instance is currently serving data to AWIPS II 14.2.1 beta testers:



IDV Application-Streaming Cloud Servers

This project is evaluating *application streaming* as a strategy for making the IDV available to a new generation of users and computing platforms. It is using the Microsoft Azure cloud platform to look at delivering cloud-based IDV-as-a-service instances to our user community on an as-needed basis. The result will be a better understanding of how the IDV works in cloud environments and any changes that might improve that performance.

This project also serves as a pilot program; with it we will further develop expertise related to cloud computing and application streaming. This will allow us to extend cloud-based software offerings beyond the IDV to other Unidata projects.

Issues

- How does technology like Docker mitigate the need for multiple VM instances.
- How best to adapt mouse-driven interfaces to a touch-based interface, while minimizing the need to re-engineer any part of the software package.
- Evaluation of bandwidth requirements for acceptable IDV use.
- How to make this transition *seamless* and *painless* to our user community.
- Evaluate the extent to which we can use "off-the-shelf" technology and under what circumstances do we need to create our own protocols and packages.

Current Status

- We are able to instantiate cloud-based IDV instances, which are then streamed via existing remote-desktop protocols to iOS devices. Nothing in the existing technology limits this to iOS devices, however; those are simply the devices on hand for testing.
- Using the Azure Web API, we are able to dynamically allocate and provision VMs for use with hosting the IDV.
- Current efforts are focused on creating a web dashboard which will allow users to register and manage IDV-streaming requests.

IDD Product Generation and Additional Experimentation

- Unidata continues to operate mid-sized virtual machine instances in both the Amazon EC2 and Microsoft Azure west clouds for the purpose of generating image products for the IDD FNEXRAD (NEXRAD Level III national composites) and UNIWISC (GOES-East/West image sectors) data streams. The EC2 instance is currently the primary source of the FNEXRAD and UNIWISC data streams to IDD participants. The plan is to transition to the use of the Azure cloud instance to reduce recurring costs of running an instance in EC2. (Ward Fisher spearheaded a Unidata effort that resulted in Microsoft awarding use of 32 small VM instances in Azure free-of-charge for approximately 1 year).

- Unidata implemented a TDS instance in the Azure west cloud for testing. This effort was put on hold pending renewal of Microsoft's grant of Azure resources.

- A mid-sized VM instance in Azure is being used to investigate running the IDV in the cloud. RAMADDA has been installed and can generate non-interactive IDV displays using Xvfb for the needed XWindow environment.

Docker

- Docker is a new cloud-centric technology that borrows from the notion of containers from the shipping industry to facilitate installation and deployment of server side applications in a cloud environment. We are investigating and exploring the possibility of creating Docker containers for cloud distribution and installation for a variety of Unidata technology offerings including RAMADDA, THREDDS, IDV, AWIPS II (EDEX/CAVE), and LDM.

- We have reserved the Unidata namespace at DockerHub, and we have a prototype Docker image for RAMADDA and Unidata Python.

- We are also educating ourselves on Docker technology and will have an internal Unidata talk on Docker, shortly.

Future Activities

The UPC is seeking User Committee input on two possible cloud experiments:

- Investigate the feasibility of replicating the RAMADDA content on motherlode.ucar.edu in either the Amazon EC2 West or Microsoft Azure West clouds.

  It is thought that providing "cloud" access to a large portion of the RAMADDA content currently hosted on motherlode would help to mitigate impacts that would be experienced in the event of a catastrophic failure on motherlode. It is also thought that the serving of IDV bundles from the cloud would not result in large amounts of outbound network traffic from "the cloud", so costs would be modest.

- Investigate replicating the decoded GEMPAK content currently available on all motherlode class machines (motherlode.ucar.edu, atm.ucar.edu, idd.ssec.wisc.edu, lead.unidata.ucar.edu and weather.rsmas.miami.edu) in "the cloud" and then encourage sites who are web scraping these data to establish their own presence in the same "cloud" and then grab the data from there. It would then be the end-user's responsibility to pay for the outbound network traffic from their own cloud instances.

Updated: March 2, 2015

Community Services

Status Report: October 2014 - March 2015

Doug Dirks, Jeff Weber, Joshua Young

Strategic Focus Areas

Community Services supports the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   We monitor and collaborate with data sources to stay apprised of impending changes and to advocate for the needs of our user community. We provide user workshops, tutorials, and community workshops to help build supportive relationships between community members.
2. **Develop and provide open-source tools for effective use of geoscience data**
   We promote Unidata tools and software for multi-disciplinary use, with an eye toward finding additional research and educational communities that can benefit from our work.
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   We work with government and industry data providers to secure access to data for Unidata community members.
4. **Build, support, and advocate for the diverse geoscience community**
   We coordinate with our governing committees to find ways to expand Unidata's community participation. We use our web site, electronic newsletters, and social media to keep community members informed about enhanced data services, software tools, and cyberinfrastructure.

   We participate in UCAR/NCAR and NSF projects for underrepresnted populations and minority communities (SOARS, AIHEC, outreach to HBCUs). We provide services and tools to facilitate education and research in diverse communities. We work to broaden the Unidata community by participating in student and professional conferences.

Activities Since the Last Status Report

News@Unidata blog

Posts to the News@Unidata blog appear regularly, but not on a specific schedule. Some highlights:

- Some highlights:
  - A Multi-platform Observing System in the Balearic Islands
  - 2015 Unidata Users Workshop: Save the Date
  - Unidata Staff at AGU Fall 2014 Meeting
  - HIWPP Seeks Engagement with Scientific Community
  - Unidata Summer Student Internship Available!
  - AMS 2015 Conference Highlights from the Unidata Staff
  - Software release information
  - Community job postings
  - Community meetings and other announcements
- Dependencies, challenges, problems, and risks include:
  - Finding community members willing to contribute stories (or story ideas) for the blog is an ongoing challenge

Community Outreach and Services

- Progress has been made on the following:
  - Representing the Unidata community in professional society events such as the AMS and AGU annual meetings
  - Actively engaged in the UCAR Education Working Group and UCAR-Connect to explore novel techniques and applications for the educational community
  - Work with UPC staff to set up and maintain EDEX servers for equipment award grantees
    - ERAU is enjoying EDEX access in the cloud for both the Daytona and Prescott campuses as well as the Phoenix and Flagstaff WFO's
  - Creation of the Teaching Resource Network on RAMADDA servers, the search page can be found here
  - Continue to work closely with NOAA to distribute The HRRR, FIM, and HIWPP model output and coordinate Unidata's tools to effectively make use of this model output
  - Maintain effective communication with CUAHSI by being an active member of the CUAHSI HIS standing committee, which we helped get going by being at the first meeting in 2000 in Austin
  - Work closely with NCAR scientists and the CUAHSI HIS group by attending the NFIE workshop March 16-19 in Tuscaloosa at the National Water Center
  - Engage with NCEP to discuss the contents of the CONDUIT and NOAAPort feeds streams to more effectively serve our community
  - Creation of a *Data Management Resource Center* to collect information about tools and best practices with the goal of helping community members prepare data management plans
  - Explored the application of agile software development principles to data management in an effort called Agile Data Curation
  - Pilot site visits at the Metropolitan State University of Denver, Colorado State University, and University of Northern Colorado
  - Engagement with NSF's Earthcube through the Liaison and Engagement Committees
  - Planning for the 2015 Users Committee Workshop
  - Release of survey on .25-degree GFS to CONDUIT users

Social Media Outreach Activities

- Progress has been made on the following:
  - We continue to update the Facebook, Google+, and Twitter feeds.
  - We continue to publish short videos/screencasts on the Unidata YouTube channel.
- Dependencies, challenges, problems, and risks include:
  - Engagement with Unidata social media streams among community members is not particularly high

Planned Activities

Ongoing Activities

We plan to continue the following activies:

- NAWIPS migration to AWIPS II, including the overall AWIPS II project
- Ongoing development of news articles for publication through News@Unidata
- Continue to support and contribute to governing committees
- Seminars
- Outreach
- Engagement with professional societies
- Triennial workshop planning and coordination
- Further development of the Data Management Resource Center
- Further work on Agile Data Curation
- Site visits as the budget allows

New Activities

Over the next three months, we plan to organize or take part in the following:

- Additional community guides
- Further CONDUIT surveys regarding future data additions

Over the next twelve months, we plan to organize or take part in the following:

- Submit and present further talks on the Teaching Resouces Network and Data Management Resources Center at geoscience professional societies.

- Facilitate community identification of preferred data management policies and enforcement

Beyond a one-year timeframe, we plan to organize or take part in the following:

- Pursue funding to support a Data Management pilot project to implement Unidata services on several small geoscience projects

Areas for Committee Feedback

Community Services is requesting your feedback on the following topics:

1. **Are there other areas that are ripe for a guide similar to the Data Management Resource Center?**
2. **What would help you to use the Teaching Resources Network?**

Relevant Metrics

Statistics from the Community pages on the Unidata web site. Comparisons are made with statistics from the previous six-month period.

