

Agenda: Fall 2023 Joint Committee Meeting

(Times are Mountain Standard Time)

Wednesday, November 1, 2023 (Users Committee)

11:00 – 12:00 New Committee Member Orientation

Meeting Commences After Lunch at 13:00

13:00 – 13:15 Welcome/Introduction (Mohan Ramamurthy/Chair)

13:15 – 13:20 Administrative Items (Chair/Tanya)

13:20 – 13:45 Users Workshop Follow-up Report on Outcomes and Impact (Chair)

13:45 – 15:15 Status Reports Discussion (Committee/Staff)

- Visualization Software/Tools
- Data Access/Formats/Dissemination
- Community Services/Educational Efforts

15:15 – 15:30 BREAK

15:30 – 16:30 Core Proposal Community Hubs and MSI Partnerships Discussion (Mohan)

16:30 – 16:45 Administrative / Wrap Up Day 1 (Review/Discussion, Actions, Spring Meeting Dates)

Thursday, November 2, 2023 (Joint Committee)

09:00 – 09:10 Welcome and Administrative Items (Chair/Tanya)

09:10 – 09:45 Thank you to parting/transitioning members and welcome to new committee members

- Introductions (abbreviated round table) – highlight what doing and excited about
- Committee Chairs Update

09:45 – 10:15 Agency Update (NOAA – Joint Update, Scott Jacobs with William Boll/Carissa Klemmer)

- 10:15 – 10:45 Agency Update (USGS – David Blodgett)
- 10:45 – 11:00 Break
- 11:00 – 11:30 Agency Update (NASA – Doug Newman)
- 11:30 – 12:00 Program Center Project Updates (CyberTraining and Pythia)
- 12:00 – 13:00 LUNCH
- 13:00 – 14:00 Director’s Report (Mohan Ramamurthy)
- 14:00 – 14:30 Joint Committee Topic (AI in ESS Coursework)
- 14:30 – 14:45 BREAK and walk to FL2-1001 (the small seminar room)
- 14:45 – 16:00 Hybrid DeSouza Presentation and Q&A (Andrea Zonca and Jeremy Fischer)
- 16:00 Meeting Adjourn, Handing over of the Users Committee Chair(s) Gavel
- 18:30 Dinner and Collaborative Discussion on the Day’s Proceedings at: [Bacco | Trattoria & Mozzarella Bar](#), 1200 Yarmouth Ave, Boulder, CO 80304 ([map](#))

Friday, November 3, 2023 (Strategic Advisory Committee, Executive Session)

- 09:00 – 09:15 Welcome and Administrative Items (Chair/Tanya)
- 09:15 – 10:15 Core Proposal – Unidata Reimagined: New Approaches to Community Data Services (Mohan Ramamurthy)
- 10:15 – 10:30 BREAK
- 10:30 – 11:00 Open Discussion
- 11:00 – 12:00 Budgets and Staffing (Mohan Ramamurthy, Joshua Neidigh)
- 12:00 – 12:15 Wrap Up, Review of Day’s Proceedings / Discussion, Admin (Actions, Spring Dates and Location)
- 12:15- 13:00 All Other Business/Lunch/Meeting Adjourned

Status Report: Users Committee Actions

April 2023- November 2023

Unidata Program Center Staff

Actions from the Previous Meeting (March 30-31 2023)

Action 1

Convene McIDAS-X power users to give guidance for easing transition as Unidata ends support [UPC: MikeZ and TomY]

Result

MikeZ continues to work closely with COD to address any concerns there; of recently there have been none aside from a request for updated McIDAS maps, which he is working on.

MikeZ and Tom recently attended the McIDAS Users' Group (MUG) meeting in Madison Wisconsin. New relationships have been formed between MikeZ and SSEC, and he gave a presentation at the MUG meeting showcasing how and the scale at which McIDAS-X is used at COD. SSEC reinforced that McIDAS-X will be supported through at least the GOES-R series lifetime (2030's), and there are no known data changes on the horizon that would require new features to be added. As long as Unidata maintains its sub award with SSEC, our users will be able to continue to use Unidata's McIDAS-X, with the understanding that support and development have ceased.

There are plans to sunset the mcidas-support email list and write the McIDAS-X Users email list to make some final announcements, including a transition to that email list for any/all inquiries. We have not yet made this announcement.

Action 2

Set up Users Committee virtual session to discuss consolidating data-centric e-mail lists [UPC: TanyaV]

Result

[In Progress]

Action 3

Remove unused NWS-Changes e-mail list [UPC: JenO or DougD]

Result

The NWS-Changes list is no longer active, but the message archives are still available to the

public for viewing.

Action 4

Start e-mail discussion with the committee regarding NEXRAD L3 products that are available but not distributed via NOAAPort. Does the community want any of these added to Unidata's distribution mechanisms? [UPC: MikeZ]

Result

There has been some interest in a number of the NEXRAD L3 products that we would be able to add to the IDD. During internal discussions on how best to proceed with that, we realized this taps into a larger issue concerning the LDM Feed Types, their purposes, and how the UPC could/would/should add data to the IDD in a way that would minimize impact and confusion to users. These additions are on hold until we can form better plans around these larger issues.

Action 5

Evaluate data on CONDUIT that could potentially be delivered via NWSTG-LDM instead; have discussion at the Fall committee meeting. [UPC: MikeZ, TanyaV. NCEP: Margaret Curtis or Anne Myckow]

Result

Little progress has been made on finding data products that could be moved to NWSTG-LDM. However, discussions have recently resumed concerning duplicate NDFD data on the IDD. A course of action has yet to be decided, but the cause of the duplicate data is being investigated on NCEP's side.

Action 6

National Association of Geoscience Teachers (NAGT) have pre-made rubrics where they assess activities. Casey Davenport can get the team in contact with someone at NAGT who might be able to provide insight on their process. [UPC: NicoleC, AnaE Committee: Casey Davenport]

Result

Casey pointed us to NAGT's Teach The Earth project, specifically their submission form for contributions and their peer review process. A similar procedure could be implemented for contributions to the Unidata Science Gateway (USG), however it is still unclear the extent to which the USG team has the capability to review contributions. Furthermore, as we envision a variety of contributions, e.g. tutorials, scripts, and workflows, maintaining those contributions takes extra time and scheduling as their dependent software packages inevitably will fall out of date. Collaboration with Unidata staff and community members may help alleviate some of these problems.

Action 7

Draft DeSouza award announcement and share with committee [UPC: DougD]

Result

While the award announcement is traditionally published after the honorees have had a chance to give their seminar, some introductory text was shared with committee members.

Action 8

Investigate whether DeSouza seminar can be shared with some other UCAR or NCAR seminar series. [UPC: DougD]

Result

Discussed with UCAR multimedia group (who stream/publish UCAR seminars). Most NCAR seminars are focused on the science of a particular lab or program, so we'd have to find a good fit for a specific DeSouza topic. Many seminar series have set schedules/days, so it might be difficult to coordinate with Users Committee meetings.

Question: should we investigate further?

Action 9

To gather future DeSouza nominations, investigate contacting committee alumni, geoscience department chairs, maybe partner with AMS? [UPC: TanyaV or DougD]

Result

[In Progress]

Action 10

Add agenda item to Fall meeting schedule to talk about the role of webinars and other outreach/branding activities as well as status checks for Equipment Awards. [UPC: TanyaV]

Result

Delayed until Spring meeting.

Status Report: AI/ML

April 2023- November 2023

Unidata Program Center Staff

Executive Summary

Unidata ML Staff have been working on educational materials, project pythia cookbooks, and blog posts. The Cybertraining Award from NSF will be a focus moving forward with our partner at MSU Denver for in classroom AI/ML content and materials.

Questions for Immediate Committee Feedback

- Is anyone using the Zarr format (<https://tp/zarr.dev/>) in the classroom or in research?

Activities Since the Last Status Report

- Won the NSF Cybertraining award! Thomas Martin is Co-PI with Nicole Corbin, Unidata's Educational designer.
- Did a virtual visit with University of Oregon to discuss Earth Systems Science Data Science careers and pathways
- Visited Colorado School of Mines to discuss future skills for geologists
- Working with NCAR researchers to get hands-on experience with cutting edge ML models and workflows.
- Office hours still available (sign up here: <https://calendar.app.google/ZsM8dLHLa65eGAr39>)
- Helped with intern activities (stay tuned for NcML improvements from the TDS team)!

New Activities

We are looking for more ways to support our community at large. This might include in-person or remote workshops, 1 on 1 mentoring and tutoring, and online asynchronous resources. Future visits to Texas Tech, and other universities are in early planning stages.

Resource development that has a high priority will be around Zarr, Dask, Kerchunk, and other technologies that allow for large dataset analysis larger than local memory.

The Cybertraining award will be a major focus for the duration of the award (2 years). Always looking for more collaboration opportunities around ML/AI education in the classroom, from guest lectures to content development.

Status Report: AWIPS

April 2023- November 2023

Tiffany Meyer, Shay Carter

Executive Summary

Since the last status report, the team has been developing the Beta, version 20.3.2 of AWIPS, while maintaining our production version of 18.2.1. Both versions have had production and development EDEX systems hosted on the Jetstream2 platform since last summer. The team is excited to announce that we are planning on releasing a production version of 20.3.2 before the end of the year, and fully transitioning away from v18 – the first major upgrade of AWIPS for this team.

Throughout the year, for version 20, we have released significantly simplified and improved installers of CAVE for both MacOS and Windows, as well as a full EDEX installation for CentOS as well as continuing a Centos7 Virtual Machine option (available on both CentOS and Windows).

We have been supporting a publically accessible beta EDEX available at edex-beta.unidata.ucar.edu for v20. For each of the subsequent releases of the beta, we have encouraged users to submit bugs and issues using [our reporting form](#). This has allowed us to track and address many user-identified problems.

Aside from AWIPS development, the team is proud to host our first AWIPS workshop at the 2024 AMS Student Conference in late January. We have also completed more than two consecutive years of releasing [AWIPS Tips blogs](#) every other week.

Questions for Immediate Committee Feedback

Do you use AWIPS in the classroom, if not, what's the biggest obstacle keeping you from doing so?

Activities Since the Last Status Report

AWIPS

Our EDEX servers have been continuously running on the new Jetstream2 platform since our transition last summer, despite numerous issues we've encountered on the new platform. The AWIPS team has worked closely with Unidata IT, Science Gateway coworkers (Ana and Julien) and the Jetstream staff to troubleshoot and resolve issues that arise. We have been maintaining both production v18 and beta v20 EDEX servers which continue to serve real-time

weather and geographic data to [CAVE clients](#) and the [python-awips](#) data access framework API.

Through the use of ancillary EDEX machines we have been able to decouple certain datasets from the main EDEX instance. We take advantage of distributing EDEX workload over three machines: a main EDEX, an ancillary radar EDEX, and an ancillary satellite EDEX. These [distributed architectural concepts](#) of AWIPS allow us to scale EDEX in the cloud to account for the size of incoming data feeds. We have been maintaining 4 of these “EDEX Systems” – a production v18, a backup v18, a production v20, and a backup/development v20.

Since the last status report, we’ve put out several beta releases of v20. Our latest release is [version 20.3.2-0.4](#). A link to all of our AWIPS release notes can be found [here](#). We have also put out a [beta release of v20 python-awips](#).

Version 20 has some major upgrades in terms of dependent software, with the following significant packages:

- Python (from 2.7 to 3.6)
- Java (Oracle Java 1.8 to OpenJDK Java 1.11)
- Eclipse (4.6.3 to 4.17.0)

For the beta, the AWIPS team has developed a new installation process for both MacOS and Windows operating systems. The CAVE application for both platforms now runs in a fully contained environment, meaning no additional downloads and no environment or user variable changes, unlike previous installs. We have also ported this streamlined installation back to our Windows installer for v18. Now we can guarantee consistency when it comes to the versions of Java and Python that are running when the user runs CAVE, all without interfering with the default host system.

Throughout the development cycles of the beta release, we have been collecting user feedback through [this reporting form](#). This method has been an invaluable tool to collect, track, and update users on issues found in version 20 of AWIPS.

Since last summer, the AWIPS team has been successfully maintaining and updating our ISatSS installation which allows us access to the NOAA-created GLM products. This work has been assisted by Brian Rapp and Lee Byerle from NOAA.

As of the beginning of October, Unidata has come to an agreement with Synoptic Data to begin a data-sharing partnership with them. Out of this agreement, AWIPS is now accessing, decoding, and displaying RAWs data. This entire collaboration was prompted from a user request for RAWs data because of its importance in the field of wildfire response and management. The AWIPS team worked closely with our Data Systems team, and Data Engineer, Mike Zuranski, on this project.

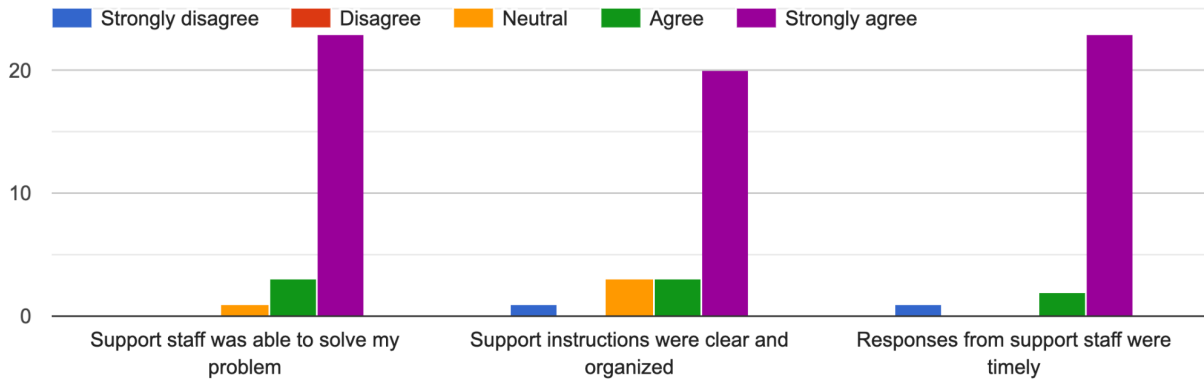
A significant portion of our documentation both for [CAVE, EDEX](#), and [python-awips](#) has been modified for easier understanding and comprehension. When the new version 20 of AWIPS goes live as the production version, we will have all new updates to documentation to make sure it reflects the most current version. We are continuing to update our python-awips example notebooks to follow our new template that contains a helpful table of contents, with consistent subsections across the various example topics.

Our blog series, [AWIPS Tips](#), has successfully been running every other week for over two years now. A current list and breakdown of all the entries is provided on our [documentation website in the Educational Resources page](#). We plan to continue the blog series for the foreseeable future and have several more ideas already planned out for upcoming entries. Announcements for new blog posts are shared through our mailing list (awips2-users@unidata.ucar.edu), and our social media accounts (Facebook, Twitter, LinkedIn, and YouTube when applicable).

We have asynchronous training available for both CAVE and Python-AWIPS on the Educational Resources website. We encourage those who have never used CAVE and those who have but still might be fairly new to the software to [sign up and try out the course for themselves](#). *Learn Python-AWIPS* is tailored for users with entry level Python knowledge, and zero python-awips knowledge. The skills and knowledge gained from the course will hopefully allow the learners to then easily understand and alter our example notebooks to be able to produce customized, useful data plots. We are actively encouraging our University members to [take the learning course](#) themselves, and recommend it to their students as well. We know Python is growing in popularity in the meteorologic field and would like to offer this learning experience as a tool for professors and students to take advantage of.

The AWIPS team still has an active [support evaluation survey](#) that is advertised in our support email signatures. The majority of our feedback has been overwhelmingly positive, and the graphic below is a summary from all responses we've received regarding the quality of service we provide:

Service Quality



Some of the latest open-ended feedback from the support evaluations includes the following:

- *“Both Tiffany and Shay always respond quickly to requests I have made.”*
- *“Shay was incredibly responsive to my suggestions for improved usability in AWiPS. A personal response was provided several times and each of the items were addressed. I'm very impressed with the support I received, indeed, it couldn't be better. Thank you!!”*
- *“To you and the team, AWIPS/CAVE is by far the most useful and mission critical software in use by myself and other IMETS on wildfires, hazardous materials spills and other high impact public safety incidents. I cannot thank you all enough for your work.”*

This next year, part of the AWIPS team will be attending AMS 2024. During the student conference we will be hosting a workshop on AWIPS and how it can be used in the classroom to prepare students for careers at the NWS and how it can benefit your school meteorology program. We will also be presenting a talk about the latest Unidata AWIPS updates.

The team has also received several requests in the recent past inquiring about workshops focused on the EDEX side of AWIPS. Most of these requests have come from either private companies or varying government agencies. The team is considering planning an EDEX workshop for sometime in the future.

Software Releases

Since our last status report we have put out three releases of the beta for [version 20](#). These updates include new, streamlined installers for CAVE on MacOS and Windows, as well as full installations of EDEX available for RHEL7 (CentOS7). Additionally, we have put out a [beta version of python-awips](#) that is also compatible with our edex-beta.unidata.ucar.edu public v20 EDEX.

We are still finalizing the last few functionality and enhancement developments for version 20 before we push it to production and decommission our v18 releases and public EDEX. We are on track to move completely to v20 before the end of the year, and will be reaching out to our University labs with CAVE and EDEX installed to help them transition their machines over the winter break.

Activities Ongoing/In-Progress

AWIPS development activities are constantly ongoing. Currently the following activities are in progress:

- The AWIPS team is looking into building and distributing the new v20 gridslice.so for MacOS CAVE, which will make it as functional as our production v18 CAVE
- The AWIPS team has been maintaining our new Jetstream2 platform with our production and development EDEX servers, with the help of the Science Gateways team.
- The AWIPS team is responding to all AWIPS support questions from the community and striving to provide realistic solutions in a timely manner.
- The AWIPS team is actively updating and refining our online documentation to be as accurate and useful as possible.
- The AWIPS team has maintained a bi-weekly blog series called AWIPS Tips that began on April 7th, 2021 and has been used to highlight useful functionality and fundamentals for CAVE, EDEX, python-awips, and general AWIPS announcements.
- The AWIPS team is actively developing the v20 AWIPS build, and aims to have a production release by the end of the year.
- The AWIPS team is preparing content for the AMS 2024 Student Conference Workshop to give a seminar on how AWIPS is useful at school and can prepare you to enter the workforce.

Future Activities

Future plans are constantly evolving to meet the needs of our users. The AWIPS team is focused on developing and improving the current beta to release a production version of 20 before the end of the year. When this goes live, we will also update our python-awips package to have the latest v20 installation available on Conda and pip. Once we put out the production version of 20, we will be focusing on a new build that will be able to be installed on RHEL8 OS (as well as Rocky8) with newer versions of python. We are also looking forward to the opportunity of hosting an in-person student workshop at AMS 2024. The team is actively participating in conferences, workshops, and virtual message boards (blogs) to expand our user base.

Metrics

Downloads March 1, 2023 - September 30, 2023

AWIPS downloads: 5,192

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

The cloud-based EDEX data server continues to see widespread use and growing adoption. More and more datasets continue to be added to the server as Unidata deploys more decode/ingest nodes.

2. **Providing Useful Tools**

All AWIPS tools (EDEX, CAVE, and python-awips) are freely available, and also incorporate LDM/IDD technology for accessing geoscience data.

3. **Supporting People**

At this juncture, we are providing full technical support with regards to AWIPS for the community.

Prepared *October 2023*

Status Report: Science Gateway and Cloud Computing Activities

April 2023- November 2023

Unidata Program Center Staff

Executive Summary

From April to October 2023, we made significant advancements in our Science Gateway and Cloud Computing offerings and infrastructure. We secured an 8 million SU allocation from the ACCESS program on Jetstream2, marking our largest allocation since 2015. We streamlined the JupyterHub request process, enhanced Docker container projects, and expanded the capabilities of the Unidata demonstration server with GPU support enabling projects in the artificial intelligence and machine learning arena. We've been actively enhancing our AWIPS EDEX server on Jetstream2 to ensure seamless data delivery and to prepare for upcoming infrastructure changes like CentOS 7's End of Life. Collaborations with institutions such as Colorado State University and the NCAR EOL for the LROSE project have been fruitful, resulting in knowledge sharing and technology development. We successfully deployed the Weather Research Forecast (WRF) system on Jetstream2, enabling numerical weather prediction in classroom and research settings. Furthermore, the redesign of the Unidata Science Gateway is underway, with key components including a JupyterHub portal. We also continued to expand on PyAOS JupyterHub offerings and actively participated in several workshops, serving many students. This period saw significant engagement and growth, serving users, expanding technological capabilities, and fostering collaborations.