All Community pages

Most recent six months

- 40,128 unique pageviews (43,313 in previous period)
- 7.1% of total unique pageviews to site (7.1% in previous period)

Top community pages

1. All blog pages
   News@Unidata blog and developers' blog
   - 27,062 unique pageviews (31,724 in previous period)
   - 67% of total community pageviews (73% in previous period)
2. Events pages (www.unidata.ucar.edu/events/)
   Information about training courses and other events
   - 4676 unique pageviews (3379 in previous period)
   - 11.6% of total community pageviews (7.8% in previous period)
3. Community pages (www.unidata.ucar.edu/community/)
   Information about Unidata community events and governance
   - 3825 unique pageviews (3217 in previous period)
   - 9.5% of total community pageviews (7.4% in previous period)
4. About Unidata (www.unidata.ucar.edu/about/)
   Information about Unidata
   - 3590 unique pageviews (3877in previous period)
   - 9.0% of total community pageviews (9.0% in previous period)

Social media statistics, March 13, 2015

1. # of Twitter followers: 397 (359 in September 2014)
2. # of Facebook followers: 383 (317 in September 2014)

Prepared March 2015

IDV with RAMADDA

Status Report: September 2014 - March 2015

Yuan Ho, Julien Chastang

This report updates the status of Unidata's Integrated Data Viewer (IDV) development efforts since the last report (September, 2014).

Strategic Focus Areas

The *IDV* group's work supports the following Unidata funding proposal focus areas:

1. **Enable widespread, efficient access to geoscience data**
   The IDV is a state of the art geoscience visualization application. It gives users the ability to view and analyze a rich set of geoscience data, including real time data, in a seamless and integrated fashion. This analysis is captured in IDV bundles. RAMADDA is a content management system and service specifically tailored towards the sharing and distribution of IDV bundles facilitating distribution of scientific data and analysis.
2. **Develop and provide open-source tools for effective use of geoscience data**
   The IDV has been an open-source project for several years. The IDV is available on the github version control platform for greater open-source collaboration. The IDV provides users the unparalleled ability to analyze, integrate, and visualize heterogeneous geoscience data in two, three, and four dimensions. The IDV coupled with RAMADDA enables geoscience specialists the capability to share and collaborate their IDV analysis via social scientific networks.
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   RAMADDA allows geoscience specialists the ability to search and publish their IDV bundles on-line. Unidata's RAMADDA installation enables the IDV team to communicate more effectively to our users concerning their IDV issues. Specifically, during support ticket conversations, the IDV team requests that users upload pertinent data to RAMADDA for analysis. The IDV team also takes advantage of RAMADDA to share instructional IDV screencasts with users.
4. **Build, support, and advocate for the diverse geoscience community**
   Unidata offers yearly multi-day training and occasionally regional workshops for IDV and RAMADDA. The IDV coupled with RAMADDA enables our earth science community partners to distribute geoscience data and metadata through web-based technologies thereby fostering scientific collaborations. Moreover, the IDV's ability to share bundles through RAMADDA creates a scientific social and collaborative network for the geoscience community.

Activities Since the Last Status Report

System Changes

Latest netCDF-Java Version: 4.5
The version of the netCDF-Java library currently distributed with the IDV is 4.5. The 4.5 release includes improvements related to I/O, NIO.2, HttpClient 4, improved projections, and updated GRIB support. See netCDF-Java Library for a more details on these changes.

IDV distribution, installation and program execution: There have been several changes related to security and distribution of the IDV. In particular, the Windows, Mac OS X, and Java Webstart versions of the IDV installers are now signed with certificates from certificate authorities (CA). We continue to fix bugs and make improvements to the mechanism that provides an IDV old version warning to users. The Install4J script continues to be maintained and improved. These changes result in a smoother installation and execution of the IDV for our users.

RAMADDA plugin: The RAMADDA IDV plugin has been enhanced to facilitate the publication of movies

and images, including images as zip files, to RAMADDA. This change improves the display of movies published from the IDV to RAMADDA.

IDV Webstart: The Java Webstart version of the IDV has been fixed following problems associated with the Java 3D migration to Java 7.

Display Changes

CSV Time Matching: The IDV now has time matching capability for CSV point data files. In this scenario, users first need to create the display, and subsequently specify a driver time. In particular, users need to select View > Times > Uses Time Driver Times in the display control of the CSV display. Reloading the data will create the display with matching driver times.

IDV Webcams: We continue to maintain a list of webcams that can be accessed through the IDV.

Data Changes

CFRadial netCDF Format Support: The IDV now has the capability of viewing CFRadial netCDF format radar data files. We collaborated with engineers from NCAR/EOL to achieve this result and we continue to make changes, improvements, and bug fixes in this area.

AMRC Text Point: IDV can now handle text point data from the Antarctic Meteorological Research Center (AMRC).

IDV and RAMADDA Training and Conference Attendance

- Produced three new IDV training videos:
    - Creating a 3D map
    - Creating a Jython Formula
    - Combining Data with Time Matching
- Presented "Using Progressive Resolution to Visualize large Satellite Image dataset" at the 2015 AMS conference in Phoenix.
- IDV introduction at WRF Tutorial, January 28-30, 2015
- Participated in the NSF DataViz Hackathon for Polar CyberInfrastructure on November 3-4, 2014

RAMADDA

Cloud Activity: Docker is a new cloud-centric technology that borrows from the notion of containers from the shipping industry to facilitate installation and deployment of server side applications. We are investigating and exploring the possibility of creating Docker containers for easy distribution and installation of RAMADDA in a cloud environment. We have a prototype RAMADDA Docker image we are are experimenting with and intend to build upon in the near future.

Planned Activities

New Activities

Preparing to submit a NASA ROSES ACCESS proposal in collaboration with SSEC.

Preparing for forthcoming WRF workshops and tutorials.

Preparing for Costa Rica IDV/McV training for May, 2015.

Preparing for University of South Florida Regional Workshop in April, 2015.

Relevant Metrics

The IDV team continues to provide the geoscience community with high-quality support through e-support software and idv-users maillist. The volume of e-support remains high and constitutes a large fraction of our daily activities. In the last half year the IDV team has closed ~175 e-support tickets. Each individual ticket may and often does involve many back-and-forth messages.

The number of both casual and regular IDV users is stable or increasing. In February 2015, there were 130 IDV users starting the IDV more than 20 times per month compared with 119 users for the same period in 2014. In the same month, there were 74 IDV users starting the IDV more than 30 times per month compared with 67 users for the same period in 2014. Each month there were about 1700 to 2000 unique IP addresses where IDVs were run at least one time. Notably, the number of the IDV users from the edu domain is the largest among the top level domains (about 10 times the second top level domains). Readers can find the raw metrics at http://www.unidata.ucar.edu/software/idv/logging/left.html.

In the area of greater collaborative development, since the migration of the IDV project to github, we have accepted 85 "pull requests" or code contributions from internal and external collaborators.

In the area of online IDV training, the Youtube IDV instructional videos have been viewed over 4,840 times compared with 3,000 from six months ago.The video on trajectories has generated an especially great interest.

Prepared March 2015

International Activities and Collaborations

Status Report: October 2014 - March 2015

Tom Yoksas

Strategic Focus Areas

The International Activities and Collaboration group's work supports the following Unidata funding proposal focus areas:

1. Develop and provide open-source tools for effective use of geoscience data
   The majority of tools downloadable from Unidata are available free-of-charge to everyone (the exception being McIDAS-X).

2. Provide cyberinfrastructure leadership in data discovery, access, and use
   Activities of the Unidata Program Center are routinely provided to the worldwide atmospheric science community. Strategic partnerships with leading organizations in other countries minimize the impact on UPC staff.

3. Build, support, and advocate for the diverse geoscience community
   By informing the international atmospheric science community of the products, data and services available in the Unidata Program, an extended community has been enabled.
   Non-U.S. users of products available from Unidata reflect, in a number of cases, minority constituencies in the U.S. atmospheric science community.

Activities Since the Last Status Report

There are no significant new activities since the last status report.

Unidata will conduct a training workshop focused on using the IDV, McIDAS-V (for Hydra) and RAMADDA in San Jose, Costa Rica in May, 2015.

Planned Activities

New Activities

Unidata will be conducting an IDV/McIDAS-V/RAMADDA training workshop in San Jose, Costa Rica during the week of May 4, 2015. This effort is being organized by Maria del Rosario Alfaro Ocampo, a UCAR/JOSS employee who works in the International Activities Office at the NWS in Washington, DC, in support of National Weather Services of Central America.

The training will focus on integrating WRF output with GOES satellite imagery in support of generation of timely weather warnings. Unidata's participation in this workshop is being funded by USAID.

Prior Activities

Unidata's Africa-related international outreach activities have largely focused on its role in the UCAR Africa Initiative (AI) which officially ended on April 15. The UPC is currently involved in transfer of technologies developed during the UCAR Africa Initiative project to the African Centre of Meteorological Application for Development (ACMAD) which is located in Niamey, Niger. The following are some highlights of the UPC's involvement in the UCAR Africa Initiative:

- The IDV was used to generate displays of forecast relative humidity that is created using TIGGE ensemble data from ECMWF (via the NCAR/CISL TIGGE repository). The 2013 products and data being

made available in the RAMADDA instance on motherlode.ucar.edu can be found in:

Motherlode Data Server

RAMADDA Data Repository

Projects -> Africa Initiative -> Data -> 2013 Prediction Exercise
-> Areal Coverage
-> Timeseries
-> TIGGE Model Ensembles

- areal distribution of the 50% quantile for RH (which means that each point in the RH field has a 50% probability of being that value or less).

  These products have been stored as animated GIFs, individual frames of the animated GIFs, and week 1 and 2 averages.

- probe timeseries plots of QC25, QC50, and QC75 fields for districts (a district is a subdivision of a region which is like a U.S. state) in a select set of countries (Benin, Burkina Faso, Cote D'Ivoire, Nigeria, Senegal, Tchad, and Togo) that are located in the meningitis belt (which is roughly the Sahel) in Africa.

  The locations for the probe time series plots are determined by meningitis ***attack rates*** (number of new cases per week normalized by population) that are reported by the national health service of each country.

- Also made available were the RH and quantile regression fields (in netCDF format) that are created from ECMWF ensembles that are part of TIGGE.

Every other Thursday UCAR/NCAR AI team members (Tom Hopson NCAR/RAL, Raj Pandya formerly of UCAR/Spark, and/or Arnaud Dumont NCAR/RAL) participated in conference calls with WHO, U Lancaster, and African nation stake holders to discuss the forecast of meningitis cases for the upcoming 1-4 weeks (focusing on the next and second weeks).

- The final stage of AI work in Unidata included automating the generation of display products upon receipt of a new forecast file produced in RAL (by Tom Hopson).

  There were a number of challenges that had to be overcome to automate the product generation process:

  - The programmatic use of the IDV

    Yuan was very helpful in making changes/additions to the IDV to enable this. Don Murray has also been contributing expertise to help Yuan in his efforts.

  - Use of RAMADDA to serve display products to the African Decision Information System (ADIS) interface that Arnaud Dumont (NCAR/RAL) created for the project.

    Jeff McWhirter (NASA, UNAVCO) readily implemented enhancements to RAMADDA for this task.

  - Scraping human-generated documents to get the list of districts for which probe time series plots should be generated.

    The issue is that the sort of information being made available to folks reading MS excel spreadsheets or MS word documents needs to be turned into machine-readable documents that can be used in the product generation workflow.

- The UCAR African Initiative transfer of technologies developed during the Google-funded Meningitis project to the African Centre of Meteorological Application for Development (ACMAD) continues at a low level.

Other activities of note:

- Data from UCAR GOES East/West ingest systems continue to be routinely accessed by international

users in North, Central and South America using McIDAS-X, IDV, and McIDAS-V.

- Use of Unidata tools, especially netCDF, the IDV and GEMPAK, continues to grow internationally.
- IDD-Brazil continues to deliver data via the LDM in Africa.

Updated: February 24, 2015

IDD and NOAAPort

Status Report: October 2014 - March 2015

Mike Schmidt, Jeff Weber, Tom Yoksas

Strategic Focus Areas

The *IDD/NOAAPort* group's work supports the following Unidata funding proposal focus areas:

1. Enable widespread, efficient access to geoscience data
   A project like the IDD demonstrates how sites can employ the LDM to move data in their own environments.

2. Develop and provide open-source tools for effective use of geoscience data
   The IDD is powered by the Unidata LDM-6 which is made freely available to all. The Unidata NOAAPort ingest package is being used by a variety of university and non-university community members. Both the LDM and NOAAPort ingest packages are being bundled by Raytheon in AWIPS-II.

3. Provide cyberinfrastructure leadership in data discovery, access, and use
   The community-driven IDDs provide push data services to users an ever increasing community of global educators and researchers

4. Build, support, and advocate for the diverse geoscience community
   Providing access to data in real-time is a fundamental Unidata activity.
   The IDD-Brasil, the South American peer of the North American IDD operated by the UPC, is helping to extend real-time data delivery outside of the U.S. to countries in South America and Africa. The Universidad de Costa Rica is experimenting with relaying data received in the IDD to Colombia.

Activities Since the Last Status Report

Internet Data Distribution (IDD)

- The UPC has been evaluating ingest and relay of the 0.25 degree GFS data that became operational in NCEP on January 14, 2015. Testing has shown that peak CONDUIT data volumes would increase from about 8 GB/hr to about 21 GB/hr if all forecast hours for the 0.25 degree GFS were added to the current set of products being delivered.

  The increase in data aggregate volume that results from the addition of the 0.25 degree GFS and HRRR data from NOAA/GSD can be seen by comparing the volume on our IDD test leaf node, **lead.unidata.ucar.edu** with that on one of the **idd.unidata.ucar.edu** real server backends shown below:

```
Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume  44003.174 M bytes/hour
Average hourly volume  21199.138 M bytes/hour

Average products per hour     322785 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)           (M byte/hour)   number/hour
CONDUIT                  7756.879    [ 36.591%]    21449.892    91513.640
FSL2                     3718.160    [ 17.539%]     7435.481     6485.200
NGRID                    3523.378    [ 16.620%]     5434.958    24026.240
NEXRAD2                  2878.497    [ 13.578%]     3605.088    53754.120
```

```
FNMOC                1146.801    [  5.410%]     5302.426     3131.360
NEXRAD3              1075.070    [  5.071%]     1262.917    79334.720
HDS                   393.330    [  1.855%]      604.352    19163.000
NOTHER                244.889    [  1.155%]      735.924     1177.960
NIMAGE                152.517    [  0.719%]      247.598      198.840
FNEXRAD                78.796    [  0.372%]       88.861      104.760
GEM                    74.915    [  0.353%]      467.891      803.200
UNIWISC                70.412    [  0.332%]      116.423       46.000
IDS|DDPLUS             55.050    [  0.260%]       66.344    42427.640
EXP                    29.312    [  0.138%]       51.651      304.600
LIGHTNING               1.054    [  0.005%]        1.707      312.920
GPS                     0.079    [  0.000%]        0.405        1.000
```

A working group composed of User Committee members and Unidata staff prepared and distributed a questionaire that is primarily aimed gauging the community's interest in getting the 0.25 degree GFS data added to CONDUIT. Responses to the questionaire will be collected for review during the March User Committee meeting.

Planned Activities

Ongoing Activities

We plan to continue the following activies:

- Unidata continues to receive High Resolution Rapid Refresh (**HRRR**) grids (both 2D and 3D fields) in an LDM/IDD feed from NOAA/GSD and feed these products to a small number (6) of university sites on **hrrr.unidata.ucar.edu**. Since HRRR and ESTOFS data were added to the NOAAPort Satellite Broadcast Network (SBN) in late September, 2014, continuing to relay the HRRR ingested from NOAA/GSD is considered to be of lesser importance and will eventually be discontinued.

  The HRRR is being experimentally served at: http://thredds-jumbo.unidata.ucar.edu/thredds/modelsHrrr.html (.xml for machines)

- Other data sets we are actively exploring with NOAA/GSD/ESRL are:

  - FIM
  - HIWPP

- HRRR and ESTOFS products were added to NOAAPort in late September, 2014. The following TINs announced these additions:

  http://www.nws.noaa.gov/os/notification/tin14-28hrrr-cca.htm
  http://www.nws.noaa.gov/os/notification/tin13-43estofs_noaaport_aaa.htm

  Briefly, these additions are comprised of:

  - HRRR: 81 products, hourly F00-15 each hour. CONUS 2.5km grid184. ~44 GB/day
  - ESTOFS: 3 products, hourly F00-F180, 00, 06, 12, 18z runs. CONUS 2.5km grid, Puerto Rico 1.25 km grid. ~2 GB/day

  HRRR fields and forecasts times that are not included in the NOAAPort expansion will be evaluated as additions to the CONDUIT IDD datastream.

- The UPC continues to relay FNMOC and CMC data model output directly to the community. FNMOC provides the COAMPS and NAVGEM model output and the CMC provides the GEM model output. Unidata has provided access to these data for the past 8 years, but on a "point-to-point" basis. GEM model output was converted from GRIB1 to GRIB2 in January. The CMC is now relaying output of there new hi-resolution (15 km) GEM model to Unidata.

NOAAPort Data Ingest

- The NOAAPort SBN, which transitioned from DVB-S to DVB-S2 in April/May 2011, was upgraded to support much higher throughput in August, 2014. Ingestion of the broadcast as been working at the UPC since the upgade, but we routinely experienced high numbers of missed frames.

  Comparison of our ingest metrics with other sites running our software (e.g., UW/SSEC, NOAA/GSD, LSU/SRCC, and a Northrup Grumman office in Northern Virginia) strongly suggested

that signal quality was a major contributing factor in the problems were being experienced. We expended **considerable** effort to understand the data ingest problems being experienced. Experimentation demonstrated that that use of an older version of firmware on out Novra S300N receivers (V2R7 the version recommended by the manufacturer for our hardware) would produce errors in the UDP output UDP when the S300Ns were interrogated for status information, and this effect was, in turn, a function of signal quality. This problem was verified by one of the commercial vendors of NOAAPort receipt systems who is currently working with Novra to correct this and other problems being experienced by S300N receivers. It was learned very recently that Raytheon was not aware of this problem, but they were aware of other problems related to the S300N receiver, and they too are working with Novra to get the problems corrected.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.12.6, and v6.12.7 is being readied for release.

- Raytheon bundles a version LDM-6 with AWIPS-II and is actively using Unidata's NOAAPort ingest code at a variety of NOAA offices. Raytheon has provided the UPC code modifications and GRIB table updates needed to support new data to be added to in the NOAAPort expansion. when possible

Relevant IDD Metrics

- Approximately **550** machines at **245** sites are running LDM-6 **and** reporting real time statistics to the UPC. Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

  A number organizations/projects continue use the LDM to move substantial amounts of data that do not report statistics to Unidata: NOAA, NASA, USGS, USACE, Governments of Spain, South Korea, private compaines, etc.).