Questions for Immediate Committee Feedback

1. As we are in the redesign phase of the Unidata Science Gateway, are there specific features or components that the committee would like to see included or prioritized?
2. With our collaborations with institutions like Colorado State University and the NCAR EOL, are there other institutions or organizations that the committee believes we should be partnering with?

Activities Since the Last Status Report

Jetstream2 2023-24 ACCESS Grant Request

Unidata has been awarded 8 million SUs from the ACCESS program in another Jetstream2 cycle to maintain continuous access to essential servers like EDEX, JupyterHub, THREDDS, RAMADDA, and LDM/IDD nodes. This grant allows access to a variety of CPU and GPU virtual machines (VMs) with various configurations. This is our largest allocation since 2015, and a significant increase compared to around 5 million SUs received in 2022-23.

JupyterHub Request Form

Science gateway staff have designed a [JupyterHub request form](#) that includes questions on:

- Purpose of the JupyterHub
- Start and end dates
- Required computational resources (CPU, GPU, RAM)
- Necessary computation environments (e.g., conda)
- Administrator accounts
- Linked GitHub accounts

This form streamlines the process of requesting JupyterHub servers for semester-long use and workshops. On our end, this form allows us to better keep track of not only the tasks that need to be completed, but also gives us an automated, centralized location to gather metrics on the requests we receive.

This form is one step in developing our Science Gateway Re-Imagined project, which, among other things, aims to enhance the user experience of using the Unidata Science Gateway and the resources we offer.

Unidata Docker Container Revamp Project

Reworked the Unidata `tomcat-docker`, `thredds-docker`, `ramadda-docker`, `ldm-docker` projects:

- Streamlined Docker files for conciseness and smaller image size
- Standardizing the structure across all READMEs

Additionally, automation scripts were written to keep these Docker containers consistently updated with the latest versions and security enhancements.

Relaunched `jupyterhub.unidata.ucar.edu` as a GPU enabled Hub

To generate more interest in the previously underutilized demonstration server, it has been upgraded to use GPU machines, enhancing its AI/ML capabilities. Unidata's Thomas Martin and Jeremy Corner, a master's student at NIU under Alex Haberlie, are utilizing this improved Hub for their respective AI/ML projects.

LROSE Collaboration between Colorado State University and NCAR EOL

Unidata science gateway staff have collaborated with Professor Mike Bell's team at Colorado State University and NCAR EOL to help build their science gateway which involves a JupyterHub equipped with LROSE radar meteorological software. We have shared our accumulated expertise in JupyterHub and related technologies with the team.

Weather Research Forecast Model on Jetstream2

Summary

For the first time in Unidata's presence on Jetstream, we have deployed a containerized version of the Weather Research Forecast (WRF) numerical weather prediction system on Jetstream2, providing two different scenarios. This new capability allows for exploration of Numerical Weather Prediction (NWP) models and subsequent analysis and visualization of the output in a data-proximate manner, for example, in a JupyterLab environment.

WRF Navajo Technical University

Unidata is collaborating with the Southwestern Indian Polytechnic Institute and Navajo Technical University to deploy an operational WRF model over the Navajo Nation. This project aims to provide Tribal Nations, and the Tribal Colleges and Universities (TCUs) with the capacity for environmental monitoring in alignment with data sovereignty objectives.

WRF Single Column Model in JupyterHub

In collaboration with Greg Blumberg at Millersville University, Unidata staff have deployed a single-column WRF model in a JupyterHub environment for undergraduate instructional objectives. As a result of this collaboration, Unidata staff will be presenting their procedures and findings at the Science Gateways 2023 Conference, hosted in Pittsburgh, PA on Oct 29 through Nov 1, 2023.

Unidata Science Gateway Re-Imagined

We continue to make progress on the design phase of the Unidata Science Gateway Re-Imagined project as time permits. After collaborating with the redesign team and Unidata management. We have settled on plan "2b" which consists of a redesigned science gateway with the following components:

- JupyterHub portal
 - Accessing customized PyAOS JupyterHub
 - Request new JupyterHub server
- Education Hub
- Data services portal
- Discussion forum
- Web accessible versions of:
 - Integrated Data Viewer
 - CAVE AWIPS visualization client
- Single sign on

Meetings have resumed twice monthly to develop an implementation strategy.

JupyterHub Servers for Summer Workshops, Spring and Fall

Semesters

Unidata is employing our Jetstream2 resource allocation for the benefit of students in the atmospheric science community by providing access to customized JupyterHub servers at an accelerating pace. Unidata tailors these servers to the requirements of the instructors so they can accomplish their Earth Systems Science teaching objectives. Since spring semester of 2023 (encompassing the length of this status report), 606 students at twelve academic institutions and various workshops have used Unidata JupyterHub servers running on Jetstream2.

Notably, we provided JupyterHub resources to:

- 2023 Unidata Users Workshop
- 2023 Summer SOARS Protégé cohort
- “Train the Trainer” workshop (CSU)
- Python Readiness Series workshop series
- I-Guide Summer School
- MetPy for Quantitative Analysis of Meteorological Data workshop
- UCAR Professional Development Workshop Series 7

University of Oklahoma REU Students

Unidata continues to collaborate with Ben Schenkel (OU) to provide data sets via the science gateway RAMADDA server. We also deployed a JupyterHub server so that NSF REU students at OU could access those data for their projects.

Ongoing Activities

NOAA Big Data Program

- Unidata continues to manage the NEXRAD archive in Amazon S3, ensuring that realtime data are successfully delivered to the noaa-nexrad-level2 bucket. LDM is employed to deliver these data.
- Unidata also continues to deliver NEXRAD level 3 products to the unidata-nexrad-level3 bucket, part of the AWS public datasets program.
- TDS on Jetstream2 for level II NEXRAD:
<http://thredds-aws.unidata.ucar.edu/thredds/catalog.html>
- AWS Explorer (Public access):
<https://s3.amazonaws.com/noaa-nexrad-level2/index.html>
- Public Bucket for level II NEXRAD: <https://noaa-nexrad-level2.s3.amazonaws.com>
- Continue to populate the NEXRAD level II archive with real time data.
- Continue to populate new GFS 0.25 degree output and NCEP HRRR output to an S3 bucket for access. We did not place a TDS on this collection as this output is available from our standard sources.

Andrea Zonca Collaboration

Unidata staff continues to collaborate with Andrea Zonca (SDSC/Jetstream2) employing his

port of the "Zero to JupyterHub with Kubernetes" project to OpenStack and Jetstream2. We give Andrea feedback by testing his [instructional blog entries](#) and workflows. When we encounter issues, we submit bug reports via GitHub and work together until the problem is resolved.

Docker Containerization of Unidata Technology

Beyond what we mentioned earlier about improvements in this area, we continue to employ Docker container technology to streamline building, deploying, and running Unidata technology offerings in cloud-based environments. Specifically, we are refining and improving Docker images for the LDM, ADDE, RAMADDA, THREDDS, and AWIPS. In addition, we also maintain a security-hardened Unidata Tomcat container inherited by the RAMADDA and THREDDS containers. Independently, this Tomcat container has gained use in the geoscience community.

AWIPS EDEX in Jetstream2 Cloud

Unidata continues to host our publicly accessible EDEX server on the Jetstream2 cloud platform where we serve real-time AWIPS data to CAVE clients and the python-awips data access framework (DAF) API. The distributed architectural concepts of AWIPS allow us to scale EDEX in the cloud to account for the desired data feed (and size). We continue using Jetstream2 to develop cloud-deployable AWIPS instances as imaged virtual machines (VMI) available to users of OpenStack CLI. Since last summer all EDEX servers have been running Jetstream2. Unfortunately, the service has not been entirely seamless and both the AWIPS team and the Science Gateways team have spent significant time troubleshooting and repairing machines to keep our servers operational. In addition, we have created a custom CentOS 7 image for deployment on Jetstream2 on which to provision new EDEX machines before CentOS 7's End of Life on June 30, 2024. Before that time EDEX will be transitioned to be deployable on Rocky or another RHEL derivative.

EDEX is designed so different components can be run across separate virtual machines (VMs) to improve efficiency and reduce latency. Our current design makes use of three VMs: one large instance to process most of the data and run all of the EDEX services including all requests, and two other ancillary machines which are smaller instances used to ingest and decode radar and satellite data individually.

We are currently supporting 4 sets of servers as described above: two sets are running our v18 software (production version of AWIPS), and two sets are running our new beta v20 software. The live backups allow us to be able to patch, maintain, and develop our servers while still having a fail-safe when something goes wrong with the current production system. Shortly after we release our production version of 20 before the end of the year, we will decommission the two v18 servers, and go back to having just two sets of servers in Jetstream.

Nexrad AWS THREDDS Server on Jetstream2 Cloud

As part of the NOAA Big Data Project, Unidata maintains a [THREDDS data server](#) on the Jetstream2 cloud serving Nexrad data from Amazon S3. This TDS server leverages Internet 2

high bandwidth capability for serving the radar data from Amazon S3 data holdings. TDS team member, Tara Drwenski, and Science gateway staff recently collaborated to upgrade this server.

Jetstream2 and Science Gateway Security

We continually work with Unidata system administrator staff to ensure that our web-facing technologies and virtual machines on Jetstream2 adhere to the latest security standards. This effort involves such tasks as ensuring we are employing HTTPS, keeping cipher lists current, ensuring docker containers are up-to-date, limiting ssh access to systems, etc. It is a constantly evolving area that must be addressed frequently.

Unidata Science Gateway Website and GitHub Repository

Website

The [Unidata Science Gateway web site](#) is regularly updated to reflect the progress of what is available on the gateway. The news section is refreshed from time-to-time for announcements concerning the gateway. The conference section and bibliography is also maintained with new information. We are in the process of redesigning this web site. See “Unidata Science Gateway Re-Imagined” section above.

Repository

All technical information on deploying and running Unidata Science Gateway technologies is documented in the [repository README](#). This document is constantly updated to reflect the current state of the gateway.

Presentations/Publications/Posters

- M. Ramamurthy, J. Chastang, and A. Espinoza. Unidata Science Gateway: A research infrastructure to advance research and education in the Earth System Sciences. In 15th International Workshop on Science Gateways, Tübingen, Germany, Jun. 13-15 2023. See also <https://science-gateway.unidata.ucar.edu>.
- M. Ramamurthy, J. Chastang, and A. Espinoza. Unidata Science Gateway: A research infrastructure to advance research and education in the Earth System Sciences. In EGU General Assembly 2023, Vienna, Austria, Apr. 23-28 2023. See also <https://science-gateway.unidata.ucar.edu>.
- A. Espinoza, J. Chastang, and W. G. Blumberg. Deploying an Educational JupyterHub for Exploratory Data Analysis, Visualization, and Running Idealized Weather Models on the Jetstream2 Cloud. Manuscript submitted for publication, 30 Oct.--1 Nov. 2023.

New Activities

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

following:

Forthcoming Conference Attendance

- Science Gateways 2023 Annual Conference
- 2024 American Meteorological Conference

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

Tomcat 8.5 End of Life

Tomcat 8.5 will reach end of life on 31 Mar 2024. This will require staff to transition the Tomcat Docker containers and any dependencies to the newer version of Tomcat.

Improved JupyterHub Kubernetes Cluster Stability

We aim to provide an optimal experience for our users, but unfortunately, we've experienced more downtimes than we'd prefer. Specifically, issues with disk attachments have disrupted users' ability to consistently access their Jupyter instances. To proactively address these issues, we plan to use cluster monitoring software like Prometheus and Grafana. This will allow us to identify and resolve problems before they impact the user experience.

Relevant Metrics

Spring/Summer/Fall 2022 JupyterHub Servers

Since spring of 2020, Unidata has provided access to JupyterHub scientific computing resources to about 1500 researchers, educators, and students (including a few NSF REU students) at 18 universities, workshops (regional, AMS, online), and the UCAR SOARS program. Below are the latest metrics since the last status report.

	<u>No. of users</u>	<u>POC</u>
<u>Spring 2023</u>		
AMS 2023 Python Workshop	87	Drew, Nicole, Ana, Julien
AMS 2023 CSU LROSE Workshop	24	Jen DeHart, Julien
AMS 2023 MetPy Short Course	30	Drew, Ryan, Kevin, Ana
LROSE University of	15	Prof Mike Bell (CSU)

Hawaii WS		
Florida State University	31	Prof Chris Holmes
Florida Institute of Technology	10	Prof Steve Lazarus
University of Oklahoma	3	Ben Schenkel
Millersville University (3 classes!)	33	Prof Greg Blumberg
Penn State University	16	Prof Paul Markowski
Saint Cloud State University	7	Prof Matthew Vaughan
University of Louisville	11	Prof Jason Naylor
University of Wisconsin	0	Pete Pokrandt
Virginia Tech University	12	Prof Craig Ramseyer
Southern Arkansas University	4	Keith Maull
Northern Illinois University (GPU)	2	Alex Haberlie
Summer 2023		
UCAR SOARS Internship	15	Keith Maull, UCAR/UCP
Unidata users workshop	66	Unidata Staff
I-Guide	16	Drew, Ryan
UCAR Professional Development Workshop Series 7	30	Unidata Staff: Drew, Nicole, Thomas
UND Summer Workshop	10	David Delene
MetPy for Quantitative Analysis of Meteorological Data	21	Unidata Staff: Drew, Nicole, Thomas
Python Readiness	10	Unidata Staff: Drew, Nicole, Thomas

Series: Train-the-Trainer		
<u>Fall 2023</u>		
Florida Institute of Technology	9	Prof Milla Costa
Metropolitan State University of Denver	19	Erin Rhoades
Millersville University	2	Prof Greg Blumberg
University of Oklahoma	2	Ben Schenkel
University of Oklahoma 2	1	Professor Sakaeda
Southern Arkansas University	33	Keith Maul
University of Louisville	7	Prof Jason Naylor
University of Wisconsin	26	Pete Pokrandt
University of Wisconsin 2	15	Prof Hannah Zanowski
CSU Python Workshop 1	25	Unidata Staff: Drew, Nicole, Thomas
CSU Python Workshop 2	14	Unidata Staff: Drew, Nicole, Thomas

Jetstream2 Allocation Usage Overview

In addition to service units (SUs) used for running various kinds of virtual machines, “regular” CPU instances, and GPU instances, Unidata was also granted a limited number of compute, storage, and network resources to carry out Jetstream2 operations. These three kinds of resources are ephemeral, being created and destroyed as necessary. Thus, metrics regarding these resources are representative of short term utilization, while SU usage is a metric that can be representative of our long-term Jetstream2 utilization. As Unidata was only recently granted a new 8M+ SU allocation, starting October 2023, SU usage may not prove a useful metric and has been omitted for this Status Report. Resource metrics current as of October 16, 2023 are presented below.

Resource Metrics

<u>Compute</u>			
Type	Used	Total	Percent Usage*
Instances	77	150	51 %
vCPUs	1034	4035	26 %
RAM	3.9 TB	15.8 TB	25 %

<u>Storage</u>			
Type	Used	Total	Percent Usage*
Volumes	206	400	52 %
Volume Snapshots	0	50	0 %
Volume Storage	31.0 TB	39.1 TB	79 %

<u>Network</u>			
Type	Used	Total	Percent Usage*
Floating IPs	47	310	15 %
Security Group	61	100	61 %
Security Group Rules	198	300	66 %
Networks	4	100	4 %
Ports	111	250	44 %
Routers	2	15	13 %

* Percent Usage is rounded to the nearest whole number

Github Statistics*

Repository	Watches	Stars	Forks	Open Issues	Closed Issues	Open PRs	Closed PRs
science-gateway	6 (+2)	17 (+1)	11	5 (+1)	167 (+1)	14 (+8)	682 (+86)
tomcat-docker	11 (+1)	60 (+1)	64 (-1)	2	40	1	83(+11)
thredds-docker	15	27 (+2)	26(+1)	4	117(+7)	0	176 (+17)
ramadda-docker	4	0	2	1	10	0	34 (+10)
ldm-docker	9(+1)	12(-3)	13	1(-4)	40(+4)	0	65(+4)
tdm-docker	5(+1)	4	7	0 (-1)	10 (+1)	0	23 (+5)

* Numbers in parentheses denote change from last stat report

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. Managing Geoscience Data

Unidata supplies a good portion of the data available on the IDD network to the Jetstream2 cloud via the LDM and the high bandwidth Internet 2 network. Those data are distributed to the TDS, ADDE, RAMADDA and AWIPS EDEX installations running on Jetstream2 for the benefit of the Unidata community. Unidata also makes the AWS Nexrad archive data accessible through the TDS Nexrad server running on Jetstream2 at no cost to the community. These data can be accessed in a data-proximate manner with a JupyterHub running on Jetstream2 for analysis and visualization. Containerization technology complements and enhances Unidata data server offerings such as the TDS and ADDE. Unidata experts install, configure and in some cases, security harden Unidata software in containers defined by Dockerfiles. In turn, these containers can be easily deployed on cloud computing VMs by Unidata staff or community members that may have access to cloud-computing resources.

2. Providing Useful Tools

Jupyter notebooks excel at interactive, exploratory scientific programming for researchers and their students. With their mixture of prose, equations, diagrams and interactive code examples, Jupyter notebooks are particularly effective in educational settings and for expository objectives. Their use is prevalent in many scientific disciplines including atmospheric science. JupyterHub enables specialists to deploy pre-configured Jupyter notebook servers typically in cloud computing environments. With JupyterHub, users login to arrive at their own notebook workspace where they can experiment and

explore preloaded scientific notebooks or create new notebooks. The advantages of deploying a JupyterHub for the Unidata community are numerous. Users can develop and run their analysis and visualization codes proximate to large data holdings which may be difficult and expensive to download. Moreover, JupyterHub prevents users from having to download and install complex software environments that can be onerous to configure properly. They can be pre-populated with notebook projects and the environments required to run them. These notebooks can be used for teaching or as templates for research and experimentation. In addition, a JupyterHub can be provisioned with computational resources not found in a desktop computing setting and leverage high speed networks for processing large datasets. JupyterHub servers can be accessed from any web browser-enabled device like laptops and tablets. In sum, they improve "time to science" by removing the complexity and tedium required to access and run a scientific programming environment.

3. Supporting People

A Unidata science gateway running in a cloud computing setting aims to assist the Unidata community arrive at scientific and teaching objectives quickly by supplying users with pre-configured computing environments and helping users avoid the complexities and tedium of managing scientific software. Science gateway offerings such as web-based Jupyter notebooks connected with co-located large data collections are particularly effective in workshop and classroom settings where students have sophisticated scientific computing environments available for immediate use. In the containerization arena, Unidata staff can quickly deploy Unidata technologies such as the THREDDS data server to support specific research projects for community members.

Status Report: Community Services

April 2023- November 2023

Nicole Corbin, Doug Dirks, Tanya Vance, Jeff Weber

Executive Summary

In addition to “normal, day-to-day” activities of communication and coordination with community members, the Community Services group’s efforts in the past six months have been focused on:

- Outreach to underserved communities and active engagement and activities with tribal colleges and universities and the Data Sovereignty Network project partners including sensor install and site configurations at SIPI, WRF output for NTU, and the addition of The Haskell Foundation and Nebraska Indian Community College to the Sovereign Network
- Outreach to the Earth System Science community through participation at ESIL Innovation Summit, Rising Voices, CUAHSI, AMS, and AGU conferences, the NCAR NESTs Coastal Resiliency Community Workshop, February AIHEC TCU Data Workshop, and Earth Data Relations and ESIL NSF working groups
- Expansion of learning materials, resources, and offerings
- Prototyping the Reimagined Science Gateway and Integrated Educational Hub to make our educational services more discoverable on the web as well as promoting the variety of educational services we are able to provide
- Received [NSF Award #2319979](#) for a series of introductory AI/ML educational resources
- Providing support for the planning and implementation of the 2023 Unidata Users Workshop
- Facilitated CSU Python Readiness Blended Learning Series including Train the Trainer, NCAR|UCAR Intern Professional Development Workshop Series on Data Management and Visualization, and the UMBC Regional MetPy for Quantitative Analysis of Meteorological Data Workshop with Howard University, Morgan State University, UMBC, NCAS-M, and John Hopkins University
- Partnered with COMET and USGS on development of a series of NetCDF/CF microlearning resources and delivery of the Metpy for Quantitative Analysis training
- Support creation of Unidata’s next core funding award proposal

Questions for Immediate Committee Feedback

No questions at this time.