- IDD toplevel relay node, **idd.unidata.ucar.edu**

  The cluster approach to toplevel IDD relay, has been operational at the UPC since early summer 2005.

  The cluster, described in the June 2005 CommunitE-letter article Unidata's IDD Cluster, routinely relays data to more than 700 downstream connections. Data input to the cluster nodes routinely averages up to 20 GB/hr (~0.5 TB/day); average data output from the entire cluster exceeds 1.3 Gbps (~14 TB/day); peak rates routinely exceed 2.2 Gbps (which would be ~24 TB/day if the rate was sustained).

  The following shows a snapshot by feedtype of the data being received on one real server backend node of the Unidata toplevel IDD relay, **idd.unidata.ucar.edu**.

```
Data Volume Summary for uni19.unidata.ucar.edu

Maximum hourly volume  23982.273 M bytes/hour
Average hourly volume  13282.657 M bytes/hour

Average products per hour     297389 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)               (M byte/hour)   number/hour
CONDUIT                  3653.337    [ 27.505%]     8376.741    73213.320
NGRID                    3523.492    [ 26.527%]     5434.958    24064.880
NEXRAD2                  2863.788    [ 21.560%]     3605.088    53421.200
FNMOC                    1146.801    [  8.634%]     5302.426     3131.360
NEXRAD3                  1073.337    [  8.081%]     1262.917    79176.360
HDS                       393.425    [  2.962%]      604.352    19170.920
NOTHER                    244.889    [  1.844%]      735.924     1177.960
NIMAGE                    151.668    [  1.142%]      247.598      197.920
GEM                        74.915    [  0.564%]      467.891      803.200
IDS|DDPLUS                 54.982    [  0.414%]       66.344    42352.160
UNIWISC                    39.047    [  0.294%]       90.063       19.760
FNEXRAD                    32.550    [  0.245%]       53.054       40.480
EXP                        29.301    [  0.221%]       51.651      304.520
LIGHTNING                   1.047    [  0.008%]        1.707      313.720
GPS                         0.079    [  0.001%]        0.405        1.000
```

  Currently six real server nodes operating in one location on the UCAR campus (in the UCAR co-location facility in FL-2) and two directors comprise idd.unidata.ucar.edu. The cluster approach to IDD relay has been adopted by NOAA/GSD, Penn State and Texas A&M.

Updated: March 2, 2015

The Unidata Program Center is a member of the UCAR Community Programs, is managed by the University Corporation for Atmospheric Research, and is funded by the National Science Foundation.

LDM

Status Report: October 2014 - March 2015

Steve Emmerson, Mike Schmidt, Tom Yoksas

Strategic Focus Areas

The *LDM* group's work supports the following Unidata funding proposal focus areas:

1. **Enable widespread, efficient access to geoscience data**
   The LDM powers the Unidata Internet Data Distribution (IDD) system.
2. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The LDM allows sites to move data in their own environments.
3. **Build, support, and advocate for the diverse geoscience community**
   The LDM is used by US universities and by entities throughout the world.

Activities Since the Last Status Report

Multicast-capable LDM-7

Work on the multicast sending and receiving components of LDM-7 is finished except for any unforeseen additional work revealed by integration testing with the Virtual Circuit Multicast Transport Protocol (VCMTP) layer. The VCMTP layer is in the last stages of development and unit-test and the LDM developer is helping with that effort. LDM-7 has the potential to greatly reduce the bandwidth used by the UPC to distribute data via the Internet Data Distribution (IDD) system.

LDM 6.12.7 released

- Made it easier to run downstream-only LDM instances that don't interfere with an LDM server on the same computer.
- NOAAPORT:
  - Diagnosed and fixed a bug in the GEMPAK GRIB2 library that caused WFO-s to miss some RUC, NAM, and GFS products (among others) when experimental multi-radar/multi-sensor products from NSSL were temporarily added to the NOAAPORT broadcast on February 17th and 18th.
  - Improved performance of finding GRIB2 parameter information by caching GEMPAK parameter tables.
  - Improved realtime performance.
  - Improved error-messages for GRIB2 products with unknown parameters or vertical coordinates.
  - Added entries to the GEMPAK NCEP parameter table.
  - Improved statistics reporting.
- Squashed bugs.

Planned Activities

Ongoing Activities

We plan to continue the following activies:

- Support LDM users

    - Email, phone, etc.
    - Training workshops
- Work on multicast-capable LDM-7
- Incrementally improve the LDM as necessary
- Incorporate additional AWIPS-II-related changes into the LDM
- Update table-driven decoding of GRIB products as necessary

New Activities

Over the next three months, we plan to organize or take part in the following:

- Deploy multicast-capable LDM-7

Over the next twelve months, we plan to organize or take part in the following:

- Deploy multicast-capable LDM-7

Prepared Month year

McIDAS, ldm-mcidas, Satellite Data Ingest

Status Report: October 2014 - March 2015

Tom Yoksas

Strategic Focus Areas

McIDAS activities support the following Unidata funding proposal focus areas:

1. Enable widespread, efficient access to geoscience data
   *McIDAS remains **the** application of choice for the satellite meteorology community.*
   The Advanced Data Distribution Environment (ADDE) componet of McIDAS was the first application offered by Unidata to provide remote, programmatic access to a wide variety of data that is important to the atmospheric science community.

2. Develop and provide open-source tools for effective use of geoscience data *The fifth generation of McIDAS, McIDAS-V, unlike its predecessors, is a fully open source application that is is in wide scale and growing use in the satellite meteorology community.*
   McIDAS ADDE continues to evolve and provide access to increasing volumes of image and non-image data.

3. Provide cyberinfrastructure leadership in data discovery, access, and use
   Concepts articulated in ADDE inspired the development of THREDDS (to address the lack of rich metadata available in ADDE) and RAMADDA. ADDE remains one of the most used data services in the Unidata suite serving over 3 TB per month.

4. Build, support, and advocate for the diverse geoscience community
   McIDAS is sought for use by those interested in satellite meteorology worldwide.

Activities Since the Last Status Report

Unidata McIDAS-X/XCD is in mostly in maintenance mode. Aside from routine updates/bugfixes to existing code and tables, the main thrust of development is to add indexing to ADDE datasets to speed access into large datasets.

Prior Activities

- Unidata McIDAS version 2014 includes all SSEC versions up to and including the current release, v2014.1 and Unidata updates and bugfixes. Changes to Unidata McIDAS continue to be made through an **addendum** process. Unidata McIDAS-X/XCD was released in September, 2014.

Ongoing Activities

- SSEC McIDAS Advisor Committee (MAC)

  The UPC (Yoksas, Ho) continues to participate as the Unidata representative to the McIDAS Advisory Committee (MAC) that is operated by SSEC.

  The MAC was assembled by UW/SSEC to advise SSEC on McIDAS-X users needs/concerns/desires for development in the next generation McIDAS, McIDAS-V. The MAC was modeled after the Unidata IDV Steering Committee.

- Interest in McIDAS by Non-core Users

  The UPC continues to receive requests for McIDAS from international university users, U.S.

government agencies and other non-traditional Unidata users (e.g., private businesses, etc.). Government agencies and non-traditional Unidata users are referred to UW/SSEC for access to McIDAS; international educational community user requests are granted on a case-by-case basis after they provide a clear statement of their acceptance of the terms of use provided by SSEC.

Planned Activities

Ongoing Activities

Continued support of existing and new community members.

New Activities

Implement indexing for ADDE image datasets to speed up access especially in large and archive datasets. A preliminary design for ADDE image dataset indexing has been made. Investigations for how to integrate the new capbilities in to the suite of existing ADDE servers is in progress.

Add support for new types of data when they become available, otherwise McIDAS-X support is in maintenance mode.

Relevant Metrics

- Internet2 (I2) bandwidth usage by the McIDAS ADDE protocol routinely exceeds several TB/week.
- McIDAS Inquiry Metrics

ldm-mcidas Decoders Activities

Development

ldm-mcidas releases are made when needed to support changes in software development and operating system environments. **ldm-mcidas** v2012 was released at the end of September, 2012. This release addressed building on newer OS versions.

Geostationary Satellite Data Ingest

Unidata continues to ingest GOES-East and GOES-West imager data at the UCAR Foothills Lab campus in Boulder. GOES-South (GOES-South America) was decommissioned on August 16, 2013, and there appears to be no current plans for repurposing an existing GOES platform for South American surveillance.

- Direct, programmatic access to real-time GOES-East (GOES-13) data via McIDAS ADDE is being downloaded an average of approx. 1.8 TB/month.
- Direct, programmatic access to real-time GOES-West (GOES-15) data via McIDAS ADDE is being downloaded an average of approx.1.5 TB/month.

Planned Activities

Ongoing Activities

Continued ingest and serving of GOES-East and GOES-West imagery. This effort sporadic requires maintenance of the satellite ingest and computer data equipment.

New/future Activities

Repurpose former USAN dish at Mesa Lab to operation as a remotely controllable ingester for any of the GOES platforms. This is a moderately low priority activity.

Proposed Activities

Begin planning for the resources it will take to ingest and disseminate GOES-R data (which is currently scheduled to become available in second quarter 2016). This activity will proceed with cooperation/coordination of NCAR/RAL, NCAR/EOL and NOAA. A draft DRAFT Executive Summary and Budget (i.e., a *non-proposal* "proposal") was developed in cooperation with RAL and EOL, and submitted to Steve Goodman who is in NOAA's GOES-R office. This funding effort was resurrected in fiscal year 2015.