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:

- [Why are GPUs Exciting for Machine Learning Research?](#)
- [Free Online Courses for Machine Learning](#)
- [Deploying The Littlest JupyterHub at the University of Wisconsin-Milwaukee](#)
- [Meet Unidata's 2023 Summer Interns](#)
- [Offer: Unidata Science Gateway JupyterHub Resources Available for Fall 2023 Courses](#)
- [Summer 2023 Unidata Interns Wrap Up Their Projects](#)
- [2023 Users Workshop Explores Storytelling with Earth Systems Science Data](#)
- [2023 Community Equipment Awards](#)
- [Unidata Welcomes New committee Members](#)
- Software release information
- Many AWIPS Tips and MetPy Mondays episodes
- 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. We're starting to make progress working with committee members, but there is more to do.
- A sufficient number of committee member nominations and diverse and inclusive representation across ESS

Community Outreach and Services

The community services group continues to actively reach out to and engage with Unidata community members.

Progress has been made on the following:

- The community services team has led efforts to craft the core proposal for the next 5 years building on the strategic plan
- The community services team has supported the Users Committee's work in planning and execution of the June 2023 Unidata Users Workshop
- Continue to work with AIHEC, NEON, ESIL, and UCAR/NCAR staff on NSF funded MSI engagement and NSF funded Data Sovereignty Project
- Successful training and microwave network bridge install, network system install, and server connection at SIPI in partnership with NTU, successfully running WRF over the Navajo Nation and pushing output to SIPI and NTU
- Participating on new one year grant NSF#2220614 "The Indigenous Data Governance in Open Data Working Group"
- Participating with members of Oglala Lakota College and ESIL focusing on data access, visualization, and analysis for Earth System Science; and efforts are underway to investigate potential pathways to foster complementary integration with the Unidata-partnered, Sovereign Network
- Continue to engage with underserved populations and institutions as part of Unidata's outreach efforts to groups such as Rising Voices, SACNAS, and AIHEC including Rising

Voices Indigenous Peoples Climate Change Working Group (IPCC-WG) and Indigenous Food, Energy, & Water Security and Sovereignty (Indigenous-FEWSS)

- Continue to serve on the CUAHSI HIS and DEI standing committees
- We continue to actively support the NCAR/SOARS program including serving on the 2023 Internship Selection Committee and Advisory Committee
- Engage with the Arctic Research Consortium of the US on multidisciplinary projects
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+)
- We continue to publish short videos/screencasts on the [Unidata YouTube channel](#).
- Represent Unidata at the National Weather Service Partners events
- Actively participate in Super Science Saturday
- Participate in EdEC's Engagement Group to support awareness of UCAR/NCAR/UCP opportunities
- Engage and support the Undergraduate Leadership Workshop (ULW) at UCAR
- Support the development and operation of the UCAR:NCAR Equity and Inclusion (UNEION) community of practice
- Support UCAR/NCAR media services by responding to requests from media outlets

Dependencies, challenges, problems, and risks include:

- Facilitating community adoption of new technological services (cloud, etc)
- Engagement with Unidata social media streams among community members is not particularly high
- Engaging with new communities that have different resources, capacities, and expectations

Learning Services

The community services group has expanded efforts to promote learning Unidata products and workflows.

Progress has been made on the following:

- Awarded [NSF #2319979](#) "Machine Learning Foundations and Applications in the Earth Systems Sciences" (Nicole Corbin, PI and Thomas Martin, Co-PI) in collaboration with Dr. Keah Schuenemann (MSU Denver) and Dr. Karen Kortz (Community College of Rhode Island)
 - Goals:
 - Elucidating the conceptual mechanisms behind machine learning models for an Earth Systems Science audience.
 - Bridging the gap between machine learning conceptual mechanisms and low-code, real-world applications in the Earth Systems Sciences.
- Supported planning and execution of the [2023 Unidata Users Workshop](#) in Boulder, CO and virtually
- Facilitated the Data Management and Visualization session as a part of the UCAR|NCAR 2023 intern professional development series (virtual).
- Facilitated MetPy for Quantitative Analysis of Meteorological Data at University of Maryland Baltimore County Aug 14-15.

- Invited students from UMBC, Howard University, Morgan State University, and John Hopkins University
- Completed the second iteration of the [Python Readiness Learning Series](#) at CSU.
 - Eight graduate students attended the [Train-the-Trainer](#) session
 - Approximately 20 new graduate students completed the full learning series.
- Two new microlearning resources (soft launched, in beta), tentatively will be shared at the 2024 AMS Education Conference
 - [Multidimensional Data Structures](#)
 - [Siphon/THREDDS Data Server Ecosystem](#)
- Prototyping the Science Gateway, Reimagined website, including a comprehensive Education Hub for all Unidata's educational offerings. See Cloud Computing Activities for more details.
- Partnered with COMET and USGS on development of a series of NetCDF/CF microlearning resources, to be completed Summer 2024
 - Will also deliver an off-the-shelf instructor-led training for USGS in Spring 2024

Dependencies, challenges, problems, and risks include:

- Assignment and allocation of resources to develop new material

Community Equipment Awards

The NSF provides the Unidata Program Center up to \$100k in equipment grant funds each year. In alignment with the Unidata 2024 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.

Progress has been made on the following:

The 2023 Equipment Awards review panel met on March 31, following the spring Users Committee meeting. Four institutions received funding this year:

- Coastal Carolina University
- Embry-Riddle Aeronautical University
- Howard University
- University of Missouri

See the [2023 Awards](#) page for details.

Dependencies, challenges, problems, and risks include:

- Increasing participation in the Equipment Awards program, especially by MSIs, smaller programs, and others who have not participated in the past.

Ongoing Activities

We plan to continue the following activities:

- Support for governing committee activities (convening and coordinating discussions, facilitating awards, etc.)
- Ongoing development of news articles and blogs for publication through News@Unidata
- Seeking partnerships to build and deliver community learning and development resources and continue to expand Unidata's educational services
- Continue to engage with underserved populations and institutions
- Engagement with other Earth System Science organizations, professional societies, and working groups
- Support the pursuit of funding and bringing greater public awareness to Unidata
- Engage other UCAR/NCAR divisions regarding Unidata software use and collaborative projects
- Ongoing work to transition Unidata's website to UCAR-mandated system
- Active participation in the Hydroshare Advisory Committee (CUAHSI)
- Continue working with AIHEC, NEON, ESIL, and UCAR/NCAR staff on an NSF funded MSI engagement and data sovereignty project

New Activities

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

- Preparation for AMS training offerings
 - Student Conference: AWIPS
 - Multidimensional Data Structures
 - Siphon/THREDDS Data Server Ecosystem
- Present poster at AGU on The Sovereign Network: A Platform for Interoperable, Convergent Science and present poster on Indigenous Science to Action: Authentic Contexts for Supporting Indigenous Priorities
- Co-planning and hosting the AIHEC TCU Data Workshop funded by NCAR CORE award and focusing on earth system science and the Sovereign Network Project
- Participation in the community design and development of the NCAR Earth System Science and Technology Hubs (NESTs) around coastal hazards, flooding, and extreme weather
- Needs assessment for AI/ML project and development of first eLearning module
- Strategic planning for Science Gateway implementation
- Delivery of final-version microlearning modules for NetCDF and Siphon
- Equipment Awards
- 6-month evaluation of Users Workshop goals

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

- Delivery of NetCDF resources in partnership with COMET and USGS
- Delivery of Unidata Education Hub web presence
- Delivery of prototypes of all three AI/ML educational modules
- Engage other underrepresented communities
- Expand sovereign network to other TCU's and AIHEC members
- Prepare for follow up proposal w NTU, SIPI, and TOCC to expand the Sovereign Network

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

- Continued maintenance of the reimagined Unidata Science Gateway, including community contributions and adding resources to the catalog
- Champion CARE principles
- Support community hubs

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:

- 34,692 unique pageviews (40,323 in previous period)
- 16.7% of total unique pageviews (11.4% in previous period)

Top community pages

1. All blog pages
26395 unique pageviews (25126 in previous period)
76% of total community pageviews (60% in previous period)
2. www.unidata.ucar.edu/community
1874 unique pageviews (8781 in previous period)
5% of total community pageviews (21% in previous period)
3. www.unidata.ucar.edu/about
1240 unique pageviews (2965 in previous period)
4% of total community pageviews (7% in previous period)
4. www.unidata.ucar.edu/events
4830 unique pageviews (2983 in previous period)

14% of total community pageviews (7% in previous period)

Social media statistics, October 11, 2023

1. # of Twitter/X followers: 2013 (up from 1978)
2. # of Facebook followers: 908 (up from 899)
3. # of YouTube subscribers: 3508 (up from 3160)
4. #of LinkedIn followers: 145 (up from 89)

Unidata eLearning statistics, October 11, 2023

1. Total unique users: 162
2. Enrolled users in Learn AWIPS CAVE: 138
3. Enrolled users in Learn Python-AIPS: 18

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. Managing 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.

2. Supporting People

We provide user workshops, tutorials, and community workshops to help build supportive relationships between community members.

We coordinate with our governing committees to find ways to expand Unidata's community participation. We use our website, 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 underrepresented 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.

Status Report: Data Services

April 2023- November 2023

Mike Zuranski

Executive Summary

Progress on the RTSTATS Revamp has been made, before pausing to get started on IDD Cataloging. RTSTATS development will resume shortly.

IDD Cataloging is underway. Presently, a week-long database of Product IDs is retained, which has already been rather helpful in assisting staff, as well as to identify duplicate data on CONDUIT which has recently been remedied.

Other primary duties include working with other teams and staff to assist in their needs. By working alongside the other engineering teams at Unidata, new opportunities have been exploited; for example the THREDDS team will be using the TIG stack as part of their performance testing and I am assisting in that deployment. I have also been helping the AWIPS team test new versions of CAVE on various platforms.

User support of course continues to be a primary focus. See the **Status Report: Support** for inquiry metrics.

Questions for Immediate Committee Feedback

None at this time.

Activities Since the last status report

RTSTATS Revamp

Work on the RTSTATS Revamp project continued for a time until a proof-of-concept reached a stable place. At that point development was put on hold, in favor of getting a foot in the door into IDD Cataloging instead.

Development is planned to resume in the coming weeks.

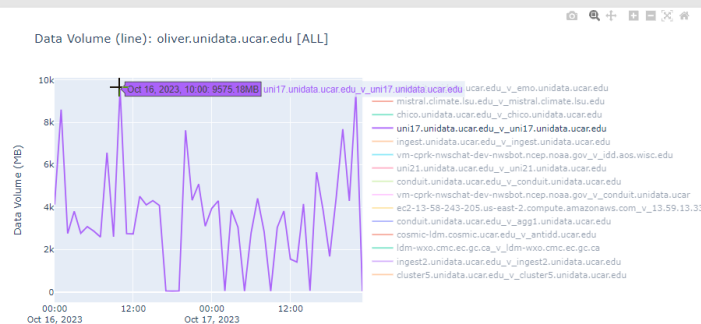
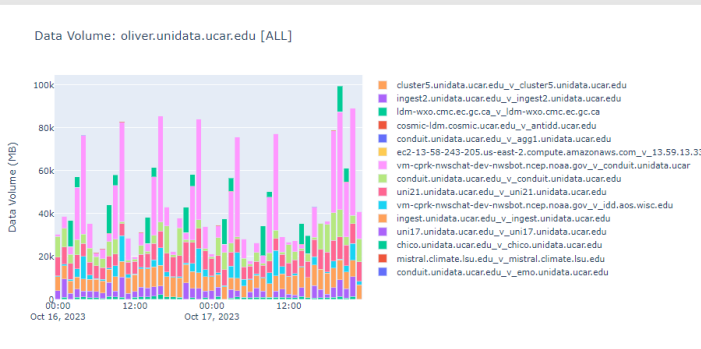
Real-Time IDD Statistics

The Unidata community of over 260 universities is building a system for disseminating near real-time earth observations via the Internet. Unlike other systems, which are based on data centers where the information can be accessed, Unidata Internet Data Distribution (IDD) is designed so a university can request that certain data sets be delivered to computers at their site as soon as they are available from the observing system. The IDD system also allows any site with access to specialized observations to inject the dataset into the IDD for delivery to other interested sites.

Host:

Feedtype:

Date Range: →



A proof-of-concept demo is running, though is not quite ready for announcing to the public. Please contact Mike Zuranski for more information on this demo.

Progress has been made on the following:

- RTSTATS Front-End using Dash
- Designing a new database schema for RTSTATS data storage

Dependencies, challenges, problems, and risks include:

- Time availability

IDD Cataloging

Work has begun on the IDD Cataloging project. Mike Zuranski has built a process that utilizes notyfme output to log all the product IDs that come out of the public feeds at idd.unidata.ucar.edu. These product IDs, along with some metadata, are stored in a PostgreSQL database. A front-end, built with Flask, presents a variety of forms that allow a

user to search through that database in different ways. Ultimately what this tool does is find product IDs for you that have already come across the IDD. The current retention period is one week, any more than that and database queries become problematically slow.

I would also like to share a recent success story. This catalog was able to reveal that NDFD data on CONDUIT was being transmitted twice, under two different sets of product IDs! After confirming the two data sets were bitwise identical we alerted NCEP that this was occurring. They were able to correct the problem as part of a set of updates to their CONDUIT environment on 10-12 (though this did lead to [some issues for users](#)).

Progress has been made on the following:

- A process that monitors the IDD and records products seen
- A front-end that can query this database multiple ways

Dependencies, challenges, problems, and risks include:

- The product IDs are all strings, and substring searching through rows of a database is slow, very slow. So slow that if any more than about one week’s worth of product IDs are stored, database queries slow down so much that there is a noticeable impact on the web server’s performance.
- There needs to be a better way to index this information, to allow for significantly faster queries.

IDD Catalog - Custom Search: idd.unidata.ucar.edu

Search String: Feedtype: Start Date/Time: End Date/Time:

Show entries

Product ID	Feed Type	Size (bytes)	Insertion Time
data2ndfd/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00130/TMKN01/2 m HGHTI 000000	CONDUIT	989719	2023-10-11 10:22:05.156287
data2ndfd/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F02530/TMKN12/2 m HGHTI 000001	CONDUIT	965525	2023-10-11 10:22:05.172760
data2ndfd/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04930/TMKN12/2 m HGHTI 000002	CONDUIT	934829	2023-10-11 10:22:05.186812
/common/data/products/NDFD/grib2/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00130/TMKN01/2 m HGHTI 000000	CONDUIT	989719	2023-10-11 10:22:05.535311
/common/data/products/NDFD/grib2/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F02530/TMKN12/2 m HGHTI 000001	CONDUIT	965525	2023-10-11 10:22:05.549503
/common/data/products/NDFD/grib2/YHUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04930/TMKN12/2 m HGHTI 000002	CONDUIT	934829	2023-10-11 10:22:05.556564
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00130/SNOW01/0 - NONEI 000000	CONDUIT	204957	2023-10-11 10:22:11.469393
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00130/SNOW01/0 - NONEI 000000	CONDUIT	204957	2023-10-11 10:22:11.470224
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00730/SNOW06/0 - NONEI 000001	CONDUIT	218760	2023-10-11 10:22:11.471447
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F00730/SNOW06/0 - NONEI 000001	CONDUIT	218760	2023-10-11 10:22:11.473466
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F01330/SNOW06/0 - NONEI 000002	CONDUIT	215605	2023-10-11 10:22:11.475105
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F01330/SNOW06/0 - NONEI 000002	CONDUIT	215605	2023-10-11 10:22:11.476488
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F01930/SNOW06/0 - NONEI 000003	CONDUIT	220887	2023-10-11 10:22:11.479632
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F01930/SNOW06/0 - NONEI 000003	CONDUIT	220887	2023-10-11 10:22:11.480865
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F02530/SNOW06/0 - NONEI 000004	CONDUIT	227345	2023-10-11 10:22:11.482579
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F02530/SNOW06/0 - NONEI 000004	CONDUIT	227345	2023-10-11 10:22:11.484158
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F03130/SNOW06/0 - NONEI 000005	CONDUIT	229789	2023-10-11 10:22:11.485796
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F03130/SNOW06/0 - NONEI 000005	CONDUIT	229789	2023-10-11 10:22:11.488184
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F03730/SNOW06/0 - NONEI 000006	CONDUIT	224979	2023-10-11 10:22:11.490569
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F03730/SNOW06/0 - NONEI 000006	CONDUIT	224979	2023-10-11 10:22:11.492195
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04330/SNOW06/0 - NONEI 000007	CONDUIT	216572	2023-10-11 10:22:11.493215
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04330/SNOW06/0 - NONEI 000007	CONDUIT	216572	2023-10-11 10:22:11.495015
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04930/SNOW06/0 - NONEI 000008	CONDUIT	205553	2023-10-11 10:22:11.497056
/common/data/products/NDFD/grib2/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F04930/SNOW06/0 - NONEI 000008	CONDUIT	205553	2023-10-11 10:22:11.498490
data2ndfd/YSUZ98_KVBN_111020.grib2.lgrib2/mvstg/NWS_0/#000/202310111030F05530/SNOW06/0 - NONEI 000009	CONDUIT	192301	2023-10-11 10:22:11.499431

Showing 1 to 25 of 31,448 entries

Previous 1 2 3 4 5 ... 1258 Next

Can you find the duplicates?

Status and Take-Aways Thus Far

Almost as soon as it was implemented, having a record of products seen on the IDD became useful. We've been able to use it to help investigate issues like product outages, the scope and duration of those outages, and locating data for users. Even though the record is only one week long, already that has proved beneficial.

The original goal was to store a year's worth of data, because some products only come around once a year. One week is quite short of one year. It's clear that storing and searching through raw product IDs will not be a viable way to manage this information.

I have developed a plan on how to move forward in a way that should not only allow for faster queries, but also longer retention and a better end use-case for users. The idea is to take product IDs that come across the IDD to build a catalog of what products to **expect**. By doing that, we can build dashboards to show the status of things such as model runs or satellite data. Imagine a page showing not only what variables to expect for each model, by both runtime and forecast hour, but also whether or not they have come in yet or if they're late. By knowing what to expect we can then know what's missing and what's new, and that opens doors to things such as alerting and automating the detailing of new data products.

I paused RTSTATS development in order to focus on this. At this time it's clear I need to change strategies a tad, and I won't have anything that will be publicly usable for some time. As such, I plan to bring things to a good stopping point soon so I can go back to focus on RTSTATS, which should prove a simpler project to bring to completion.

I will maintain a version of this current IDD Cataloging running for my own purposes, mostly to help me assist users & staff because again it has already helped me immensely. Unfortunately it's not stable enough to make public at this time.

Improved Monitoring with the TIG stack

We have begun hosting a publicly available Grafana page to showcase various metrics. The page is available at status.unidata.ucar.edu. No account or login is required. Currently, dashboards exist for our LDM metrics and our NOAAPort ingest metrics. There are also plans to work with the THREDDS team to utilize the TIG stack to host and visualize TDS performance test results.