Updated: February 24, 2015

NetCDF

Status Report: September 2014 - February 2015

Ward Fisher, Dennis Heimbigner, Russ Rew

Strategic Focus Areas

The *netCDF* group's activities support the following goals described in Unidata's Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   by developing netCDF and related cyberinfrastructure solutions to facilitate local and remote access to scientific data.
2. **Develop and provide open-source tools for effective use of geoscience data**
   by supporting use of netCDF and related technologies for analyzing, integrating, and visualizing multidimensional geoscience data; enabling effective use of very large data sets; and accessing, managing, and sharing collections of heterogeneous data from diverse sources.
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   by developing useful data models, frameworks, and protocols for geoscience data; advancing geoscience data and metadata standards and conventions; and providing information and guidance on emerging cyberinfrastructure trends and technologies.
4. **Build, support, and advocate for the diverse geoscience community**
   by providing expertise in implementing effective data management, conducting training workshops, responding to support questions, maintaining comprehensive documentation, maintaining example programs and files, and keeping online FAQs, best practices, and web site up to date; fostering interactions between community members; and advocating community perspectives at scientific meetings, conferences, and other venues.

Activities Since the Last Status Report

New Features, Performance Enhancements, and Bug Fixes

We use Jira and GitHub tools for C, Fortran, and C++, interfaces to provide transparent feature development, handle performance issues, fix bugs, deploy new releases, and collaborate with other developers. We currently have 91 open issues for netCDF-C, 18 open issues for netCDF-Fortran, and 3 open issues for netCDF-C++. The Unidata CDM/TDS group maintains the netCDF Java interface, also using Jira and GitHub, and we collaborate with external developers in maintaining the Python interface.

- In the netCDF group, progress has been made in the following areas since the last status report:
  - Integrate and test new floating-point compression plug-in technologies for use with netCDF-4
  - Improve ease of building Fortran interface
  - Fix organization of on-line documentation
  - Support continuous integration for development
- Dependencies, challenges, problems, and risks include:
  - Small group of developers for supporting large project
  - Dependency on HDF5, controlled by external group
  - Slow progress in user adoption of netCDF-4 features

Planned Activities

Ongoing Activities

We plan to continue the following activities:

- Provide support to a large world-wide community of netCDF developers and users
- Continue development, maintenance, and testing of source code for multiple language libraries and generic netCDF utility programs
- Improve organization of Doxygen-generated documentation for the netCDF-C and Fortran libraries

New Activities

Over the next three months, we plan to organize or take part in the following:

- Prepare material for the Unidata Python training workshop in July
- Respond to Naval Research Lab patent application for "System and Method for Importing NetCDF Data"
- Incorporate support for 64-bit-everything netCDF format from parallel netCDF project at Argonne and Northwestern
- Transition to new netCDF project head (Ward Fisher replacing Russ Rew)

Over the next twelve months, we plan to organize or take part in the following:

- Submit an abstract for a netCDF update talk at annual AMS meeting
- Deploy a release with compression competitive with GRIB2
- Participate in development of new CF 2.0 conventions for climate and forecast simulation output and observational data in netCDF-4 form
- Continue to encourage and support use of netCDF-4's enhanced data model by third-party developers

Beyond a one-year time frame, we plan to organize or take part in the following:

- Implement DAP-4 client support in netCDF C library
- Provide thread-safety for netCDF C library
- Improve scalability to handle huge datasets and collections

Areas for Committee Feedback

Community Services is requesting your feedback on the following topics:

1. **Are there any HDF5 features that you wish netCDF supported?**
2. **If netCDF compression were better than GRIB, would you still have uses for GRIB?**
3. **Should netCDF be ported to and maintained for any other programming languages or development environments?**

Relevant Metrics

There are currently about 140,500 lines of code in the netCDF C library source.

The Coverity estimate for defect density (the number of defects per thousand lines of code) in the netCDF C library source has been reduced slightly from **0.36** six months ago to **0.35** today. According to Coverity's analysis of over 250 million lines of open source projects that use their analysis tools, the average defect density with 100,000 to 500,000 lines of code is **0.50**.

Google hits reported when searching for a term such as netCDF-4 don't seem very useful over the long term, as the algorithms for quickly estimating the number of web pages containing a specified term or phrase are proprietary and seem to change frequently. However, this metric may be useful at any particular time for comparing popularity among a set of related terms. Currently, Google hits estimates are:

- **528,000** for netCDF-3
- **511,000** for netCDF-4

- **375,000** for HDF5
- **132,000** for GRIB2
- **273,000,000** for "Taylor Swift"

Google Scholar hits, which supposedly count appearances in peer-reviewed scholarly publications, are:

- **213** for netCDF-3
- **312** for netCDF-4
- **5,260** for HDF5
- **428** for GRIB2
- **2,540** for "Taylor Swift"

Google Patent hits, computed by searching both filed and published patent applications, are:

- **1,450** for netCDF-3
- **1,350** for netCDF-4
- **284** for HDF5
- **3** for GRIB2
- **42** for "Taylor Swift"

So, we finally found a metric where netCDF beats Taylor Swift, and by a **large** margin.

Prepared February 2015

Python

Status Report: October 2014 - March 2015

Ryan May, Sean Arms, Julien Chastang, Ward Fisher, Russ Rew, Ben Domenico

Strategic Focus Areas

Python activity at Unidata supports the Unidata strategic goals in the following ways:

1. **Enable widespread, efficient access to geoscience data.** Python can facilitate data-proximate computations and analyses through IPython (now Jupyter) Notebook technology. In particular, IPython Notebook web servers can be co-located to the data source for analysis and visualization through web browsers. This capability in turn, reduces the amount of data that must travel across computing networks. There are also external providers such as Wakari and coLaboratory that help to promote the use of this technology as a cloud service.
2. **Develop and provide open-source tools for effective use of geoscience data.** Our current and forthcoming efforts in the Python arena will facilitate analysis of geoscience data. This goal will be achieved by continuing to develop Python APIs tailored to Unidata technologies. For the summer 2013 Unidata training workshop, we developed an API to facilitate data access from a THREDDS data server. This effort was later encapsulated with the new pyUDL (a collection of Python utilities for interacting with Unidata technologies) project. Moreover, Python technology coupled with HTML5 IPython Notebook technology has the potential to address "very large datasets" problems. In particular, an IPython Notebook can be theoretically co-located to the data source and accessed via a web browser thereby allowing geoscience professionals to analyze data where the data reside without having to move large amounts of information across networks. This concept fits nicely with the "Unidata in the cloud" vision. Lastly, as a general purpose programming language, Python has the capability to analyze and visualize diverse data in one environment through numerous, well-maintained open-source APIs.
3. **Provide cyberinfrastructure leadership in data discovery, access, and use.** The TDS catalog crawling capabilities found in pyUDL will facilitate access to data remotely served by the Unidata TDS, as well as other TDS instances around the world. The desired goal of pyCDM is to construct a geoscience focused data model in Python, based heavily on the netCDF-Java implementation of the Common Data Model (CDM). pyCDM is anticipated to provide a simple, pythonic API to the higher level functionality of the FeatureType layer of the CDM.
4. **Build, support, and advocate for the diverse geoscience community.** Based on grassroots interest from the geoscience community, Unidata, as part of its annual training workshop, will host a three day session to explore “Python with Unidata technology”. Also, to try to help the use of NetCDF in Python, Unidata has promoted Jeff Whittaker’s NetCDF4-python project, including hosting its repository under Unidata’s GitHub account.

Activities Since the Last Status Report

- Users’ Workshop
  - 21 people attended the Python Users’ workshop in October 2014
  - Takeaways:
    - Cartopy worked out well
    - Anaconda worked out very well
- CDMRemote
  - Prototyped Python implementation of THREDDS’ CDMRemote protocol to provide OPeNDAP-like access to data
  - Pure Python implementation facilitates remote data access from cloud-based services using Wakari and/or the IPython Notebook
  - Prototype sets stage for additional development of protocol to expose more details and semantic information available from the CDM in netCDF-java

- WAVE
  - Tech demonstration of using client-side javascript and WebGL in the browser to talk to IPython notebook-based server
  - Explores cloud-based visualization of data in the web browser using open standards
  - Utilizes existing Python base for data access (from THREDDS server) and analysis
  - Presented at 2015 AMS Annual Meeting in Python symposium
- Matplotlib support
  - Worked variety of support issues with Matplotlib's animation support, which is popular with the community
- Unidata Python on Docker
  - We have a prototype Docker image for Unidata and Python related technologies. See Unidata cloud status report for more information.
- 2015 Unidata Summer Python Training Workshop
  - Expanding the Python workshop to 3 days.
- Discussion with UKMET
  - John Caron met with the team from the UKMet office responsible for development of the IRIS package. They seem very keen for collaboration and an invitation was extended for some developers to come here after the 2015 SciPy conference.
  - Risk: Any collaborative activities will be band-limited by staffing resources available to commit to Python development

Planned Activities

Ongoing Activities

We plan to continue the following activies:

- Continue planning for the Unidata Summer Python Training Workshop
- As time allows, continue to explore WebGL + IPython Notebook + THREDDS access within WAVE
- IPython Notebook
  - Install IPython Notebook on server to explore data-proximate analysis in Python
  - Looking at IPython community solutions for multi-user management and authentication: tmpnb and Project JuPyteR Hub
- CDMRemote
  - Continue getting all corners of CDMRemote protocol implemented
  - Currently pursuing funding to implement CDM in Python (PyCDM), including support for CDMRemoteFeature
- Continue relevant Matplotlib support

New Activities

- Actively working on contribution to matplotlib to facilitate its use for making station plots--this fills a large missing piece in using matplotlib for day-to-day meteorological analysis
- Open up MetPy as a useful landing point for some Python work: Skew-T, NEXRAD Reader, full station-plots

Areas for Committee Feedback

The Python group is requesting your feedback on the following topics:

1. What are the biggest obstacles that you see to the use of Python with other Unidata technologies, or for use in meteorology in general?
2. Given the limited staffing resources for Python activities, how would you order in priority the various on-going and new activities listed above?

Relevant Metrics

- Matplotlib Animation Support
  - 3 Pull Requests reviewed and committed
  - 4 Issues supported and closed
- NetCDF4-Python

36 Issues opened (28 closed)
- 22 Pull Requests (all closed)
- 99% Jeff Whittaker (and Stephan Hoyer)

Prepared March 2015

Rosetta

Status Report: September 2014 - March 2015

Sean Arms, Jen Oxelson, Jeff Weber

Strategic Focus Areas

Community Services supports the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   The initial goal of Rosetta is to transform unstructured ASCII data files into the netCDF format; once in this format, standard tools, such as the THREDDS Data Server, IDV, Python, and other analysis packages, can take advantage of these datasets with relative ease.
2. **Develop and provide open-source tools for effective use of geoscience data**
   Although the primary goal of Rosetta is to get data into the netCDF format, the transformation process does not stop there. The Rosetta group realizes that not everyone knows how to work with netCDF files, and may feel more comfortable working with other formats. Therefore, Rosetta includes the ability to transform from one format to another (e.g. netCDF to .xls), thereby reducing data friction.
3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   Metadata contained in netCDF format file (no longer locked away in a separate README file) can be automatically extracted, facilitating the discovery of data in these files. Additionally, the Rosetta development plan includes the creation of a standard ASCII and spreadsheet representations of the CF-1.6 DSGs.
4. **Build, support, and advocate for the diverse geoscience community**
   Promote the use of standard formats in the dissemination of data, while allowing flexibility to transform into other formats, as needed, to enable users to "do science". For commonly used formats, such as User Defined ASCII format or an unstructured spreadsheet, create and advocate for the use of a standard representations based on the CF-1.6 DSGs.

ACADIS Project

The ACADIS project is finishing its final year of funding. Current activities focus on NSF panel review recommendations, as well as extending the usefulness of Rosetta in context of arctic datasets. While the ACADIS project is winding down, development on Rosetta will continue into the future.

Activities Since the Last Status Report

Basic Documentation

Transitioned to using Doxygen for user and developer documentation.

- Progress has been made on:
  - Transition of old documentation from Sphinx to Doxygen
  - Extension of documentation
  - http://www.unidata.ucar.edu/software/rosetta/dox/html/index.html
- Dependencies, challenges, problems, and risks include:
  - Duplication of documentation effort is a risk, as ACADIS requests documentation specific to the project. However, most of these requests are centered around â€œbrandingâ€ , and can easily be handled by Doxygenâ€™s use of CSS.

Accomplishments of Note

- Added the ability to publish converted files directly to RAMADDA and the ACADIS Gateway
- Live instance of Rosetta hosted at Unidata for testing
- Released the Rosetta source code on github
- Started using Coverity static analysis on the Rosetta source code

Planned Activities

Ongoing Activities

While the ACADIS project is winding down, Unidata plans to continue on the following lines of development:

- Increase the number of CF-1.6 discrete sampling geometries handled by Rosetta. For example, this will enable Rosetta to transform data from moving platforms and profiler data. A specific arctic related dataset would be an observation tower on a drifting iceberg.
- Solicit examples from the community (*hint, hint...that's you guys!*)

New Activities

Over the next three months, we plan to organize or take part in the following:

- Investigate csv and xls(x) representations of the CF-1.6 Discrete Sampling Geometries
- Create infrastructure to collect use metrics for Rosetta
- Transition from the Maven build system to Gradle
- Continue documentation efforts, including the creation of screencasts for User documentation

Over the next twelve months, we plan to organize or take part in the following:

- Enable Desktop (local) use of Rosetta
- Incorporate TDS capabilities into Rosetta, allowing for TDS services (like point subsetting of grids) to easily be applied to local files

Beyond a one-year timeframe, we plan to organize or take part in the following:

- We would like to extend Rosetta such that it can be used as a plugin for the THREDDS Data Server. One of the goals of this plugin would be to enable Rosetta to publish files into THREDDS Data Servers (TDS) as well as automatically generate the THREDDS Configure Catalogs needed to serve out the newly translated datasets.

Areas for Committee Feedback

Community Services is requesting your feedback on the following topics:

1. We would love your input as to where our priorities should be in terms of these New Activities. Let's chat! And, yes, **please...send example ASCII data**

Relevant Metrics

We've received a handful of support questions regarding the availability of Rosetta, as well as requests for demonstrations.

Prepared February 2015

User Support

Status Report: October 2014 - March 2015

Tom Yoksas, Jen Oxelson, UPC Staff

Strategic Focus Areas

The Support group's work supports the following Unidata funding proposal focus areas:

1. Enable widespread, efficient access to geoscience data
   Unidata User Support enables access to geoscience data by supporting the use of tools created and/or supported by the UPC.

2. Build, support, and advocate for the diverse geoscience community
   The user support provided by the UPC is recognized throughout the atmospheric science community. Unidata's outreach efforts are routinely called out in surveys of the NCAR/UCAR community.

Activities Since the Last Status Report

Training

- The UPC hosted its annual training workshop series from October 21 - November 6, 2014.
- The UPC hosted a special 4 day training for representatives from the Taiwanese Civil Aviation Authority (CAA) from September 8 - September 11, 2014. This activity was funded by NCAR/RAL.

Relevant User Support Metrics

Since January 26, 2006 over 44250 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system.

Support by Category

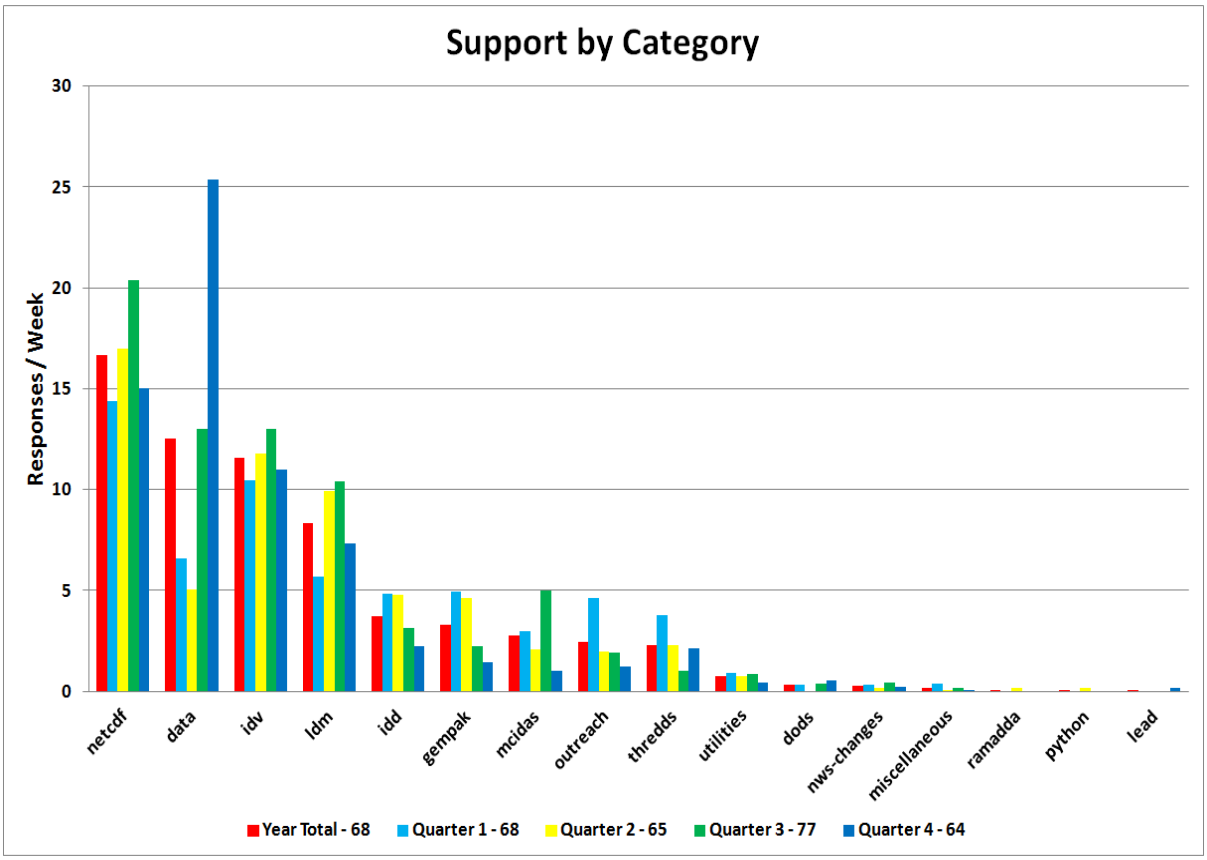


Fig. 1: Above are histograms that portray the number of Unidata email responses for categories of support for a one year period ending February 25, 2015. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period is depicted from oldest to newest from left to right. The number of responses has been normalized to weekly averages so that the support load over the various periods can be easily compared.