Progress has been made on the following:

- Telegraf deployment to all LDM and NOAAPort internal machines
- Scripting to collect LDM and NOAAPort metrics
- Public page: <https://status.unidata.ucar.edu>
- Multiple LDM and NOAAPort dashboards

Dependencies, challenges, problems, and risks include:

- Working alongside the UCAR security perimeter
- Ensuring proper security policies for public access

Add Data to the IDD

The SPIRE Radio Occultation data that the [UCAR COSMIC Program](#) provided us had ceased once that particular project ran its course. Through a new partnership with [PlanetIQ](#), COSMIC has resumed processing and sharing this data, which we continue to push to the IDD. Details of the data have not changed since last time; these data products include three netCDF and one BUFR product, at rates of roughly 80-110MB per hour.

Dependencies, challenges, problems, and risks include:

- While we can see how much of this data is going into the IDD with RTSTATS, it's nigh on impossible to gauge how much data is actually being used at sites. COSMIC wanted to know how much additional reach adding this data to the IDD gave them, and that information is difficult to determine.

Ongoing Activities

We plan to continue the following activities:

- Evaluate current methods for information dissemination (e.g. community email lists)
- Establish better lines of communication with data providers
- Explore new tools for UCP staff
- Explore new data sets
- Assist users in finding and using data
- Better understand and help manage geoscience data in our industries today

New Activities

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

- Resume development on the RTSTATS Revamp project
- Continue to investigate a new support/ticketing system
- Revamp the data pages on the Unidata website

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

- Advanced IDD Cataloging, making it easier to search for what's available

- Leverage the IDD catalog to assist in decision making in regards to what data we provide. Learn more about what's being used and what isn't, what can we add and take away.
- Begin offering a "data listing" page, to help find external sources of data

Relevant Metrics

N/A at this time

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

Gain a detailed understanding of what data is available, where and in what format. Convey that information to our users and partners. Work with data providers to help distribute and support their data sets. Support users and UCP staff in finding and working with data. Identify opportunities to reduce our data footprint, including the removal of duplicate data sets.

2. **Providing Useful Tools**

RTSTATS modernization, IDD Cataloging, monitoring and alerting solutions.

3. **Supporting People**

User email support, training workshops, transcend UPC silos and support UPC staff.

Prepared *October 2023*

Status Report: Data Standards and Technical Outreach

April 2023- November 2023

Ethan Davis, Ward Fisher, Hailey Johnson, Dennis Heimbigner, and Ryan May

Executive Summary

Unidata's netCDF teams continues to engage with the Zarr community on:

- 1) Zarr support in both the netCDF-C and netCDF-Java libraries;
- 2) the development of the Zarr version 3 specification; and
- 3) the development of the GeoZarr convention.

Unidata continues to be active in efforts to advance the Climate and Forecast (CF) Conventions for netCDF.

Unidata continues to be active in several international standards bodies and other communities focused on data and technology including the World Meteorological Organization (WMO), the Open Geospatial Consortium (OGC), and the Earth System Information Partners (ESIP).

Questions for Immediate Committee Feedback

No questions at this time.

Activities Since the Last Status Report

NCZarr/Zarr Specification Efforts

As part of implementing Zarr support in both the netCDF-C and -Java libraries, the NCZarr convention/extension has been developed to provide a clean and complete mapping between the netCDF and Zarr data models. During this work, the netCDF developers have been participating in discussions around clarification and evolution of the Zarr (version 2 and 3) specifications.

Progress has been made on the following:

- Members of Unidata's netCDF teams have membership on the Zarr Implementation Committee and the Zarr Enhancement Protocol (ZEP) Committee and regularly participate in the bi-weekly Zarr Community and ZEP calls.
- Members of Unidata's netCDF team participated in discussions around the GeoZarr convention. GeoZarr builds on the Climate and Forecast (CF) Conventions for netCDF and will be developed within the OGC community standards process.

CF Conventions for netCDF activities

Unidata has a long history of involvement in the development of the [Climate and Forecast \(CF\) Conventions for netCDF](#). These efforts continue with ongoing participation in development conversations on the [CF GitHub repositories](#), participation in and help in organizing the annual CF Workshops, and participation in the governance of CF.

Progress has been made on the following:

- The [2023 CF Workshop](#) was held virtually on 3-5 Oct 2023.
- Ethan Davis continues serving as chair of the [CF Governance Panel](#).

Ongoing Activities

We plan to continue the following activities:

- Track and engage in WMO data standards efforts
 - Ethan Davis is a member of the WMO Expert Team on Data Standards (ET-Data) and its Task Team for CF-netCDF (TT-CFNetCDF)
 - WMO CF-netCDF Profiles have been developed for radar data, oceanographic glider data, and aircraft data. Experimental distribution of CF-netCDF data on the WMO Information System (WIS) 2.0 is planned.
 - The WIS 2.0 provides similar functionality to the GTS as well as more interactive access to data. We plan to take a closer look at WIS 2.0 technologies to better understand possible connections with the IDD/LDM and how WIS 2.0 may impact and benefit the University community. We have had meetings with NWS and Environment Canada on how the LDM and THREDDS might interact with WIS 2.0 and it's underlying OGC standards.
- Continue efforts to update and reorganize the NetCDF User's Guide (NUG)
 - Separate the aspect of netCDF that are useful to any user/developer, independent of which library or tool they use (i.e., data model, file formats, CDL definition, conventions, and best practices) from those that are library or language specific and
 - Clarify where and how the netCDF community can ask questions about the NUG as well as discuss and contribute to the development and advancement of the NUG.
- Continue efforts to registering netCDF Media Type (application/netcdf) with IANA
 - The netCDF media type has been added to IANA's provisional registry list with Unidata listed as the standards-related body supporting the effort.
 - Next Step: Complete documentation and metadata needed for full registration. The updated NUG will feed into this effort
- Represent Unidata in Earth System Information Partners
 - Unidata has been a Type II ESIP Partner Organization since 1999
 - Ethan Davis is currently the Unidata voting representative to ESIP.
- Represent UCAR and Unidata in OGC and various OGC working groups
 - Ethan Davis is the UCAR voting representative to the OGC Technical

Committee, Kevin Sampson (NCAR/RAL, GIS group) is alternate voting representative.

- Participate in OGC MetOcean Domain Working Group (DWG) meetings.
- Ethan Davis is co-chair of the OGC netCDF Standards Working Group (SWG)
- Track and participate in the OGC Environmental Data Retrieval (EDR) SWG meetings.
- Track and participate in OGC Community Standard process for CoverageJSON.

New Activities

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

- Deploy a draft version of the new, library independent, NetCDF User's Guide (NUG).

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

- Submit request for full registration of the netCDF media type with IANA

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

- 1. Managing Geoscience Data**

Unidata's various data standards efforts contribute to important tools for data producers, especially those that design and develop new data products, and for those that develop software tools for data management, analysis, and visualization.

Prepared *October 2023*

Status Report: GOES-East/West, NOAAPort and Other Satellite Imagery

April 2023- November 2023

Mike Schmidt, Tom Yoksas, Mike Zuranski, Stonie Cooper

Executive Summary

Unidata continues to operate satellite downlink facilities for the NOAAPort Satellite Broadcast Network (SBN) and GOES-East and GOES-West rebroadcast services on behalf of UCAR/NCAR and the Unidata community. All received products are then provided via the Internet Data Distribution system (IDD) in various feeds and via remote access provided by AWIPS EDEX, McIDAS ADDE and THREDDS Data Servers.

Details on various efforts related to maintaining this capability are presented below.

Questions for Immediate Committee Feedback

Does the committee have any recommendations for other products or services based on our existing satellite ingest that would be beneficial for the community?

Activities Since the Last Status Report

- Repurposed a 3.7 m mesh dish located at FL-2 from GOES-15 (GVAR) ingest to NOAAPort SBN ingest, and re-pointed the existing FL-2 NOAAPort dish at the new SBN satellite, Galaxy 31.

These moves were made for three reasons related to the NWS movement of the SBN from Galaxy 28 (located at 89W) to Galaxy 31 (located at 121W):

- The need for continued periodic GOES-15 (which formerly acted as GOES-West) operations was obviated by the successful launch, test out, and christening of GOES-18 as GOES-West, and the movement of GOES-17 to a parking orbit at 104.7W where it will serve as an on-orbit spare for GOES-East and GOES-West
- A second FL-2 mesh dish is configured to ingest GOES-14 data whenever it is returned to service either for checkout or if it assumes an active role as a spare for either GOES-East or GOES-West
- The line-of-sight from the existing NOAAPort dish on the south side of FL-2 to Galaxy 31 is impaired by large branches from trees outside of FL-3 and FL-2

Getting the worst offending tree removed or, at least, pruned enough to mitigate the problems being caused by the branches is unlikely.

- Work towards establishing an additional GOES-R downlink in a locations that has an unimpeded view of the southern sky continues

See Ongoing Activities for additional information.

- Establish a new NOAAPort downlink at the NSF-owned Marshall field site

An effort to establish a satellite downlink facility at the NCAR Marshall field site (just south of Boulder) has been slowed by an NSF moratorium on any ground penetrations until an environmental impact assessment (NEPA) has been completed. A non-penetrating ground mount will be installed in the Marshall compound, and a 3.8 m dish will be installed on the mount in the coming weeks. Following the satellite pad and dish installations, electronics needed to complete the downlink will be installed, and ingest testing will begin.

After the Marshall installation is complete, and assuming that high quality NOAAPort ingest can be achieved, and the interference of the trees at FL-2 can not be mitigated, the existing FL-2 NOAAPort solid dish will be converted to GRB downlink as it has an unobstructed view of the GOES-East orbital slot. This conversion would require that existing quad-shielded RG-11 coax be replaced by a dual run of LMR-400 coax from the dish to the 2nd floor computer room, and the LNB on the dish outside of the FL-2 cafeteria be moved to the dish being repurposed.

Ongoing Activities

We plan to continue the following activities:

- Participate in UW/SSEC's "fanout server" sharing of GOES-R/S data (redistribution of the GRB-200 UDP unicast stream over TCP) for GOES-16/18 GRB products.

We are feeding from SSEC's GOES-16/18 fanout servers, and they are feeding from the ingest machine that we operate. Sharing of the feed streams has allowed SSEC and Unidata to minimize effects of solar and terrestrial interference.

- Ingest GOES ReBroadcast (GRB) streams from GOES-16 and GOES-18 in real-time

As described in previous status reports, the 4.5 m dish located on the eastern satellite pad at the NCAR Mesa Lab has been used for GOES-West ingest since terrestrial interference (TI) was observed when pointing at GOES-East (GOES-16) in the fall of 2017. Replacement of power poles and lines to the south and downhill from the Mesa Lab prompted us to run a test to see if TI was still a problem when pointing at GOES-16. The results of the multi-day test convinced us that we could return the dish on the eastern satellite pad to GOES-East ingest, and proceed with installation of a dish on the western pad and use it for GOES-West ingest.

In the spring of 2022, we were given a 3.8 m satellite dish that was being excessed by a private company that was relocating their operations. This dish will be installed on the western satellite pad at the NCAR Mesa Lab. The running of dual coax cables from the western pad to the main Mesa Lab machine room has been completed, so the next step is the physical installation of this dish on the existing mounting pole

- Continue to distribute GOES-16 and GOES-18 data via the LDM/IDD and serve the data via the TDS, ADDE and EDEX

The volume of data available in the SATELLITE datastream can be seen in:

http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?SATELLITE+oliver.unidata.ucar.edu

Future Activities

CSPP GEO Gridded Geostationary Lightning Mapper (Gridded GLM)

On March 21, 2021 Graeme Martin (UWisconsin/CIMSS) announced the initial release of **Gridded Geostationary Lightning Mapper (Gridded GLM)** software package:

The software is capable of processing GOES-16 and GOES-17 GLM Level 2+ products in mission standard format, generating a new set of products which have been gridded to the Advanced Baseline Imager (ABI) 2-km resolution, and are aggregated at one-minute intervals. Spatial extent information that is not readily available in the GLM L2+ data is recovered and used to create the gridded products.

The following products can be produced:

- *Minimum Flash Area*
- *Flash Extent Density*
- *Total Optical Energy*

AWIPS-compatible tiles can optionally be generated, using functionality that was developed within the open source Python SatPy library.

Input GLM L2+ files can be obtained from the CSPP Geo GRB software running at a direct broadcast site, or from NOAA CLASS. Output is in NetCDF4 format.

We intend to implement this software, evaluate the products, and distribute them in the IDD when appropriate.

Gridded Geostationary Lightning Mapper (Gridded GLM) products from Amazon AWS S3

We obtained access (effort spearheaded by Tiffany Meyer) to Gridded GLM products being

created by the NWS for use in forecast offices. Redistribution of these products in the IDD as replacements for the Gridded GLM products previously created by Eric Bruning of Texas Tech University was implemented in early summer 2022.

Himawari Imagery and Level 2 Products

We have also obtained access to Himawari imagery and Level 2 products from Amazon AWS S3. We are asking the User Committee to weigh in on the importance/need of adding some of these products to the IDD. One thing that must be kept in mind is the volume of Himawari data is *large*, so the ability of end user sites to handle real-time feeds of the full set of data is in question.

IDD FNXRAD, NEXRAD3 and NIMAGE Datastreams

As noted in previous status reports, the IDD **FNXRAD** datastream was enhanced by the addition of MRMS products we receive in an LDM feed from NOAA/NCEP.

The IDD **NIMAGE** feed was repurposed a few years ago from a feed that only contains satellite image products distributed in NOAAPort to one that can include value-added satellite products. The question for the committee is if there are other products that should be added to the **NIMAGE** feed?

VALUE-ADDED Products

We welcome contributions of additional value-added Level 2 satellite products by community members.

To date, Texas Tech University (Eric Bruning), CSU/CIRA, and NOAA's Vlab have provided value-added Level 2 products created from satellite image and lightning scans, and these have been distributed to the community in the NIMAGE IDD feed.

SSEC Collaboration

Continue working with SSEC on their *fanout* approach that insulates GRB ingestion from expected (e.g., twice per year solar interference periods; etc.) and unexpected (e.g., TI caused) service interruptions

L2 Product Creation Testbed

We still want to establish a test bed for the creation of Level 2 (L2) products from GOES-16/18 imagery, model output and observational data.

The objective would be to provide the capability of running user site submitted algorithms to create L2 products and make them available for testing for a short period of time via the IDD,

the TDS, McIDAS ADDE and AWIPS EDEX. This initiative has been slowed by the inability by most staff to work on-site.

Relevant Metrics

- Lots O' Data!

That the volume of GOES-16 and GOES-18 GRB products, 15 GB/hour average and 20 GB/hour maximum, is the second most voluminous IDD feed can be seen in the real-time statistics listings from any of the accumulators for our toplevel IDD relay clusters. For instance :

https://rtstats.unidata.ucar.edu/cgi-bin/rtstats/rtstats_summary_volume?oliver.unidata.ucar.edu

- Feeding data to a slowly growing list of sites via the IDD:

We are distributing all or part of the GOES-East/West GRB products to:

- Groups within UCAR/NCAR (3: all products Unidata, EOL, RAL)
- U.S. Universities (25: variety of feeds; GLM very popular)
- U.S. Government (3: all products to 2 NOAA sites and one Military site)
- International (3: Full Disk imagery and GLM L2 products)

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

Providing TDS, ADDE and EDEX servers for GOES-16/17 imagery and products benefits the greater community by providing access to real-time observations from the U.S. operational satellite constellation.

2. **Supporting People**

Providing access to data in real-time has been a fundamental Unidata activity since its inception. Continuing to provide data enables Unidata sites to focus on their educational and research activities.

Status Report: Internet Data Distribution(IDD)

April 2023- November 2023

Mike Zuranski, Tom Yoksas, Stonie Cooper, Steve Emmerson, Mike Schmidt, Jeff Weber

Executive Summary

Unidata continues to support, update, and enhance the data available via the IDD for the benefit of research and education. Included but not limited to adding new data formats, bridging the knowledge gap in newly introduced data, and providing statistics of data flow and composition.

Questions for Immediate Committee Feedback

Unidata is looking into the possibility of repurposing several current LDM Feed Types. **Is anyone using any of the GPS, GPSSRC, FAA604, FNMOC or WSI feeds?** Would you have any negative impacts if we repurposed these feeds for other uses?

Activities Since the Last Status Report

Internet Data Distribution (IDD)

IDD data volumes continue to increase especially when new datasets are made available.

The following output is from a Linux-based data server that the UPC operates on behalf of the community, lead.unidata.ucar.edu:

```
20231016
```

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

```
Maximum hourly volume 131656.311 M bytes/hour
```

```
Average hourly volume 84361.581 M bytes/hour
```

```
Average products per hour 516181 prods/hour
```

Feed	Average		Maximum	Products
	(M byte/hour)		(M byte/hour)	number/hour
CONDUIT	17436.980	[20.669%]	59207.864	96062.196
SATELLITE	15340.787	[18.185%]	20318.566	6676.370
NIMAGE	9254.217	[10.970%]	13105.174	7751.674
NEXRAD2	8793.045	[10.423%]	11522.887	102478.696
NGRID	6986.409	[8.282%]	11976.030	67110.783
HDS	5053.065	[5.990%]	10176.131	31136.348

EXP	4842.265	[5.740%]	8720.515	26837.978
FSL2	4823.666	[5.718%]	8383.821	5794.283
FNEXRAD	4219.474	[5.002%]	4586.080	9867.413
NEXRAD3	3191.932	[3.784%]	4233.136	82942.978
GEM	3021.219	[3.581%]	13356.547	5104.109
UNIWISC	1014.499	[1.203%]	1150.341	968.543
NOTHER	296.262	[0.351%]	750.932	60.935
IDS DDPLUS	86.732	[0.103%]	101.062	72969.043
LIGHTNING	1.031	[0.001%]	2.439	419.391

Data Distribution:

IDD CONDUIT feed:

NCEP migrated their CONDUIT systems to RHEL 8 during the week of 10-9-2023. As a result, the duplicate NDFD set has gone away. An example of the pattern that remains is below:

```
/common/data/products/NDFD/grib2/YEUZ98_KWBN_151952.grib2
!grib2/nwstg/NWS_0/#000/202310152000F052/TMPK/2 m HGHT! 000040
```

Unfortunately, NCEP's migration had the ill effect that dates, variable and level names are not being included in the product IDs. Mike Zuranski opened a ticket with their dataflow team on 10-16; they were quick to respond, they believe they know what the issue is and are discussing the next steps. Hopefully these issues will have been resolved by the time of the committee meeting.

IDD FNEXRAD NIMAGE and UNIWISC feeds:

We continue to create the content for the FNEXRAD (NEXRAD Level III national composites), NIMAGE (GOES-East and -West Level 2 images and products, fully reconstituted images from NOAAPort tiles and with broadcast headers and footers stripped off to leave "bare" netCDF4 files), and UNIWISC (select GOES-East and -West images converted to McIDAS AREA format for use in legacy systems like GEMPAK) feeds.

Experimental HRRR feed to eventually be replaced by RRRS:

Unidata is ingesting a subset of RRRS (Rapid Refresh Forecast System) products from the NOAA/GSL LDM feed. Last Mike Zuranski checked, it was limited to the first three forecast hours, though some sub-hourly data was there as well. This data is available on the FSL2 feed at hrrr.unidata.ucar.edu. The data is made available at the discretion of NOAA/GSL and may change or be removed at any time without notice.

Existing Data Distribution:

The primary top level IDD relay cluster, `idd.unidata.ucar.edu`, has been operating well since its move to the NCAR Wyoming SuperComputer (NWSC) facility in Cheyenne, WY.

The data volume seen in the **SATELLITE** (which is known as **DIFAX** in LDM distributions prior to v6.13.6) listing above represents all products received in the GOES ReBroadcast (GRB) downlinks that we installed in UCAR (currently GOES-18 at the NCAR Mesa Lab and GOES-16 at UCAR Foothills Lab 2). The data volume seen in the **NIMAGE** entry represents GOES-East/West ABI Level 2 imagery that has been reconstituted by stitching together tiles that are distributed in NOAAPort and all other Level 2 products. In both cases, binary headers and footers that are added to products before distribution in NOAAPort have been stripped off leaving "raw" netCDF4 files. The **UNIWISC** feed represents the volume of 3 select channels (0.64um VIS, 6.2um WV and 10.3um IR) for all coverages (CONUS, FullDisk, Mesoscale-1 and Mesoscale-2) of GOES-East/West image products that are in PNG compressed McIDAS AREA format that is suitable for use in GEMPAK, the IDV and McIDAS-V, McIDAS-X, and AWIPS.

Challenges, problems, and risks:

More sites, including UCAR, are installing intrusion detection/prevention systems (e.g., Palo Alto), which can adversely affect LDM throughput if not configured correctly.

Ongoing Activities

We plan to continue the following activities:

- Unidata distribution of GPS radio occultation products from COSMIC was resumed on August 7th, 2023.
- Many, but not all, of the products in NCEP operational HRRR are being distributed in the NOAAPort SBN and relayed in the IDD NGRID feed. Fire weather products (HRRR Smoke) that are being made available by NOAA/GSL in an EXP feed were added to the set of HRRR products that are available from `hrrr.unidata.ucar.edu`. All of these products along with with other model output are available via the TDS and Unidata AWIPS EDEX:
- Other data sets we continue to explore with NOAA/GSD/ESRL are:
 - [FIM](#)
 - [HIWPP](#)
 - [RRFS](#)

NOAAPort Data Ingest

- Ingest of the DVBS-2 NOAAPort Satellite Broadcast Network (SBN) products and their relay to end-users via the IDD has been “operational” at the UPC since August 2014.

Considerable effort has been expended in streamlining our NOAAPort ingest systems and assisting sites (UWisc/SSEC, NOAA/GSL, NOAA/SPC, Fox13 TV) in troubleshooting problems being experienced in their systems.

- The NOAAPort-derived data streams (**HDS, IDS|DDPLUS, NGRID, NIMAGE, NEXRAD3** and **NOTHER**) are redundantly injected into the IDD at five geographically separate locations: UCAR/Unidata, UWisc/SSEC, LSU/Climate, Allisonhouse.com and Fox13 TV in Tampa, FL.
- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.15.0.

Relevant Metrics

- Approximately **506** machines at **177** sites are running LDM-6 **and** reporting real-time statistics to the UPC.

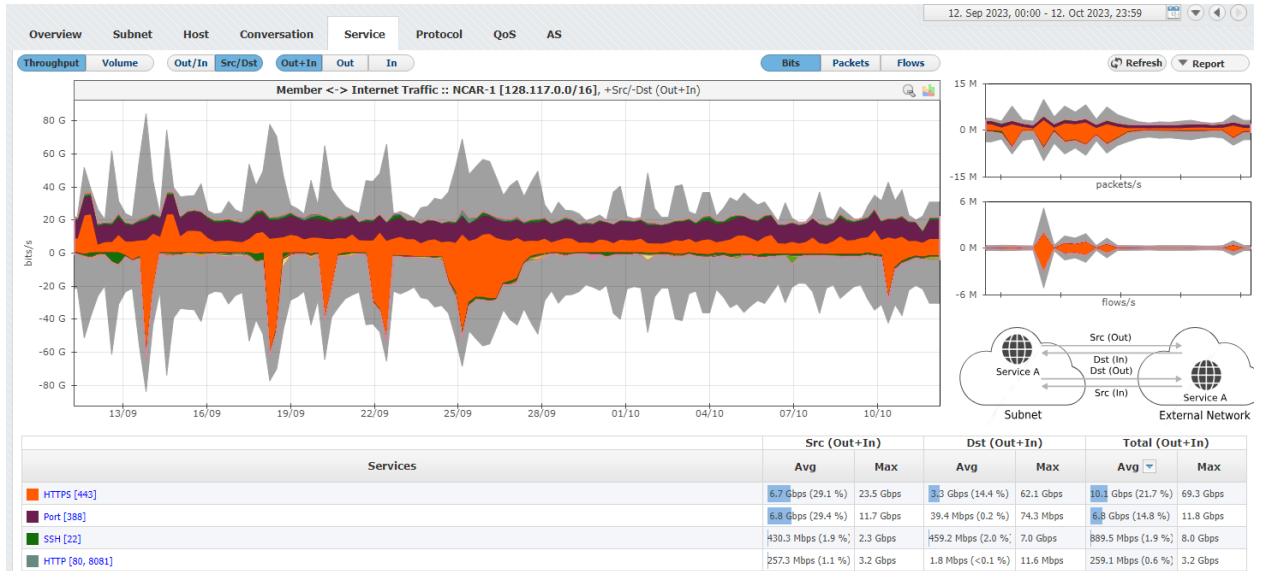
We routinely observe that the number of sites reporting real-time statistics fluctuates. We are not 100% certain why this may be the case, but our best guess is that some sites do not keep their LDMs running all of the time; campus firewall adjustments block the sending of the statistics; and/or sites decide to stop sending statistics. The latter possibility seems to be happening more frequently.

NB: We know that there are a number of sites that are participating in the IDD, but are not reporting real-time statistics back to us. Reporting of real-time statistics is not and never has been mandatory.

Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service. We have learned about sites not sending real-time statistics during these kinds of support activities, and a number of times the impediment to sending in stats is firewall configurations at the user sites.

- A number of organizations/projects continue to 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 companies, etc.).
- UCAR IDD toplevel relay clusters, **idd.unidata.ucar.edu** and **iddb.unidata.ucar.edu**

The IDD relay clusters, described in the June 2005 CommunitE-letter article Unidata's IDD Cluster, routinely relays data to more than 1250 downstream connections. The primary IDD relay cluster, **idd.unidata.ucar.edu**, was moved to the NCAR/Wyoming Super Computing facility in Cheyenne, WY in late August 2019.



Over the period from September 10, 2023 through October 13, 2023 (IDD volume snapshots are taken during periods that do not have monitoring dropouts in NetVizura plots) the average volume of LDM/IDD data flowing through the Front Range GigaPop averaged around 6.8 Gbps (~73.44 TB/day), and peak rates reached 11.7 Gbps (which would be ~126TB/day if the rate was sustained (which it is definitely **not**)).

The following table of volume snapshots shows that the volume of data flowing to downstreams out of UCAR has been reasonably consistent:

Date range	Src		Dst		Total	
	Ave	Max	Ave	Max	Ave	Max
20200508 - 20200630	5.4	7.5	42.1	52.9	5.5	7.5
20200701 - 20200930	5.4	7.9	41.9	60.3	5.4	7.9
20201001 - 20201231	5.2	6.9	39.9	55.9	5.3	7.0
20210101 - 20210331	5.5	8.0	42.3	59.9	5.5	8.1
20210401 - 20210415	6.1	15.5	46.4	112.7	6.1	15.7
20210601 - 20210719	6.6	9.2	50.5	73.0	6.6	9.2
20210908 - 20211005	7.6	14.9	59.3	121.7	7.7	15.0
20211101 - 20211231	6.7	9.1	52.4	71.4	6.8	9.2
20220208 - 20220311	6.6	15.2	53.5	114.8	6.6	15.3

20220412 - 20220521	7.2 14.5	52.6 103.7	7.3 14.6
20220717 - 20220831	7.3 13.3	46.3 86.1	7.3 13.4
20220714 - 20230313	7.8 11.7	51.1 77.4	7.8 11.7
20230910 - 20231013	6.8 11.7	39.4 74.3	6.8 11.8

NB: The units for Src and Total Ave and Max are Gbps (gigabits per second), and the units for Dst are Mbps (megabits per second).

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

The IDD project demonstrates how sites can employ the LDM to move and process data in their own environments.

2. **Providing Useful Tools**

The freely available LDM software and the IDD project that is built on top of the LDM have served as a demonstration for distribution of real-time data for a variety of organizations including the U.S. National Weather service.

The cluster approach for LDM/IDD data relay that Unidata pioneered has been adopted by several Unidata university sites, and is currently being implemented at U.S. government sites.

Unidata's NOAAPort ingest package, which is bundled with LDM-6, is being used by a variety of university, U.S. government, and private sector entities.

Both the LDM and NOAAPort ingest packages are bundled with AWIPS.

3. **Supporting People**

The IDD is the primary method that core Unidata sites use to get the meteorological data that they need. Providing access to data in near real-time is a fundamental Unidata activity. The IDD-Brasil, the South American peer of the North American IDD, and IDD-Caribe, the Central American peer of the North American IDD, are helping to extend real-time data delivery throughout the Americas

Status Report: IDV with RAMADDA

April 2023- November 2023

Yuan Ho, Julien Chastang

Executive Summary

We continue to support, update, and enhance the 3D data visualization and analysis tool IDV for our community. Our current activities include: coordinating with netCDF-Java group to add new data formats, collaborating with the SSEC developers to enhance the VisAD library, and working with our community to promote the usage of the IDV in research and education.

Questions for Immediate Committee Feedback

We have noticed that many advanced features of the IDV, such as formulas and trajectory displays, have not been widely used in the community and many data servers that the IDV can directly access are less well known to IDV users. We would like to provide help to classes, research groups and project teams to use these resources. Can committee members help to establish such connections?

Activities Since the Last Status Report

IDV Releases

The IDV 6.2 was released in September 2023.

The IDV 6.2u1 was released in October 2023.

IDV System Changes

__IDV Certificates__

Java Webstart, Windows app and MacOS certificates have been renewed and will be valid until at least May 30, 2021 (MacOS certificate is valid until 2024). Moreover, as properly signing the IDV under these different environments can be an involved process, this information has been thoroughly [documented here](#).

__Changes to nightly release that will eventually be incorporated into into stable version__

- IDV uses the latest Java 8 AdoptOpenJDK
- IDV employs latest Java3D (1.6.2)
- Updated the IDV code signing certificates on all platforms (i.e., MacOS, Windows, Webstart)
- IDV now “notartized” on MacOS
- Updated Unidata's Install4J license from version 5 to 8.
- Updated the IDV Install4J configuration.

IDV Display Changes

__Level II Radar Grid Displays __

The new and enhanced feature of the Level 2 radar grid display reads radial parameters from the Level 2 volume scan dataset and automatically generates 3D Cartesian grid parameters. The Integrated Data Viewer (IDV) employs an 8-point weighted interpolation method for this purpose. Efficient search algorithms are utilized to identify the 8 potential valid numbers of data points from 4 available surrounding rays. These algorithms significantly contribute to expediting the process. The resultant 3D Cartesian grid parameters can be showcased through various visualization techniques such as volume rendering, iso-surface representation, cross-sectional display, and plan-view presentations.

__ Point Clouds and Point Grids __

The IDV has introduced a novel data source type named "Point Cloud Data" exclusively tailored for CF netCDF station time series feature data. This innovation facilitates the presentation of time series point clouds and point grids for such point feature data. The IDV's point cloud display effectively visualizes substantial quantities of point data (often numbering in the millions) by portraying them as a collection of points, each colored based on its corresponding parameter. Additionally, the point grid function adeptly interpolates the point station data onto a 2D regular grid, thereby furnishing a comprehensive array of visualization and analytical tools for this multidimensional dataset. This includes scatter displays as well as associated statistical analyses.

__New Boundary Layer Parameters__

The planetary boundary layer (PBL) encompasses the volume of the lower troposphere which trades mass, energy, and momentum between the surface and free troposphere via turbulence. The algorithm was developed to identify PBL depth via two methods: by locating the potential temperature inversion above the surface (temperature gradient method) and by finding altitude corresponding to the critical bulk Richardson number (Ri method) which identifies PBL based on the turbulence profile. With the contribution of 2022 summer intern

Hassanpreet Kaur Dhaliwal , the algorithm was verified and integrated into the IDV.

__Latest Version of VisAD__

The SSEC team at UW, Madison has made a number of improvements to support 3D trajectories.

IDV Community Support

In the hybrid environment of in person and remote-learning system as a result of the COVID-19 pandemic, we keep helping universities and research institutes to run Unidata's Integrated Data Viewer (IDV) at home.

- Yuan provided a remote IDV training class to students of the satellite meteorology course from the University of Millersville in September 2022.
- Yuan collaborated with the research project team including UPC, MMM, Univ. of Hawaii, and CMB (Central Weather Bureau of Taiwan).
- Yuan provided a remote IDV training class to the Universitat de Barcelona students in March 2022
- Yuan provided a remote IDV training class to the weather briefing class of Florida Tech in October 2022.
- Yuan worked with Jeff Weber to provide support to the NSF funded project "The Indigenous Data Governance in Open Data Working Group".
- Yuan mentored this year summer intern Hassan and developed a few functions of calculating the boundary layer depth.

MSU IDV Project

In collaboration with MSU Department of Education, the project aims to enhance the educational experience of students and researchers by developing and implementing advanced data visualization and mathematical modeling tools using the Integrated Data Viewer (IDV).

IDV Publication Highlights

[Synoptic–Dynamic Meteorology in 3D: Introducing an IDV-Based Lab Manual](#) by Gary Lackmann, B. Mapes and K. Tyle

A [Google Scholar Search](#) reveals a number of publications that cite use of the IDV ([doi:10.5065/D6RN35XM](https://doi.org/10.5065/D6RN35XM)).

IDV and RAMADDA Training, Conference Attendance and Presence

__2023 AGU Fall Meeting__

- Pre and Post-processing of Machine Learning Datasets with UNIDATA's IDV

__2024 AMS Annual Meeting__

- Three Dimensional Gridded Visualization and Analysis of individual NEXRAD Level2 Volume Radar Data in Unidata's IDV

Ongoing Activities

We plan to continue the following activities:

__Investigation of Java 3D Alternative__

Because of concerns about the long-term viability of the open-source Java 3D project, the IDV team has begun discussions with our University of Wisconsin, SSEC collaborators to replace Java 3D with a more viable alternative within the VisAD API. We have started investigating whether the [Ardor 3D](#) can meet that objective. Looking into alternatives to Java 3D was also a goal described in the [Unidata 2018 Five-year plan](#).

New Activities

Over the past few months, we plan to organize or take part in the following:

We plan to upgrade the version of OPenJDK Java. This change will necessitate in depth testings and the IDV building and distribution workflow.

Relevant Metrics

__E-Support__

The IDV team continues to provide the geoscience community with high-quality support through e-support software and idv-users mail list. In the last half year the IDV team has closed ~40 e-support tickets. Each individual ticket may and often does involve many back-and-forth messages. There is an especially large number of support requests coming from international users.

Top ten universities running IDV are: Millersville, Oklahoma, University of Utah, St Cloud state, Plymouth, NC State, West Kentucky, Lyndon State, University of Illinois, and San Francisco State.

__GitHub Pull Requests__

In the area of greater collaborative development, since the migration of the IDV project to github, we have closed a total of 125 “pull requests” or code contributions from internal and external collaborators.

__Youtube IDV Instructional Videos__

In the area of online IDV training, the Youtube IDV instructional videos have been viewed thousands of times.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing 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. **Providing Useful Tools**

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. **Supporting People**

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.

Status Report: Information Technology

April 2023- November 2023

Mike Schmidt, Matt Perna, & Jennifer Oxelson

Executive Summary

Our role is to maintain and enhance the productivity of the staff and assist with the resolution of issues in service to the community. Primarily, that consists of keeping end-user and developer systems secure, and keeping servers and services highly available, patched, and operational for the community. This report is informational and there are no pressing issues.

Questions for Immediate Committee Feedback

Unless committee members or the community are experiencing performance issues that we could help resolve, no other feedback is requested.

Major Activities

- Unidata staff have been installing/vetting new support ticketing systems to replace the currently used eSupport package.
- UCAR continues to migrate select services to a centralized model and Unidata is involved in implementing the service(s) for our systems and users. Recently completed and/or anticipated are;
 - centralized DNS (InfoBlox, complete)
 - centralized backups (CrashPlan, mostly complete)
 - centralized telephony agent (TBD, soon)
 - centralized end-user security agent (TBD, soon)

As of writing this, we have implemented 2 MDM solutions to manage our Macbooks as well as our remote windows clients. Firewalls, antivirus and overall configurations are monitored constantly and report back when there is an anomaly on both Operating systems. Cloud backups have been implemented on all Unidata client workstations for the past year and we see a foreseeable upgrade in the way we back up clients to the cloud in the coming year. Disk encryption will be enforced across the board on all client operating systems. Unidata IT has the ability to remote manage any Unidata client workstation if it is connected to the internet.

This is probably the third consecutive report proclaiming UCAR is nearing completion of their Mesa Lab Data Center (MLDC) co-location facility upgrade. There have been significant procurement and supply chain delays that have mostly been resolved at this point. We have been putting pieces in place to facilitate the relocation of our MLDC based servers when the time comes. There may be the need to roll some of our Internet based services to the NWSC in Cheyenne for extended outages.

Daily, we continue efforts to keep services and systems secure which takes consistent attention and occasional herculean efforts (to patch everything all at once). UCAR has embarked on a number of new initiatives to segment the network into smaller and smaller zones and gain a more dynamic inventory of assets on the network. Unidata continues to play a role in these efforts.

We continue to maintain a LDM7 test node at the Front Range GigaPOP (FRGP) just off downtown Denver in co-location with the major backbone networks supporting FRGP participants (UCAR, ..). We expect to support intensive data movement and LDM testing for the next few years on this effort.

Ongoing Activities

We plan to continue the following activities:

- Day-to-day system and network support to the community as needed
- Resolve daily staff help desk issues
- Maintain security profile and exceed UCAR security standards
- Following UCAR directives regarding DNS and Palo Alto zones centralization

Prepared *October 2023*

Status Report: LDM

April 2023- November 2023

Steve Emmerson, Tom Yoksas, Mike Schmidt, Stonie Cooper

Executive Summary

Unidata's LDM team continues to update source code and operating paradigms with ever-changing user demographics and user requirements, particularly in the area of security and inclusiveness of data.

Questions for Immediate Committee Feedback

As intimate stakeholders in data distribution and access for research and education, please let us know about any data distribution software needs that you personally experience that are not currently fulfilled by use of LDM.

Activities Since the Last Status Report

The LDM is the primary software package by which research and education institutions obtain near real-time meteorological and related data.

Progress has been made on the following:

- Build: Modified process to accommodate non-standard installations
- ldmadmin(1): Improved the "showsettings" reporting command
- pqinsert(1):
 - Added the ability to read a single data-product from standard input
 - At NOAA's request, added a "-r" option to use regular I/O on the input files instead of memory-mapping them. This was necessary because the input files were not on the local disk.
 - Added a "-h" (help) option
 - Enhanced the man(1) page
- ldmd(1):
 - Added a warning about transmitting the oldest product in the queue because this can indicate that the queue is too small
 - Improved error messages
- notifyme(1):
 - Added a "-O" option to include the origin host in the output
 - Replaced the use of synchronous LDM-5 protocols with asynchronous LDM-6 protocols in order to greatly increase performance when used for a long time

- Eliminated a latent SIGSEGV bug when the upstream LDM encounters a severe error
- NOAAPort(1):
 - Configured the UDP socket to allow its input to be shared with other processes
 - Replaced the Python-based noaaportBlender(1) program with a shell-based one to greatly reduce the number of created processes
- rtstats(1): Improved robustness
- Documentation: Made slight improvements and corrections
- Released version 6.15.0
- Support:
 - Answered many questions from Universities, NOAA, US Military, and corporations.
 - Gave a virtual workshop on the LDM to NCEP Central Operations of NOAA/NWS.

Dependencies, challenges, problems, and risks include:

The NOAAPort component of the LDM is sometimes held responsible for decisions made by the NWS when they don't follow their own policy on how to categorize and name data products (not a new challenge).

More sites are installing intrusion detection/prevention systems (e.g., Palo Alto), which can adversely affect LDM throughput if not configured correctly (again, not new).

Ongoing Activities

We plan to continue the following activities:

- Support and maintain the LDM
- Convert the LDM build process from the current one based on GNU automake(1) to one based on CMake. Beside being cleaner, this will enable the automatic creation of binary distributions (e.g., RPM and DEB files).
- Stonie Cooper is gradually assuming responsibility for the LDM

Relevant Metrics

- [Data on LDM downloads](#)
- The LDM system at the Unidata Program Center powers the Unidata IDD (Internet Data Distribution) system. Metrics on that program can be found in the IDD status report.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**
By enabling researchers, teachers, and students to process a wide variety of meteorological and related data in near real time.