Individual support activities included in the categories depicted above are listed in the following table.

Category	Packages, Groups, and Lists
data	casestudies, casestudies-list, conduit, craft, craft-ty, craft-nws, datastream, difax, eumetsat, level2, level2-ty, noaaport, noaaport-ty, noaaportldm
dods	dods, dods-core, dods-list, dods-tech, dods-mlgui-tc, dods-pm, dods-tac, dods-team, opendap, opendap-core, opendap.forward, opendap-list, opendap-tech
gempak	gempak, gembud-list, gempak-ty, awips-ty
idd	cluster, gozer, idd, idd-antarctic-ty, idd-brasil, idd-brazil-ty, idd-caribe-ty, idd-inject, idd-status, scoop, scoop-ty, suominet, tigge, tigge-ty, venezuela-ty
idv	idv, idvlist, idvsteering, java-gui, metapps, visad, visad-list, visad-renderer
ldm	ldm, ldm-users-list
lead	lead, leadusers
mcidas	mcdevelop, mcdevelop-ty, mcidas, mcidas-list, mcidas-ty
miscellaneous	esupport, fxlinux, misc, license, network, notrack, platforms, wxp wxp-lis
	data-models, libcf, ncml, netcdf, netcdf-miss, netcdfgroup-list,

netcdf	netcdf-hdf, netcdf-hdf-list, netcdf-java, netcdf-perl
nws-changes	nws-changes
outreach	agu-ty, announce, argentina-ty, barbados-ty, brazil-ty, cathalac-ty, chile-ty, costarica-ty, mexico-ty, support-ty, cbmet-ty, community-list, eletter, egrants, eumetsat-ty, external, iai-ty, international-ty, joss-ty, k12-list, korea-ty, meteoforum-ty, unidata, workshop
python	python
ramadda	ramadda
thredds	java-dev, java-dev-list, thredds
utilities	decoders, ldm-mcidas, udunits

Support by Topic

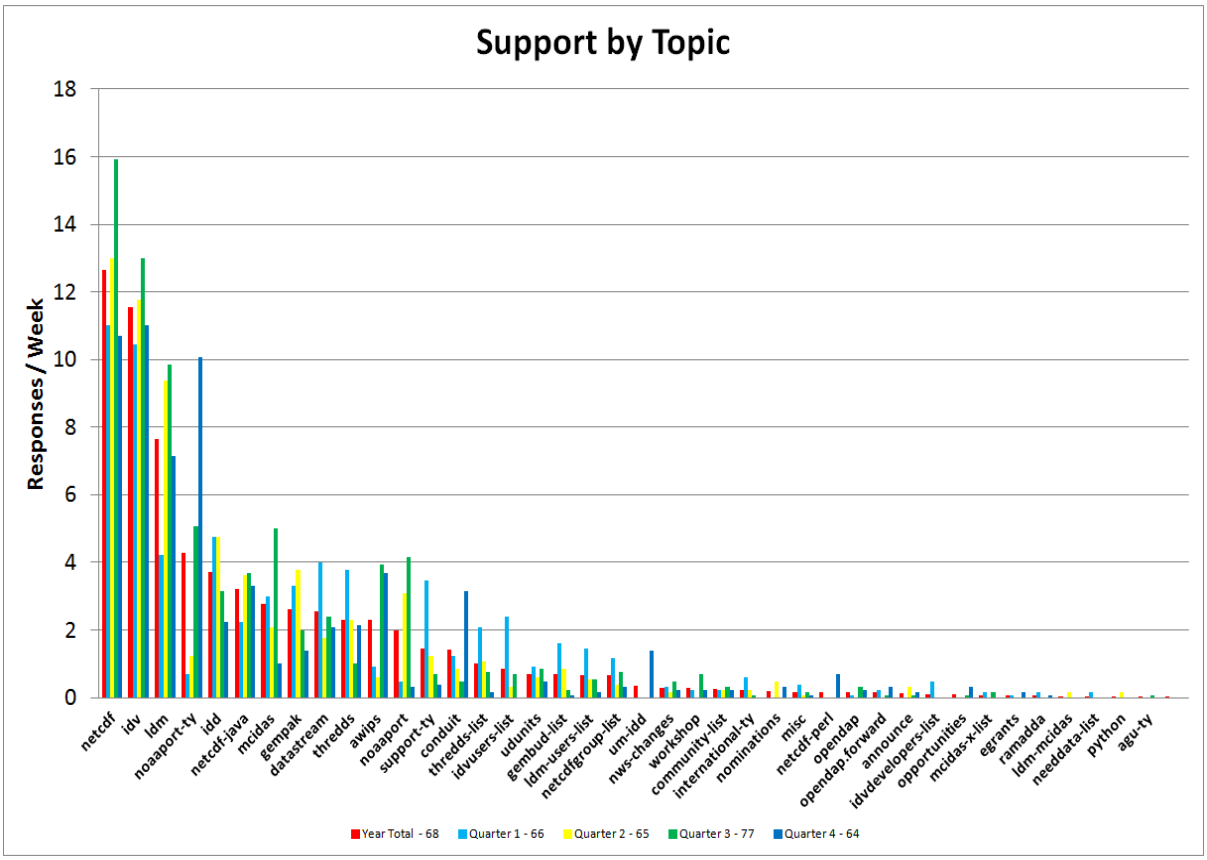


Fig. 2: Above are histograms that portray the number of Unidata email responses for individual topics of support for a one year period ending February 25, 2015. The histograms are arranged by yearly activity averages with the highest on the left and lowest on the right. Each quarter year within the period is depicted from oldest to newest from left to right. The number of responses has been normalized to weekly averages so that the support load over the various periods can be easily compared.

Comments

- The total support provided by the UPC remains high, and yearly totals have been relatively constant for the past two years. Overall support activities vary by somewhat by quarter. Spikes in support for individual packages is largely correlated with the release of new distributions, and, for the IDV in particular, jumps after training workshops.
- Support for netCDF continues to be substantial, and is understandable given the **large** number of users of the package worldwide.
- The IDV support load is second only to that for netCDF; no large increases have been seen over the past 6 months.
- Support for netcdf-java continues to grow steadily.

Support for the legacy visualization packages GEMPAK and McIDAS continues to be substantial.
- Support for LDM, IDD, and data continues at a high level and shows some variability throughout the year.
- Taken as a whole, the support required for visualization packages (GEMPAK, IDV, and McIDAS) is comparable to the support related to data reception (LDM, IDD, noaaport).
- The high numbers for outreach reflect the high level of activity in a variety of activities including organizing sessions at various national meetings.

Notes

These numbers and conclusions should not be taken too literally, for several reasons:

- For some packages, multiple responses in the same thread may be bundled into a single archived email. Other packages have each response in a thread counted separately.
- After a new release of software, there may be a flurry of the same or similar questions, which can be answered in separate emails or in a single mailing list posting.
- The graph primarily represents support of end users and site administrators, not developers. Support for non-Unidata developers in projects such as THREDDS, IDV, GEMPAK, and McIDAS requires significant resources, but is difficult to assess.
- Not all support records were indexable for this report. Given this, the above numbers are an ***underestimate*** of the actual support being provided by the UPC.

Additional User Support Inquiry Metrics

Updated: February 25, 2015

THREDDS

Status Report: September 2014 - March 2015

John Caron, Ethan Davis, Dennis Heimbigner, Sean Arms, Christian Ward-Garrison, and Ryan May

Strategic Focus Areas

The THREDDS group's work supports the following Unidata funding proposal focus areas:

1. Enable widespread, efficient access to geoscience data
   The work of the THREDDS group is comprised of two main areas: the THREDDS Data Server (TDS) and the Common Data Model (CDM) / netCDF-Java library. The TDS provides catalog and data access services for scientific data using OPeNDAP, OGC WCS and WMS, HTTP, and other remote data access protocols. The CDM provides data access through the netCDF-Java API to a variety of data formats (e.g., netCDF, HDF, GRIB). Layered above the basic data access, the CDM uses the metadata contained in datasets to provide a higher-level interface to geoscience specific features of datasets, in particular, providing geolocation and data subsetting in coordinate space. The CDM also provides the foundations for all the services made available through the TDS.

   The data available from the IDD is a driving force on both the TDS and netCDF-Java development. The ability to read all the IDD data through the netCDF-Java library allows the TDS to serve that data and provide services on/for that data.
2. Develop and provide open-source tools for effective use of geoscience data
   Unidata's Integrated Data Viewer (IDV) depends on the netCDF-java library for access to local data, and on the THREDDS Data Server (TDS) for remote access to IDD data. At the same time, the CDM depends on the IDV to validate and test CDM software. Many other tools build on the CDM / netCDF-Java library (eg ERDDAP, Panoply, VERDI, etc) and on the TDS (ESGF, LAS, ncWMS, MyOcean, etc).
3. Provide cyberinfrastructure leadership in data discovery, access, and use
   The Common Data Model (CDM) / netCDF-Java library is one of the few general-purpose implementations of the CF (Climate and Forecast) metadata standards. Current active efforts in CF that we are involved with include use of the extended netCDF-4 data model (CF 2.0) and for point data (Discrete Sampling Geometry CF-DSG).