2. **Providing Useful Tools**

By enabling researchers, teachers, and students to obtain a wide variety of meteorological and related data in near real time and at no cost via the Internet.

By using the LDM to move data into the cloud and developing multicast technologies.

3. **Supporting People**

By answering support questions, writing documentation, and conducting workshops.

Prepared *October 2023*

Status Report: McIDAS

April 2023- November 2023

Tom Yoksas, Mike Zuranski

Executive Summary

Unidata McIDAS-X/-XCD development has officially ceased. Our MUG membership with SSEC will be maintained; Unidata sites can continue to download and use Unidata McIDAS-X. SSEC will continue to support their McIDAS-X software through at least the GOES-R series (through December 2036).

Future updates to Unidata McIDAS-X/-XCD will be few and far between, and will primarily be done at SSEC rather than in-house at Unidata. These updates will likely be limited to critical fixes and possibly small enhancements.

Users should continue to contact support-mcidas@unidata.ucar.edu for download requests. We will continue to host the mcidas-x@unidata.ucar.edu email list.

This will be the final McIDAS status report.

Activities Since the Last Status Report

Aside from routine updates/bug fixes to existing code and tables, the main area of activity has been enhancing the GOES-R servers.

Other recent changes and other news:

- McIDAS-X updated for new NAAPort Level 2 products
- Culling of now archaic ADDE datasets on our public-facing ADDE server instances
- Mike Zuranski and Tom Yoksas attended the McIDAS Users' Group (MUG) Members Meeting on September 25-28. Mike gave a presentation on McIDAS-X usage at College of DuPage which was very well received. Mike also formed numerous relationships at SSEC, timely considering Tom's retirement...
- Tom retired from Unidata on September 16th. He plans on continuing to collaborate with SSEC contributing to McIDAS-X development.

Current/Ongoing Activities

- Support use of McIDAS-X/XCD in-house

- Ensure that the Unidata instances of McIDAS ADDE continue to function efficiently (ADDE serves on the order of 1.5 TB/day from three servers that Unidata operates on behalf of the community)
- McIDAS-X is used to convert GOES-East/West ABI imagery that is in netCDF4 format to McIDAS AREA format that is usable by all supported display and analysis packages (except Python/MetPy) for the Unidata-Wisconsin (**UNIWISC** aka **MCIDAS**) IDD feed.

New Activities

- Mike is working on making updated map files for McIDAS. Many of the maps, especially the ones added by Unidata years ago, are significantly out of date. There have been several user requests for updated maps, though SSEC hasn't updated theirs in a while either.
- With code contributed by SSEC, Mike is building a utility to update numerous base maps in bulk:
 - New GIS data is downloaded, their URLs saved in the scripting which scrapes the latest data from official sources.
 - Maps that will be updated include:
 - County, forecast/fire zones and CWA boundaries
 - Roads, rivers, railroads and Canadian provinces
 - Newly added marine zones, including offshore and high seas zones
 - AWIPS Basemap map files are complete, but may be optimized further
 - Will include a feature for users to convert locally stored shapefiles for custom maps
- SSEC has received a preliminary copy of some of the new maps for evaluation, and are excited that this utility is coming.
- This project is currently a low priority. The goal is to have this completed by the end of the year. Completion of this project may be delayed by more urgent matters if they arise.

Relevant Metrics

- Data delivered by the Unidata McIDAS ADDE servers exceeds 1.7 TB/day. The great majority of the data being served is imagery from GOES-16/17/18.
- See the **Status Report: Support** for 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. Recently, the ldm-mcidas code was moved to GitHub.

Geostationary Satellite Data Ingest and Data Serving

Unidata continues to ingest GOES-East and GOES-West imager data at the UCAR Foothills Lab and NCAR Mesa lab campuses in Boulder.

- Direct, programmatic access to real-time GOES-East (GOES-16) and GOES-West (currently GOES-1) data via McIDAS ADDE services on three publicly accessible servers (lead.unidata.ucar.edu, atm.ucar.edu and adde.ssec.wisc.edu) has been averaging on the order of 1.6 TB/day for the past two years. **Ongoing Activities**
- Continued ingest, distribution via the IDD and ADDE serving of GOES-East and GOES-West imagery from the GRB downlinks we installed in UCAR
- Continued ingest and ADDE serving of GOES-14 imagery when available. GOES-14 remains in its standby location (104W) and is turned on for periodic testing as needed. GOES-17 has been replaced by GOES-18 as the West geostationary platform, and it was recently drifted to approximately 105W where it will function as an in-orbit standby for GOES-16 and GOES-18.

These efforts require maintenance of the satellite ingest and data serving equipment.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

Remote, programmatic access to data provided by the Abstract Data Distribution Environment (ADDE) environment of McIDAS has been a model for the development of remote access methodologies since 1994. 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. ADDE servers operated by Unidata are currently serving in excess of 1.6 TB/day.

2. **Providing Useful Tools**

McIDAS remains the application of choice for the satellite meteorology community. The Abstract Data Distribution Environment (ADDE) component 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.

The fifth generation of McIDAS, McIDAS--V, unlike its predecessors, is a fully open source application that is in wide scale and growing use in the worldwide satellite meteorological community

McIDAS ADDE continues to evolve and provide access to a rapidly increasing volume of imagery and non-image data.

3. **Supporting People**

McIDAS is still in active use by those interested in satellite meteorology worldwide.

Prepared *October 2023*

Status Report: netCDF

April 2023- November 2023

Ward Fisher , Dennis Heimburger , Hailey Johnson , Ethan Davis

Executive Summary

The netCDF team continues to work towards maintaining the reliability of the netCDF libraries, while keeping one eye forward as to the future needs of our community. We have continued our community engagement efforts and collaborations whenever and wherever possible; examples of this include our involvement with the Zarr Community meetings and our membership on the Zarr Enhancement Protocol (ZEP) committee. We have also expressed interest in being part of the forthcoming reboot of the HDF Group's Technical Advisory board.

We continue to address the issues associated with the proliferation of new mainstream architectures (Apple Developed M1-M2/ARM), evolving compilers and standards, and extending our collaborations with tangential, but related, projects (conda-forge libnetcdf feedstock, for example).

Questions for Immediate Committee Feedback

No questions at this time.

Activities Since the Last Status Report

Snapshot of NetCDF Development Status

We are using GitHub tools for C, Fortran and C++ interfaces to provide transparent feature development, handle performance issues, fix bugs, deploy new releases and to collaborate with other developers. Additionally, we are using docker technology to run netCDF-C, Fortran and C++ regression and continuous integration tests. We currently have **260** open issues for netCDF-C, **103** open issues for netCDF-Fortran, and **55** open issues for netCDF-C++. The netCDF Java interface is maintained by the Unidata CDM/TDS group and we collaborate with external developers to maintain the netCDF Python interface.

In the netCDF group, progress has been made in the following areas since the last status report:

- Support for Amazon S3 access via libnetcdf (using either the Amazon S3 SDK library, or an internal interface layer) has been improved.
- The netCDF and netCDF-Java teams continue to participate in the Zarr Community meetings, in order to help guide the development of the Zarr v3 and future specifications in a way that promotes broad compatibility across Zarr implementations.

- The netCDF and netCDF-Java teams have also joined with the Zarr Enhancement Protocol (ZEP) committee, in an effort to help codify the process by which features are added to the Zarr v3 specification.
- The release of NCZarr (netCDF with native Zarr support) has been improved as of netCDF-C version 4.9.2.
- Continuing improvement for the NUG: We previously migrated the NetCDF User's Guide to a new, separate repository. This repository will contain the concise, language-agnostic summary of the netCDF data model. Language-specific documentation (primarily used by developers) will remain associated with the individual code repositories.
- Further enhancements to the netCDF-C documentation, modernization of the netCDF-Fortran and netCDF-C++ documentation.
- We continue to see a high volume of contributions to the netCDF code base(s) from our community, for which we are grateful. While these contributions require careful review and consideration, it is encouraging to see this model of development (enabled by our move to GitHub) being more fully embraced by our community.
- Improvement and collaboration on additional filter and plugin support for dynamic, selective compression, based on work contributed by Charlie Zender and Ed Hartnett.
- As a result of increased interest, the DAP4 functionality has been significantly improved. A corresponding set of changes was propagated to the NetCDF-Java code base. Some discrepancies in the DAP4 specification were discovered, and resolution is on-going.

Dependencies, challenges, problems and risks include:

- The small group of netCDF developers is under a lot of pressure to provide project management as well as implement new features, fix bugs, provide support, etc. With 1.5 FTE assigned to the project, the workload is significant.
- Difficult issues which require intense debugging can bog down progress in other areas of netCDF and related projects.
- Rapid evolution of the Zarr standard is very useful, but also provides a bit of a moving target.
- Increase in external contributions has greatly increased the project management overhead for netCDF-C/C++/Fortran.
- Advances in compilers (GCC 10.x) and newer architectures (such as Apple's ARM M1/M2 architecture) are requiring additional overhead to ensure compatibility.
- The proliferation of cloud environments requires specific attention.

Ongoing Activities

We plan to continue the following activities:

- Continue work towards adoption of additional storage options, separating out the data model from the data storage format (as much as possible).
- Improve the messaging around the expanded functionality of netCDF.
- Provide support to a large worldwide community of netCDF developers and users.
- Continue development, maintenance, and testing of source code for multiple language

libraries and generic netCDF utility programs.

- Continue modernizing the documentation for netCDF-C, Fortran and C++ libraries.
- Extend collaboration as opportunities arise, for increasing the efficiency of parallel netcdf-3 and netcdf-4.

New Activities

Improved NetCDF/Zarr Integration

The netCDF team has now released multiple releases of netCDF-C which support the ncZarr protocol. This work has been well received, and we continue to make improvements. We are now focused on improving the S3 support for libnetcdf/ncZarr. Work continues in collaboration with the Zarr community group and the Zarr Enhancement Protocol group. **The netCDF team recognizes the need to improve messaging around the new functionality which has been implemented**, and will be working to make these features more widely known.

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

- Release iterative versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Continue modernizing/editing the netCDF documentation to provide easy access to documentation for older versions of netCDF.

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

- Release an official Windows port of the netCDF-Fortran and netCDF-C++ interfaces.
- Continue to encourage and support the use of netCDF-4's enhanced data model by third-party developers.
- Expand support for native object storage in the netCDF C library.
- Continue to represent the Unidata community in the HDF Technical Advisory Board process.
- Continue to represent the Unidata community in the Zarr/n5 collaboration conference calls.

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

- Improve scalability to handle huge datasets and collections.
- Improve the efficiency of parallel netcdf3 and parallel netcdf4.
- Continue to add support for both file-storage and object-storage options.

Relevant Metrics

Google Metrics

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, for comparison, are:

- **1,200,000** for netCDF-3
- **1,430,000** for netCDF-4
- **5,410** for ncZarr
- **2,180,000** for HDF5
- **174,000** for GRIB2
- **3,490,000** for ZARR

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

- **448** for netCDF-3
- **1,380** for netCDF-4
- **36** for ncZarr
- **42,700** for netCDF
- **25,000** for HDF5
- **1,760** for GRIB2
- **8,340** for ZARR

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

- 1. Managing Geoscience Data**
by supporting the 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.
- 2. Providing Useful Tools**
by developing netCDF and related software, and creating regular software releases of the C, C++ and Fortran interfaces; providing long-term support for these tools through the various avenues available to the Unidata staff (Github, eSupport, Stackoverflow, etc).
- 3. Supporting People**
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.

Prepared *October 2023*

Status Report: Python

April 2023 - November 2023

Ryan May, Drew Camron, Julien Chastang, Nicole Corbin, Thomas Martin

Executive Summary

Unidata's Python efforts continue to encompass: training on the use of Python for the community; development and maintenance of several tools for the community (most notably MetPy but also Siphon and data processing scripts); and participation within the broader scientific Python community. We continue to lead and support a variety of educational efforts, including our first on-site non-AMS regional Python workshop since the COVID pandemic. We are also furthering development of asynchronous training materials through Project Pythia, where we are working to migrate our existing workshop and gallery materials into a so-called "cookbook" within the broader project. MetPy development continues with the 1.5.0 feature release (largely based around plotting surface analyses) and a 1.5.1 bugfix release. We are planning a 1.6 release which will largely include community contributions, many of which were created at a sprint at SciPy 2023. Community use of the library is proving extensive, with 74 theses and peer-reviewed publications mentioning or citing MetPy in 2023; this brings the total count to over 250. We continue to assist the broader community with participation within matplotlib, cartopy, and conda-forge, though it has become increasingly difficult to dedicate time to these efforts given the full portfolio of responsibilities.

Questions for Immediate Committee Feedback

Nothing at this time.

Activities Since the Last Status Report

Python Training Efforts

Our content, expertise, and Science Gateway JupyterHub infrastructure led and supported multiple synchronous teaching/training workshops throughout 2024, including:

- 2023 Unidata Users Workshop session on Exploratory Data Analysis
- 2023 Pythia Hackathon for teaching and supporting Cookbook development
- 2023 UCAR Internship professional development workshop on data and viz
- 2023 I-GUIDE Summer School session on MetPy
- 2023 AMS Radar Conference session on MetPy
- 2-day Unidata *MetPy for Quantitative Analysis of Meteorological Data* workshop in support of UMBC, Howard, Morgan State, and the surrounding region
- 2-day *Python Readiness* workshop re-created by CSU graduate students after a Jupyter-focused train-the-trainer series from Unidata Instructional Design

We continue to support Project Pythia's Cookbook ecosystem, and the infrastructure and educational content it provides to the community. We directly engage with content

development, infrastructure and maintenance, and community engagement. We will support organizing the 2024 Pythia Hackathon. *MetPy Cookbook* development has begun with support from interns and community development.

Progress has been made on the following:

- In support of the offerings above, we have generated and refined our self-contained workshop content for easier offering and re-packaging for new learners and institutions
- AMS Annual Meeting 2024 organizers have admitted Unidata Python for both, in-person, a Student Conference workshop session and a half-day Short Course
- Unidata continues to be a primary collaborator on Project Pythia, particularly through support of the growing Cookbook ecosystem. Development has begun on the *MetPy Cookbook*
- John Leeman continues to lead the “MetPy Mondays” effort. We have engaged community members with plans to broaden the impact of MetPy Mondays content with code snippets and reference examples

MetPy

Development continues to be driven by requirements for our dedicated awards (in addition to bug reports and pull requests from community members). MetPy 1.5.0 was released in mid-May 2023, with a 1.5.1 bug fix release following up in early July 2023, which included:

- Added support for drawing lines and paths styled as fronts and other well-known boundary styles
- Added parsing of WPC surface analysis bulletins
- Enhanced CAPE/CIN to include virtual temperature buoyancy effects
- Feature examples of simplified plotting interface predefined areas within the documentation

Moving forward, 1.6 is planned for release late October/early November 2023. This release will feature a variety of community contributions for more calculations, including Corfidi MCS motion and Galvez-Davison index, as well as enhancements to the simplified plotting interface from Kevin Goebbert. We’ve also been working extensively to resolve some internal consistency issues, raised by multiple prominent users, around differing definitions of relative humidity. We have broadly been able to adhere to the more rapid release cadence, though the 1.6 release has slipped largely due to some resources being focused on Unidata’s next 5-year core award; we plan on the release cadence returning to normal moving forward. More broadly, we will also be continuing the performance improvement work that is the focus of the CSSI award, as well as include support for accessing data from the radar and satellite S3-based data archives.

On the community participation front, the “MetPy Community Call” continues to be challenging to get off the ground. The most recent call was not well attended outside the core development group (including many community members); much of this can be attributed to inconsistent scheduling of the meeting and we hope to improve this going forward. This remains a challenging focus area, though, given the variety of items that are

prioritized across the team's limited time. On a more positive note, due to strong interest from community members, a small MetPy development sprint was held at SciPy 2023 by Drew Camron. Approximately 5 members of the community participated, with numerous Pull Requests coming out of the sprint, which will be included in the next release of MetPy.

Progress has been made on the following:

- MetPy 1.5.0 released mid-May 2023
- MetPy 1.5.1 released early July 2023
- Work towards requirements of MetPy-related NSF awards
- Community awareness continues to grow, with the volume of engagement (especially support requests) and mentions on social media growing; the MetPy twitter account has reached 2742 followers (good growth despite the platform's overall instability)
- 74 theses or peer-reviewed publications have cited or mentioned MetPy so far in 2023; this is in comparison to 53 in all of 2022, 56 in 2021, and 43 in 2020. We just passed 250 total theses and peer-reviewed publications that cite or mention MetPy.

Siphon and Data Processing

Siphon continues to exist in a steady state—continued maintenance and use, but minimal feature advancement. Some of this is due to limited development resources being focused on MetPy's needs; it is also due to limited pressing needs on the data access side. Largely, Siphon meets the needs we have identified for Python data access (that aren't also already met by zarr, xarray, etc.). With that said, Siphon does remain an important part of the stack used by our training work, and by Unidata's community of Python users in general. The most pressing developments we anticipate for Siphon, outside of the need for a bit of focused maintenance, are improvements to working with Siphon in interactive sessions, like the Jupyter notebook environment: improved catalog crawling interface, better string representations, and tab completion. The decision has been made to separate **non-TDS functionality** (e.g. Wyoming Upper Air archive access) out into a new remote-access toolset contained within MetPy, and we hope to begin this transition work soon.

We also continue to maintain the LDM Alchemy repository as a collection of LDM processing scripts in Python. Currently this includes the code powering the AWS NEXRAD archive as well as the program that reconstitutes NOAAPORT GOES-16/17 imagery. As we transition more of our internal data processing to Python, this repository will hold those scripts. We have seen several community questions regarding both the GOES and NEXRAD processing software.

External Participation

The Python team attends conferences as well as participates in other projects within the scientific Python ecosystem. This allows us to stay informed and to be able to advocate for our community, as well as keep our community updated on developments. As participants in a broader Open Source software ecosystem, the Python team regularly encounters issues in other projects relevant to our community's needs. As such, we routinely engage these projects to address challenges and submit fixes. We also continue to host Jeff Whittaker's netCDF4-python project repository; Jeff continues to be the active maintainer of the project. The overall involvement helps ensure that important portions of our community's Python

stack remain well-supported. Ryan May continues to serve as a core developer for CartoPy as well as a member of Matplotlib's Steering Council and conda-forge's core team. *It should be noted, though, that it's becoming increasingly difficult to dedicate time to these efforts given the full portfolio of responsibilities on the team.*

Progress has been made on the following:

- We continue to engage with the [Pangeo](#) project, a grass-roots effort to develop a community stack of tools serving the atmospheric, oceanic, land, and climate science. This engagement is enhanced by work on the Pangeo EarthCube award, which will likely drive some contributions to the XArray project.
- Ryan May continues to work as a developer on the matplotlib and CartoPy projects, and as a member of conda-forge core team.
- We also continue to actively engage with the xarray and pint projects.

Ongoing Activities

We plan to continue the following activities:

- Supporting Unidata's collection of online Python learning materials
- Engaging in synchronous Python teaching opportunities, virtual or otherwise
- Growing Siphon as a tool for remote data access across a variety of services
- Growing and developing MetPy as a community resource for Python in meteorology
- Continued participation in the scientific Python community as advocates for the atmospheric science community
- Working with JupyterHub as a way to facilitate data-proximate analysis
- MetPy Mondays for engaging the community

New Activities

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

- Release first version of *MetPy Cookbook* with support from community and intern development
- Release MetPy 1.6 with a variety of features, including enhancements for the simplified plotting interface and community contributed calculations like Corfidi MCS motion and Galvez-Davison Index.
- Lead Short Course and Student Conference workshop at 2024 AMS Annual Meeting
- Engage in continued support of Project Pythia and adjacent UCAR Python education efforts
- Present on updates to MetPy at the 2024 AMS Annual Meeting

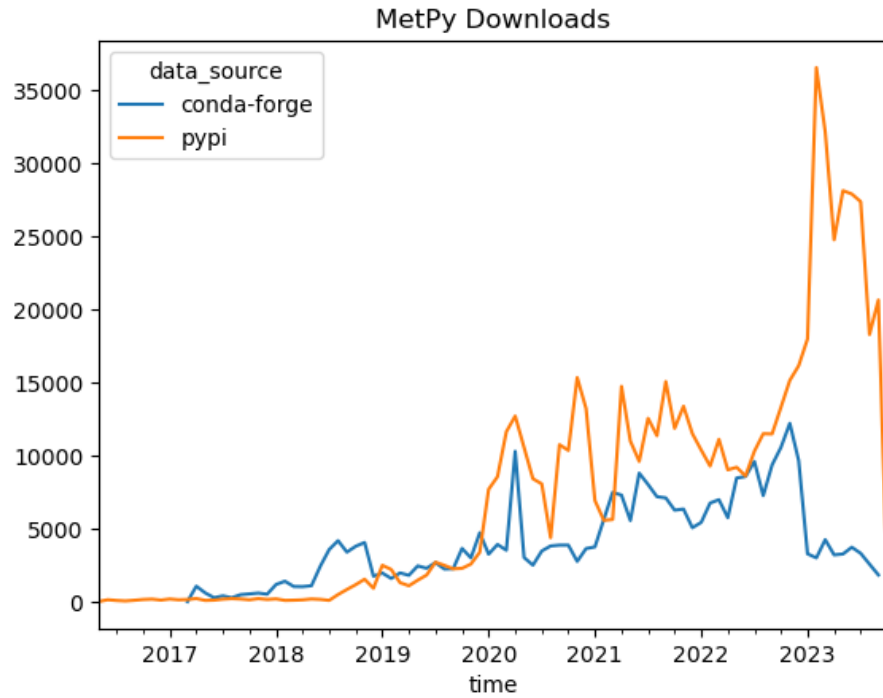
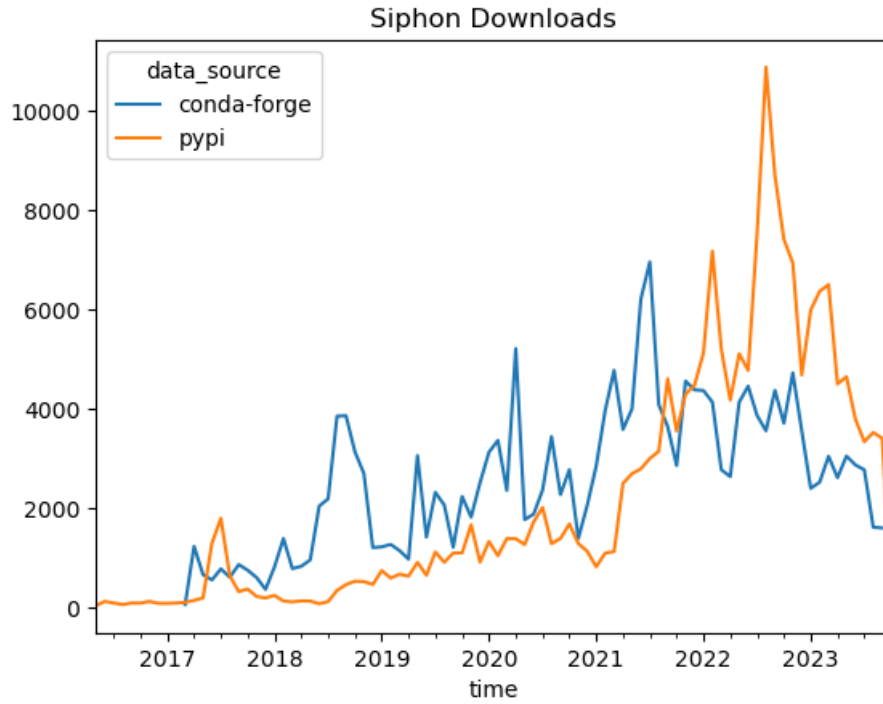
Over the next twelve months, we plan to organize or take part in the following:

- Offer additional virtual or in-person MetPy workshops
- Separate non-TDS siphon capability into new MetPy remote functionality
- Explore ways to leverage Web Assembly to provide MetPy as an in-the-web-browser

experience for users

Relevant Metrics

NOTE: conda-forge numbers are problematic due to upstream issues tracking downloads



MetPy

- 96% test coverage
- Watchers: 59
- According to GitHub, 431 repositories and 47 packages depend on MetPy
- Downloads for the releases made in the last year (Conda + PyPI):
 - 1.4.0: 50182
 - 1.4.1: 42644
 - 1.5.0: 34563
 - 1.5.1: 37202
- Since 1 April 2023
 - Active Issues: 80 (32 created, 35 closed)
 - Active PRs: 213 (194 created, 194 closed)
 - External Issue Activity: 25 opened, 74 comments
 - External PR Activity: 30 opened, 65 comments
 - Unique external contributors: 47
 - Stars: 92 (1099 total)
 - Forks: 2 (357 total)
 - Commits: 425
- Since 1 November 2022
 - Active Issues: 119 (73 created, 61 closed)
 - Active PRs: 349 (331 created, 328 closed)
 - External Issue Activity: 60 opened, 135 comments
 - External PR Activity: 37 opened, 96 comments
 - Unique external contributors: 60
 - Stars: 178 (1099 total)
 - Forks: 6 (357 total)
 - Commits: 749

Siphon

- 98% test coverage
- Watchers: 16
- According to GitHub, 172 repositories and 19 packages depend on Siphon
- Since 1 April 2023
 - Active Issues: 5 (5 created, 1 closed)
 - Active PRs: 72 (62 created, 58 closed)
 - External Issue Activity: 3 opened, 8 comments
 - External PR Activity: 2 opened, 4 comments
 - Unique external contributors: 5
 - Stars: 13 (192 total)
 - Forks: 0 (65 total)
 - Commits: 17
- Since 1 November 2022
 - Active Issues: 10 (8 created, 3 closed)
 - Active PRs: 119 (108 created, 104 closed)
 - External Issue Activity: 5 opened, 11 comments

- External PR Activity: 2 opened, 4 comments
- Unique external contributors: 10
- Stars: 20 (1192 total)
- Forks: 1 (65 total)
- Commits: 35

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Providing Useful Tools**

Python has become a key tool in the atmospheric sciences, and the geosciences in general. MetPy leverages the rest of the scientific Python ecosystem to provide a suite of documented and tested domain-specific functionality, supporting greater use of Python by the community. Siphon serves to provide access to the growing collection of remote data sets. Together, MetPy and Siphon give the community a platform for scripted analysis of real-time and archived weather data. These tools are also readily used in the Jupyter Lab/Notebook environment, for ease of use in cloud and HPC computing environments, facilitating data-proximate analysis. We also participate in a variety of projects in the broader scientific Python ecosystem, to help ensure the ecosystem's viability and that it continues to meet our community's needs.

2. **Supporting People**

We provide a variety of online training resources to facilitate our community's education and use of Python. We also regularly conduct training workshops to teach attendees how to use tools and apply them to their problems and challenges in research and education.

Prepared October 2023

Status Report: Support

April 2023- November 2023

Jennifer Oxelson, UPC Staff

Executive Summary

Unidata staff have started vetting replacement packages for our current eSupport system. We hope to have a new support package in use before the end of 2023.

With the retirement of Tom Yoksas, McIDAS development at Unidata has been curtailed, but support for the current software will continue for the immediate future. Unidata will continue to license the software on behalf of the community (via a MUG membership) for the foreseeable future as well.

Questions for Immediate Committee Feedback

What changes/recommendations would you like to see in Unidata support or how Unidata conducts support? Are you subscribed to any of the unidata software package mailing lists?

Activities Since the Last Status Report

Retirement of Tom Yoksas

Tom Yoksas has left for greener pastures (retirement), but will still be volunteering with Unidata Support efforts when applicable.

McIDAS development at Unidata has been curtailed, but support for the current software will continue for the immediate future. Unidata will continue to license the software on behalf of the community (via a MUG membership) for the foreseeable future as well.

Looking for a new in-house support package

- The currently-used eSupport package is long in the tooth. Unidata staff have started vetting replacement packages. We hope to have a new support package in use before the end of 2023.

Training

Unidata training/workshop information can be found in the [Community status report](#).

New Activities

In order to fulfill our objectives articulated in the Unidata 2018 Proposal, focused efforts are needed in two major areas:

- Enhance electronic support offerings
- [Create instructional materials for online virtual training](#)

Relevant Metrics

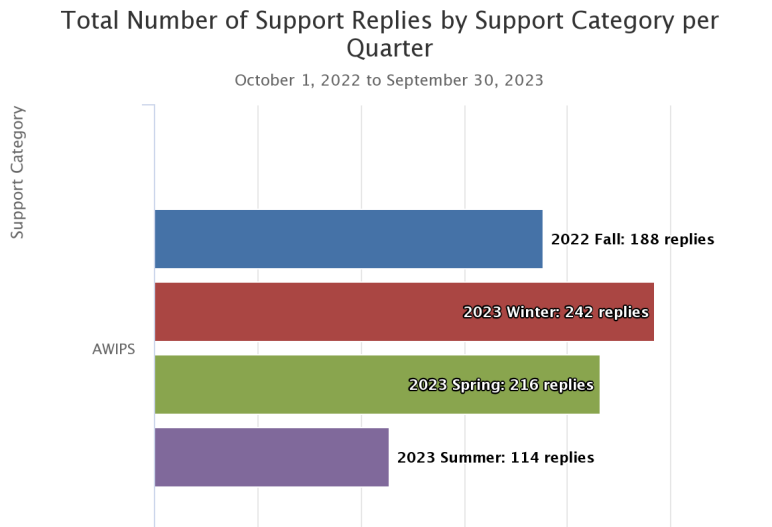
User support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system (eSupport). Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

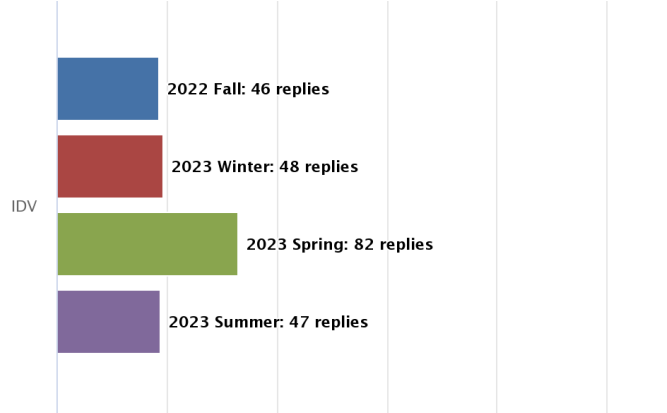
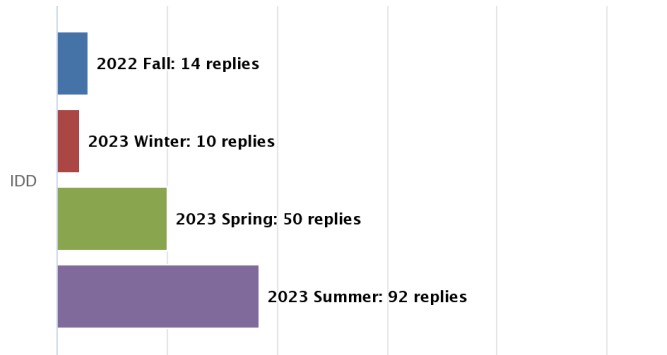
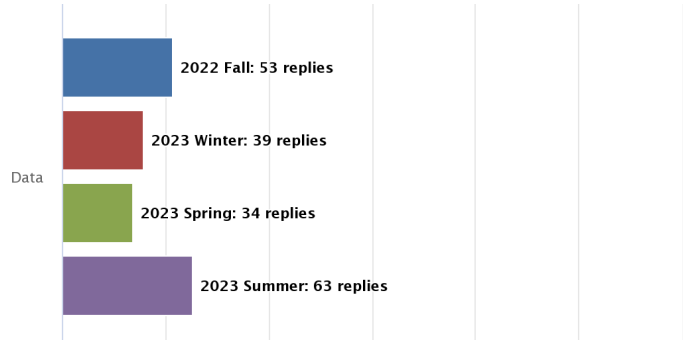
Additional metrics may be found in the [Comprehensive Metrics Data](#) portion of this meeting's agenda.

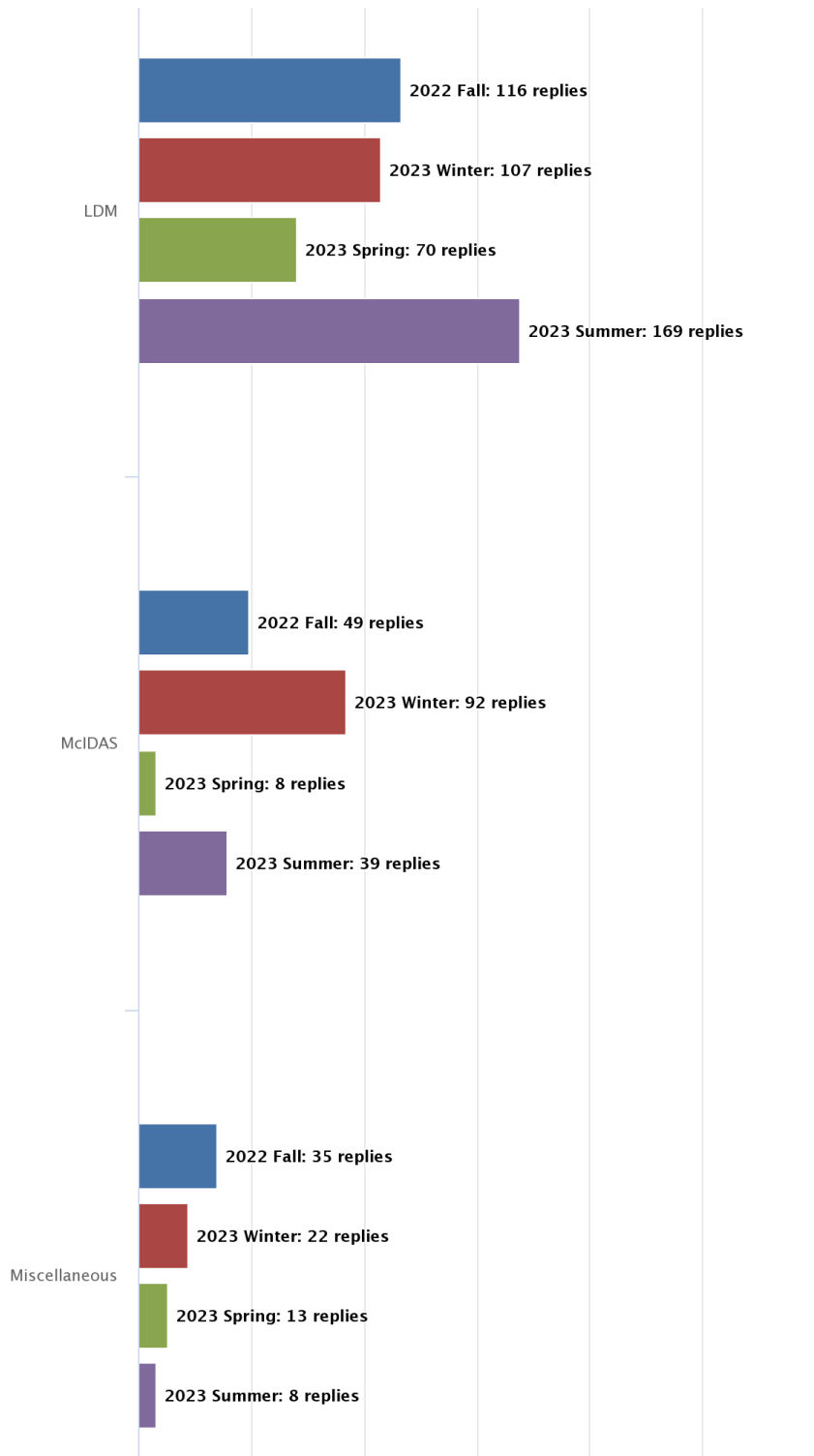
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 12 month period from **October 1, 2022 until September 30, 2023**.

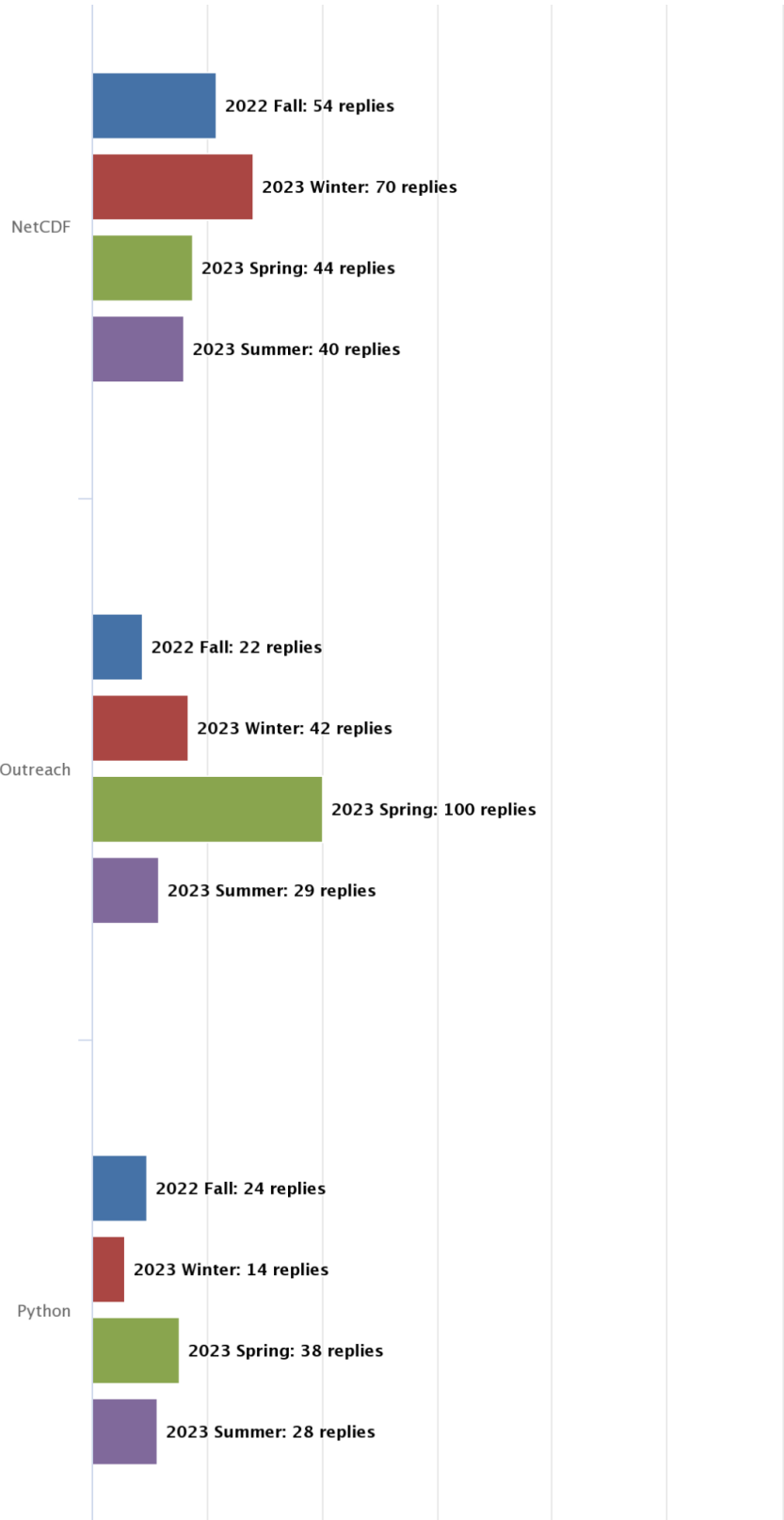
The quarters shown are defined as:

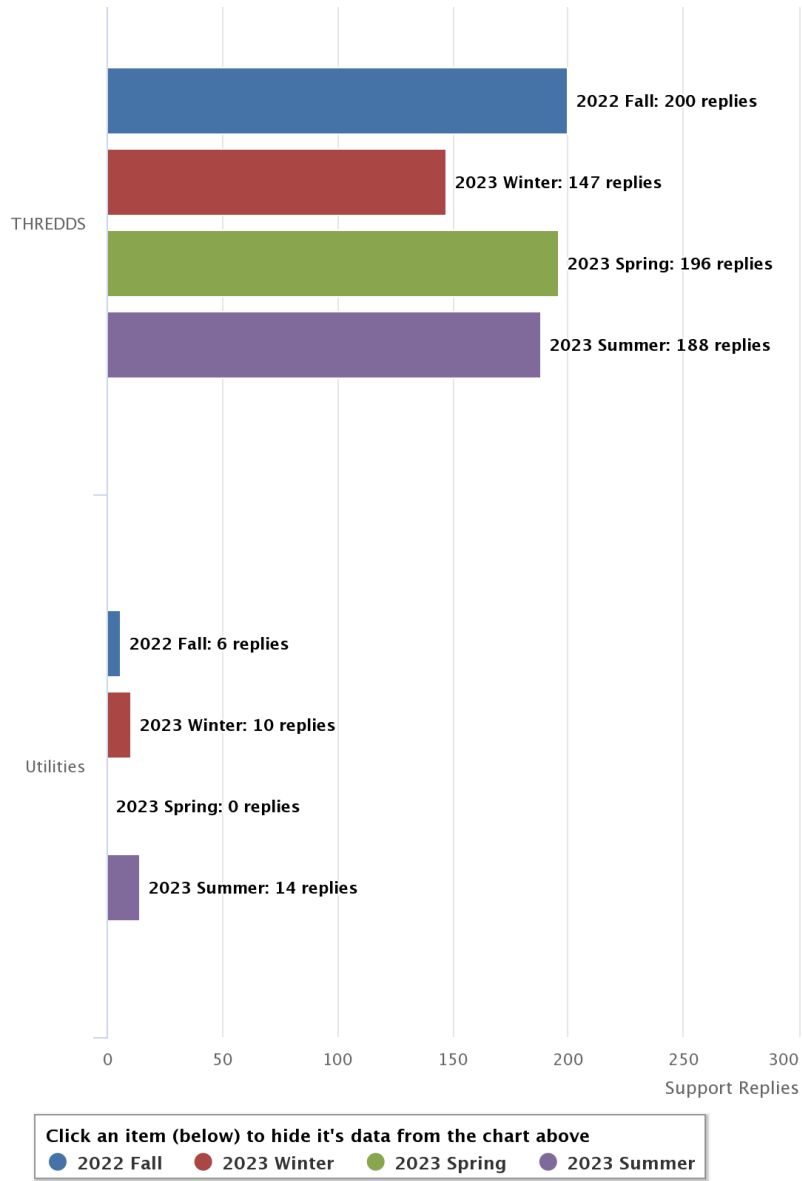
Winter: January, February, March **Spring:** April, May, June **Summer:** July, August, September **Fall:** October, November, December











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

Category	eSupport Departments
AWIPS	Support AWIPS
Data	Support CaseStudy, Support CONDUIT, Support Datastream, Support LEAD, Support Level II, Support NOAAPORT, Support SUOMINET
GEMPAK	Support GEMPAK
IDD	Support IDD, Support IDD Antarctica, Support IDD Brasil, Support IDD Cluster, Support IDD SCOOP, Support IDD TIGGE
IDV	Support IDV, Support IDV Storm, Support McV, Support VisAD
LDM	Support LDM
McIDAS	Support McDevelop, Support McIDAS, ldm-decoders
Miscellaneous	Administration, Development, Plaza, Staging Folder, Support, Support eSupport, Support Miscellaneous, Support Platforms, Support Plaza, Student Interns, Systems, Support Machine Learning
NetCDF	Support LibCF, Support netCDF
Outreach	Outreach, Polcomm, Science Gateway, Support Egrants, Support eLearning, Support News, Support Outreach, Support Workshop, Usercomm, Student Interns
Python	Support Python
RAMADDA	Support RAMADDA
THREDDS	Support netCDF Java, Support THREDDS
Utilities	Support LDM-McIDAS, Support netCDF Decoders, Support netCDF Perl, Support OPeNDAP, Support Rosetta, Support UDUNITS

Comments

- The total support provided by the UPC continues to be substantial: yearly totals have shown a slight decline over the past several years, but this is most likely attributable to the increased ways support is being provided. Overall support activities vary by somewhat by quarter. Spikes in support for individual packages is largely correlated with the releases of new distributions of the packages.
- Support for netCDF continues to be substantial, and is understandable given the very large number of users of the package worldwide.
- GEMPAK is now fully community supported, with occasional input from Unidata staff.
- McIDAS development at Unidata has been curtailed, and support for the current software will continue for the immediate future. Unidata will continue to license the software on behalf of the community (via a MUG membership) for the foreseeable future as well.
- Support for AWIPS has steadily increased and has exceeded that for GEMPAK over the past couple of years.
- Support for Python scripting using MetPy and Siphon are growing markedly.
- Support for LDM, IDD, and Data continue at a high level and show some variability throughout the year.

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, etc., 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 Metrics](#)

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**

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

2. **Providing Useful Tools**

A significant part of providing useful tools is providing support for those tools. Unidata has always provided world class support for all of the tools that it makes freely available to the greater geoscience community.

3. **Supporting People**

The user support provided by the UUPC is recognized throughout the atmospheric science community. Unidata's outreach efforts are routinely noted as being exceptional in surveys of the NCAR/UCAR community.

Prepared *October 2023*

Status Report: THREDDS

April 2023- November 2023

Hailey Johnson, Tara Drwenski, Megan Lerman, Jennifer Oxelson, Ryan May, Ethan Davis, Dennis Heimbigner

Executive Summary

TDS version 5 is now the only supported version of the TDS, and inflow of bug reports is slowing. We have started plans for netCDF-Java version 6 and begun development on TDS microservices, to include: gCDM, File Service, and Catalog Service.

Questions for Immediate Committee Feedback

None at this time.

Activities Since the Last Status Report

General

- Tara Drwenski mentored a TDS summer intern, Jessica Souza, on a very successful project to implement new AI/ML focused enhancements in netCDF-Java and the TDS.
- We have begun implementing a weekly user support rotation amongst the THREDDS team, allowing each team member to focus fully on other projects for two weeks at a time.

NetCDF-Java

- Wrapping up the release of version 5.6.0.
- Work has begun to plan for the development of netCDF-Java 6.x API, which will remove the large number of deprecated methods and limit the public-facing API.

TDS

- Version 5.5 will be released soon (really) and will address a number of user concerns and support tickets; we additionally anticipate a performance improvement in the NCSS and WMS services for enhanced datasets.
- The main development focus for the TDS currently is performance and preparing for microservice development.
- A performance testing pipeline has been set up and is being tested.
- We are revamping the TDS release process to allow for more frequent stable releases.

Ongoing Activities

Server management

- We have continue to maintain our three primary Unidata THREDDS servers:
 - thredds.ucar.edu always runs the latest stable release of the TDS (unless a quick security update is required)
 - thredds-test.ucar.edu automatically deploys new versions of the TDS when netCDF-Java or the TDS GitHub repositories update; it is therefore always running the latest development version
 - thredds-dev.ucar.edu is intended to be used by THREDDS developers, rather than THREDDS users; we use this domain to test changes that require access to “real” data
- Additionally, the Cloud Activities group manages cloud-hosted TDS instance (see cloud activities report for details).

Maintenance

- Maintain thredds.ucar.edu and continue to add new datasets as they become available or are requested by users.
- Closely monitor the security status of our project dependencies, and provide updated versions of our libraries and server technologies to address as needed.
- Continue to respond to user feedback regarding TDS 5.x and transitioning servers to the latest version.
- General maintenance and bug fixes:
 - Fixed a race condition in reading hdf5 and netcdf4 files
 - Various bug fixes and improvements to the Zarr IOSP
 - Removed unused external binaries and updated our 3rd party licenses
 - Fixed custom Thymeleaf fragments
 - Removed unused upload/download pages
 - Upgraded from javax to jakarta namespace
 - Several improvements to NCSS
 - Added several cache configuration properties as requested by users

Development

- DAP4 support
 - Continued work to support the DAP4 protocol in netCDF-Java and the TDS
- Performance and benchmarking
 - Create automated benchmarking and regression testing tools for both netCDF-Java and the TDS.
- Implemented DatasetScans for S3 datasets
- TDS microservices
 - Continue development of microservices
 - Completed backporting gCDM from v6.x to the current version of ncj
- Added UGrid support to the TDS
- Completed a new API for filters and enhancements

- Created new ncml preprocessing classes: Standardizer/Nomalizer
 - Improved performance in the enhancement pipeline
- Added a new AI/ML targeted notebook to dataset on the TDS and implemented regex matching for notebook viewers

New Activities

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

- netCDF-Java
 - Release version 5.6.0
 - Continue to participate in the Zarr development community
 - Work on supporting new features that require later version of Java prompted by [this GitHub issue](https://github.com/Unidata/netcdf-java/issues/753).
- TDS
 - Release version 5.5 of the TDS.
 - Execute new performance and benchmarking tools for the TDS.
 - Create a public repository and community forum for the microservices project.
 - Separate ThreddsIso from the TDS and automate the release process.

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

- netCDF-Java
 - Continue to develop gCDM services
 - Continue work on a version 6 API
 - Develop an extensible NcML framework allowing for pluggable `Enhancements`
- TDS
 - Complete a Python-based File Service
 - Complete a Java-based Catalog Service

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

- netCDF-Java
 - Release a version 6 of netcdf-Java that fully supports Java 11 and the Java Platform Module System (end of Java 8 support)
 - Fully support the Zarr and NCZarr data models, including new iterations of the specifications.
 - Expand use of gRPC to create multiple RPC endpoints in netCDF-Java
 - Support newer version of Java for new features in the netCDF-Java library
- TDS
 - Continue development of standalone TDS services
 - Continue to improve data access performance, exploring the possibility of asynchronous requests.

The following active proposals directly involve THREDDS work:

- The THREDDS team is not participating in any active proposals at this time.

Relevant Metrics

THREDDS Startup Metrics

	2023-03 — 2023-10	2014-08 — 2023-10
TDS Startup (unique IP address count)	2211	42668

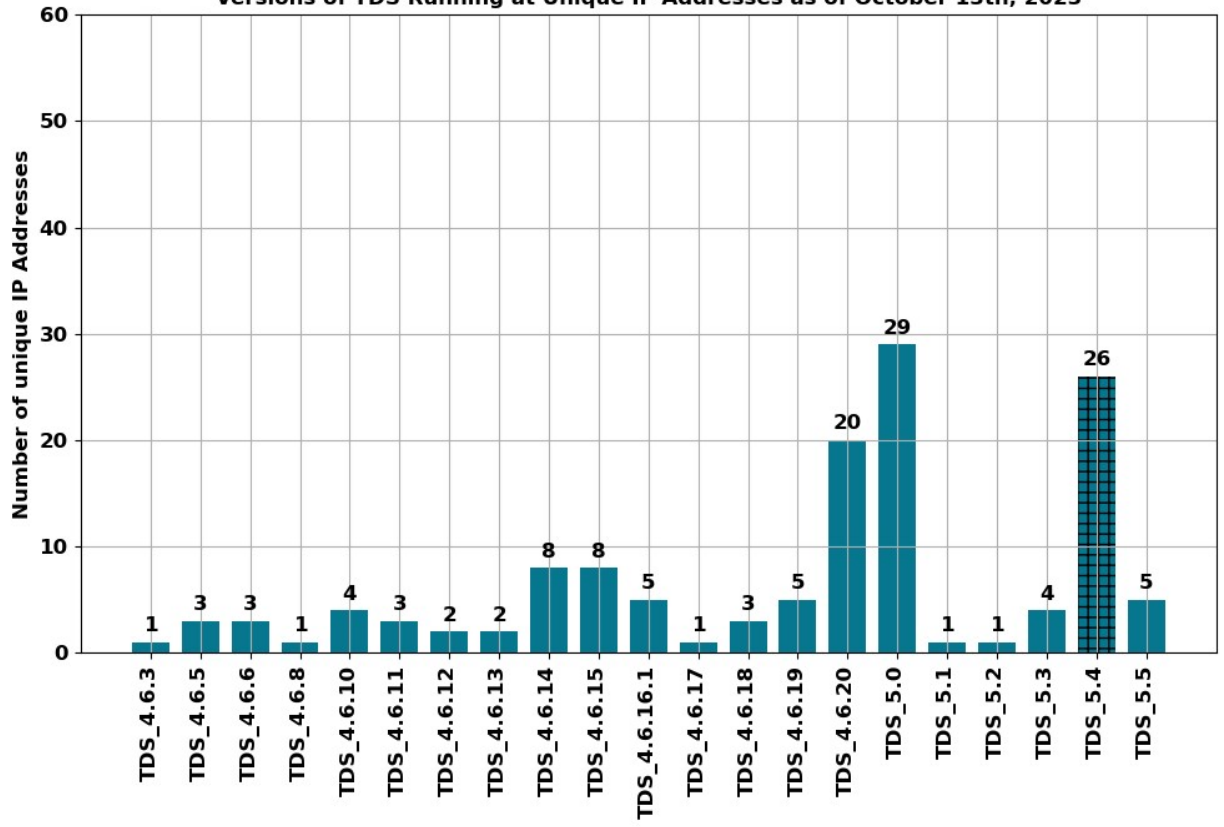
	Total Servers	Information page updated
Publicly Accessible ¹ TDS count	135	76

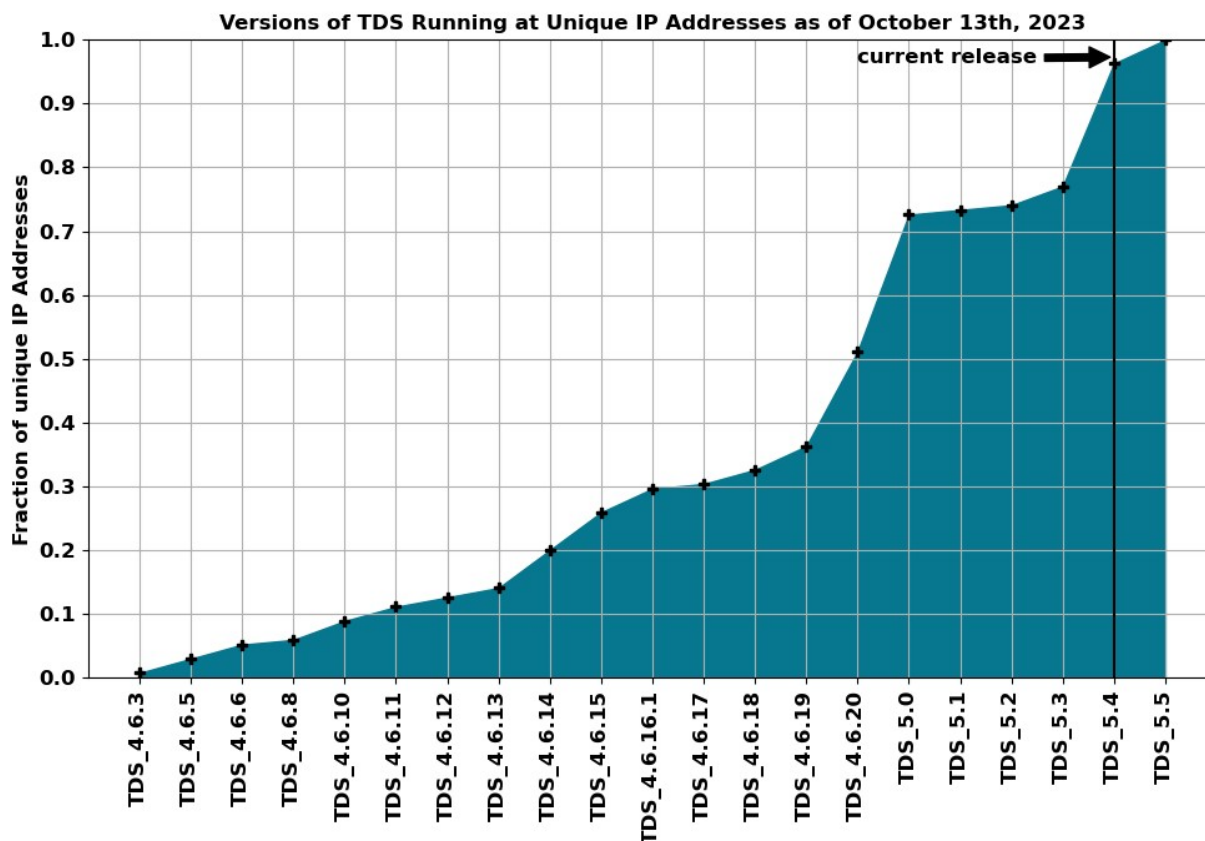
Over the past 4 months, ****2,211**** unique IPs started up the TDS (March 2023 through October 2023). Since we've started tracking these metrics (v4.5.3, August 26th, 2014), we've seen the TDS startup from ****41,107**** unique IP addresses. There are currently ****135**** publically accessible TDSs running "in the wild". Of the ****135**** publically accessible servers, ****76**** have updated the name of their server in their server configuration file (taken as a sign that they are maybe, possibly, intended to be used by others...maybe...).

The figures below show the distribution of TDS versions (top), and the fractional share of servers running version X or older (bottom). Each labeled version includes betas and snapshots, not just the official release of that version, for presentation simplicity. TDS v5.0 remains the dominant specific version running in the wild, although there are almost as many instances of 5.4.

¹ "Publicly accessible" means we could find a top-level THREDDS Client Catalog. We checked <server>/thredds/catalog.xml (version 4), <server>/thredds/catalog/catalog.xml (version 5), including the most common ports of 80, 8080, 443, and 8443.

Versions of TDS Running at Unique IP Addresses as of October 13th, 2023





Strategic Focus Areas

The THREDDS projects covered in this report support the following goals described in Unidata Strategic Plan:

1. Managing Geoscience Data

The component software projects of the THREDDS project work to facilitate the management of geoscience data from four points of view: __Making Geoscience Data Accessible, Making Geoscience Data Discoverable, Making Geoscience Data Usable, and Enhancing Community Access to Data__. As a client-side library, **netCDF-Java** enables end users to read a variety of data formats both locally and across numerous remote technologies. Less user-friendly formats, such as GRIB, are augmented with metadata from community driven metadata standards (e.g. Climate and Forecast metadata standards), and viewed through the more user friendly Common Data Model (very similar to the netCDF Data Model), providing a single set of Java APIs for interacting with a multitude of formats and standards. The **THREDDS Data Server** exposes the power of the netCDF-java library outside of the Java ecosystem with the addition of remote data services, such as __OPeNDAP__, __cdmremote__, __OGC WCS__ and __WMS__, __HTTP direct download__, and other remote data access and subsetting protocols. The TDS also exposes metadata in standard ways (e.g. ISO 19115 metadata records, json-ld metadata following schema.org), which are used to drive search technologies. **Rosetta** facilitates the process of translating ascii based

observational data into standards compliant, archive ready files. These files are easily read into netCDF-Java and can be served to a broader community using the TDS.

2. **Providing Useful Tools**

Through Rosetta, the THREDDS project seeks to intercede in the in-situ based observational data management lifecycle as soon as possible. This is done by enabling those who produce the data to create archive ready datasets as soon as data are collected from a sensor or platform without the need to write code or intimately understand metadata standards. NetCDF-java and the TDS continue to support legacy workflows by maintaining support for legacy data formats and decades old data access services, while promoting 21st century scientific workflows through the creation of new capabilities and modernization of existing services (e.g. Immutability, upgraded technical stack, microservice development).

3. **Supporting People**

Outside of writing code, the THREDDS project seeks to support the community by __providing technical support, working to build capacity through Open Source Software development, and by building community cyber-literacy__. The team provides expert assistance on software, data, and technical issues through numerous avenues, including participation in community mailing lists, providing developer guidance on our GitHub repositories, and leading and participating in workshops across the community. The team also actively participates in “upstream” open source projects in an effort to help sustain the efforts of which we rely and build upon. We have mentored students as part of the Unidata Summer Internship Program, and worked across organizations and disciplines in support of their internship efforts.