   The TDS has pioneered the integration of Open Geospatial Consortium (OGC) protocols into the earth science communities. Strong international collaborations have resulted in WCS and WMS services as part of the TDS.

   The CDM and TDS are widely used implementations of the OPeNDAP DAP2 data access protocol. Unidata has worked with the OPeNDAP group to design, develop, and implement a new version of the DAP specification, DAP4, which is now available in the TDS server and the netCDF-Java client software stack.

Release Status of CDM and TDS

- CDM 4.5.4 and TDS 4.3.23 are the current stable releases.
- TDS 4.5 is still beta, since not all features are completely stable, eg FMRC
- TDS 4.5 is running on Unidata's TDS server (thredds.ucar.edu)
- TDS 4.6 is available as a Beta release
- TDS 4.6 is running on Unidata's TDS development server (thredds-dev.unidata.ucar.edu)

Activities Since the Last Status Report

Collaboration Activities

- Active non-core awards that involve THREDDS development:

- EarthCube Building Blocks award: "Integrating Discrete and Continuous Data" with Univ of Texas, Austin and others. Period of performances: Oct 2013 - Sept 2015.
- EarthCube Building Blocks award: "Specifying and Implementing ODSIP, A Data-Service Invocation Protocol" with OPeNDAP, Inc.
- EarthCube Building Blocks award: "Deploying Web Services Across Multiple Science Domains" with IRIS, UNAVCO, and others. Period of performances: Oct 2013 - Sept 2015.
- NASA ROSES ACCESS award: "High Performance Multidisciplinary Open Standard Data Services to Serve Terrestrial Environmental Modeling" with USGS CIDA. Period of performances: 1 Jan 2014 - 31 Dec 2015.

- Two NASA ROSES AIST proposals that involve THREDDS development were not funded:

- one with SSEC on GPM data
- the second with PO.DAAC/JPL on polar orbiting satellite data.

CDM / TDS 4.5 Development details

- OPeNDAP 4.0 protocol TDS server (beta)
- Improvements to netCDF-4 writing, including

- writing CF 1.6 DSG (Discrete Sampling Geometry) files.
- default and user-specified chunking strategies

- Improvements to the TDS for handling of GRIB files, including

- two time dimension view for forecast models.
- scaling to large collections
- many bug fixes

- Improvements to the TDS for handling collections of point, station, and sounding data

- Point Feature Collections use the NCSS interface.
- Automatic metadata extraction into the THREDDS catalog

- NcML logicalReduce allows you to remove length 1 dimensions
- Add WaterML as output format from NCSS point service
- Ugrade CDM/TDS use of HttpClient from 3.x to 4.x (3.x is end-of-life)
- Require Java 7 and Tomcat 7 (start using java.nio.file)
- Updated ncWms and ncIso to latest version

Ongoing Activities

International Standards Development

- Track and participate in OGC MetOcean Working Group discussion.
- Track and participate in OGC WCS development.
- Track and participate in OGC CF-netCDF development.
- Track and participate in CF Standards development

Cloud Activities

- Instance of the TDS, a clone of http://thredds.ucar.edu, is up and running in the Microsoft Azure Cloud
- Investigating use of Docker for deployment of new TDS instances

Collaboration Activities

- Continuing work with NCAR/CISL DSS group to make their large GRIB model collections available using the new TDS 4.6 GRIB Feature Collections.
- Working with US IOOS to distribute the ncSOS plug-in with TDS (4.6 or 5.0).

- Working with Earth System Grid Framework (ESGF) to upgrade TDS functionality for serving CMIPx data.

Planned Activities

CDM / TDS 4.6.X Development

- Grid Feature Collection version 3

- scaling to very large collections (again!)
- make forecast times available on Best datasets
- best / best complete / analysis as separate datasets

- Grid Feature Collection version 3

- scaling to very large collections (again!)
- make forecast times available on Best datasets
- best / best complete / analysis as separate datasets

- File caching improvements
- Rewrite of client side catalog processing
- Add WRF initialization service to NCSS
- Improve the TDS Radar Server services and harmonize it with the NCSS interface.
- Provide NCSS access for all CF-DSG feature types
- ncSOS plug-in

CDM / TDS 5.0 Development

- Rewrite of server side catalog processing
- Improvements for ESGF use of TDS
- Catalog caching improvements
- Auto tracking of catalog changes
- Simplify process for adding new TDS web services. Document simplified process so third-party development groups can more easily add new web services to the TDS.
- Improvements on GRID / SWATH / IMAGE data types in the CDM

Areas for Committee Feedback

- We have requested community input on GRIB collections “Best” Datasets here. Any committee input would be appreciated.
- We are expecting to move to using Java 8 soon, as Java 7 will be at End-of-Life in April 2015. Due to security concerns, we have to push our community to adopt Java 8, really for their own good, especially on the server. Thoughts?
- We would like to restructure of thredds catalog layout. This would not affect IDV bundles, but would would change the way you would explore thredds.ucar.edu. Please see the proposed new layout here, and give us feedback!

Unidata Outreach

Status Report: February 2015

Ben Domenico

Strategic Focus Areas

Unidata Outreach supports the following goals described in Unidata Strategic Plan:

1. Enable widespread, efficient access to geoscience data
2. Develop and provide open-source tools for effective use of geoscience data
3. Provide cyberinfrastructure leadership in data discovery, access, and use
4. Build, support, and advocate for the diverse geoscience community

Below are a few excerpts from the current Unidata Strategic Plan that highlight the importance of the outreach activities summarized in this status update.

- ...to build infrastructure that makes it easy to integrate and use data from disparate geoscience disciplines

  Data formats like netCDF, together with community-based data standards like the Climate and Forecast metadata convention and the Common Data Model are enhancing the widespread usability and interoperability of scientific datasets.

- ...advance geoscience data and metadata standards and conventions
- ..our experience shows us that robust solutions arise from community and collaborative efforts
- ...close partnerships and collaboration with geoscience data providers, tool developers, and other stakeholders, and the informed guidance of our governing committees

Activities Since the Last Status Report

This work involves contributing to and formalizing community and formal international data system standards.

- Unidata Outreach negotiated a Memo of Understanding has been negotiated between UCAR and the Open Geospatial Consortium
  - Annual membership and Technical Committee meeting registration fees have been waived
  - Unidata hosts the OGC Technical Committee meetings in June 2015
- Unidata Outreach represented UCAR/Unidata at the OGC Technical Committee meetings in Calgary and Tokyo
- The NetCDF GML encoding standard has been approved for public comment by the OGC Architecture Board
- The OGC has asked Unidata to participate in a new initiative related to Application Programming Interface standards (as opposed to the usual OGC Web Services standards.)

CUAHSI Hydrological Information Systems (HIS) Standing Committee

The CUAHSI HIS Standing Committee is roughly equivalent to the Unidata Strategic Advisory Committee (USAC). As Unidata's Policy Committee in the past and more recently with USAC, there are difficult decisions about what to focus on in a era of limited resources.

- Unidata Outreach has servied on the CUASI Hydrological Information System standing committee
- Jeff Weber and Ben Domenico participated in several teleconferences for this group and provided

input on the proposal for the Hydrological Water Data Center and on the responses to reviewers questions

EarthCube Discrete Continuous Building Block Project

The goal is to investigate better ways of connecting gridded (field based) data & time-varying properties of discrete spatial objects (points, lines, polygons). This is to be accomplished by developing a common conceptual model and tools that can map between gridded (netCDF/CF) and vector/time series (WaterML 2.0) data encodings, with visualization of time series parameter values. Develop tools for guiding usage.

This project now is focused on one use case: namely, the National Flood Interoperability Experiment for its final year. Unidata Outreach is coordinating the aspect of the NFIE that involves the NCAR WRF-Hydro model and the group that supports it.

CyberConnector Building Block Project

This project is building an EarthCube building block, called CyberConnector, for facilitating the automatic preparation and feeding of both historic and near-real time Earth Observation (EO) customized data and on-demand derived products into Earth science models (ESMs). CyberConnector will free scientists from the laborious preparation of model inputs and release of model outputs. It will automatically process the EO data into the right products in the right form needed for ESM initialization, validation, and inter-comparison. It can support many different ESMs through its standard interfaces under a unified framework. Unidata's main role in this project will be to assist in the installation of a Unidata IDD/LDM node at the main site at George Mason University.

Planned Activities

Ongoing Activities

We plan to continue the following activies:

- Standards -- OGC TC Meeting participation
- EarthCube projects
- ODIP Phase 2 initiation
- Work on cloud based Python client software development
- Coordination of brokering system integration with TDS in the cloud

Collaborations

- NCAR GIS Program
- Collaboration with ESSI Labs to experiment with their brokering layer in conjunction with THREDDS Data Servers
- UCAR wide representative to OGC Technical Committee
- ODIP Steering Committee
- CUAHSI Hydrological Information System Standing Committee
- EarthCube Discrete/Continuous Building Block
- EarchCube CyberConnector Building Block
- Wakari Cloud-based Collaborative Python Development Environment

New Activities

- ODIP Phase 2 (at a reduced level)

Areas for Committee Feedback

Any suggestions for better approaches to these efforts

Relevant Metrics

Much of this work involves international collaborations. The OGC NetCDF Standards Working Group, for example, has representatives from Italy, Germany, Australia, New Zealand, England, France, Netherlands, Japan, United Arab Emirates, Spain, Greece, China, and Finland

Prepared February 2015