Agenda: Spring 2023 Users Committee Meeting

(Times are Mountain Daylight Time)

Thursday, March 30th

09:00 - 09:15 Open Meeting

09:15 – 09:30 Updates from Members (Committee)

09:30 - 10:30 Director's Report (Mohan Ramamurthy)

10:30 - 10:45 Break

10:45 - 12:00 Staff Status Reports (All)

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

12:00 - 13:15 Lunch at the Foothills Lab Cafeteria

13:15 - 13:45 Status Report Version Discussion

13:45 - 14:00 Update on the June 2023 Triennial Users Workshop (Planning Chairs)

14:00 - 14:30 NCEP Report and Questions (Margaret Curtis)

14:30 - 15:00 Unidata's Educational Services Going Forward (Nicole Corbin)

15:00 - 15:15 Break

15:15 – 15:45 Unidata Science Gateway Reimagined (Ana Espinoza)

15:45-16:00 Wrap-up Day One

16:00 Adjourn

18:30 Collaborative Discussion on the Day's Proceedings Over Dinner at Boulder Social, 1600 38th Street, Boulder, CO 80301 (map)

Friday, March 31st

09:00 - 09:15 Convene and Outstanding Items from Previous Day

09:15 – 09:45 Unidata Strategic Plan (Tanya Vance)

09:45 - 10:30 DeSouza Award. Discuss Candidates for the 2023 Honor

10:30 - 10:45 Break

10:45 - 11:45 Community and Committee Needs and Challenges Discussion

- What are the needs for research and teaching?
  Use cases, tools currently used/misused, current barriers
- What training is most needed?
  Training that requires teaching by Unidata and training that can be provided by the community with support from Unidata

11:45 - 12:15 Wrap-up Day Two, Action Items, Fall Meeting Dates, etc.

12:15 Close Meeting

13:00 Convene Equipment Award Panel

Status Report: Users Committee Actions

November 2022- March 2023

Unidata Program Center Staff

Actions from the Previous Meeting (June 2-3 2022)

Action 1

During discussion about how the netCDF-C library Zarr implementation can support various backend storage systems/services, the idea of splitting the backend storage support out as a separate library that could be community maintained separate from netCDF-C came up. GDAL was mentioned as a possible partner or maybe already having such a library. The netCDF team to investigate collaboration with outside groups on this front? [netCDF team]

Result

The netCDF team is in collaboration with the ZARR community group, we are members of the Zarr Enhancement Protocol (ZEP) committee, and we are trying to get our foot in the door with the (potentially) rebooted HDF Technical Advisory Board. We will continue investigating the possibility of a back-end storage library.

Prepared *March 2023*

Status Report: AI/ML

November 2022- March 2023

Unidata Program Center Staff

Executive Summary

Unidata ML Staff has been working on educational materials, project pythia cookbooks, and blog posts. 2023 will mark more in person events (Triannual meeting, in person workshops, SciPy) along with virtual office hours. ML staff also submitted a NSF proposal with other Unidata staff to fund specific in-classroom training and materials creation around machine learning skills for earth systems scientists.

Questions for Immediate Committee Feedback

- How to increase visibility and interactions with Unidatas core community?

Activities Since the Last Status Report

- Submitted NSF grant for more formal workforce development training in the classroom
- Added online office hours (sign up here: https://calendar.app.google/ZsM8dLHLa65eGAr39)
- Attended AGU and AMS to see state of the science
- Python resources have been shared on Project Pythia and other openly available github repositories
- Continued testing GPU (graphics processing units) resources on Jetstream2 for wider adoption for Jetstream users. Hope to launch service to wider community in the coming months.

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. Expect these activities to increase throughout 2023. Development of a 1 day ML workshop is a goal for 2023 to be delivered at future workshops and events.

Prepared March *2022*

Status Report: AWIPS

November 2022- March 2023

Tiffany Meyer, Shay Carter

Executive Summary

Over the last year, the AWIPS team has been hard at work to put out its first major version release. The current AWIPS team has been maintaining and adding updates to the last major version released, which was a variant of the NWS' AWIPS v18.1.1. We now have a beta version of AWIPS based off of NWS' v20.3.2, available with key updates such as Python3 and Java11.

At the end of February 2023, we released our first beta of v20.3.2 (called 20.3.2-0.1) with two installation options for CAVE:

- A direct CentOS7 install via RPMs
- A CentOS7 Virtual Machine that runs with VMWare Workstation Player (available on both CentOS and Windows)

We also have a publically accessible beta EDEX available at edex-beta.unidata.ucar.edu. Along with the release, we updated our documentation to have a dedicated page for installing the beta, and a new Google form specifically for reporting bugs/unexpected behavior with v20.

As of early March, we have resolved significant challenges we were having with regards to building CAVE directly for MacOS and Windows. We are hoping to put out a second release of the beta (v20.3.2-0.2) by the end of March that would include:

- Two bug fixes for issues reported with v20.3.2-0.1
- A direct, signed CAVE installation for Windows (.msi)
- A direct, signed, and notarized CAVE installation for Mac (.dmg)

Aside from AWIPS development, the team is proud to have the first two Unidata online training modules in *Learn AWIPS CAVE*, and *Learn Python-AWIPS*. We also just completed two consecutive years of releasing AWIPS Tips blogs every other week. The AWIPS team is also looking to host a workshop at AMS next year on CAVE.

Questions for Immediate Committee Feedback

As stated, the AWIPS team is looking to host CAVE training at AMS next year, what format would be more beneficial to the community - short course or student workshop?

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 various hiccups with the new platform. The AWIPS team was also able to experiment with a newly-offered "large" virtual machine instance during testing for the v20 beta. We currently have 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.

The main challenge we investigated with the "large" VM instance, was whether one machine could handle all of the data we process and make accessible through our EDEX. Unfortunately, with the given constraints of queuing architecture in EDEX, even the "large" instance was not capable of sustaining our entire real-time data flow. However, through the use of ancillary EDEX machines we have been able to decouple certain datasets from the main EDEX instance. We still 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.

Texas A&M has also been using a similar distributed architecture since Summer of 2021. We have worked closely with them since they first installed their system, as we've released new versions of AWIPS, to help them transition between versions and identify weaknesses in their EDEX server.

Since the last status report, we've put out a beta release for version 20 of AWIPS (20.3.2-0.1). A link to all of our AWIPS release notes can be found here.

This release is the first major version upgrade since the current AWIPS team has started. Version 20 has some major upgrades in terms of dependent software as well, with the following significant upgrades:

- 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 this beta release the AWIPS team has put out a direct install option for CentOS7 via new RPMs, and a Virtual Machine option that runs on both Windows or CentOS. We have a public v20 EDEX available at edex-beta.unidata.ucar.edu. Our beta release of CAVE includes all the latest NWS updates and features, as well as a few of our Unidata updates. Any bugs or unexpected functionality can be reported through this brief form.

The AWIPS team is currently working on releasing direct install options for both Windows and MacOS. Our next release will also include a few changes in response to issues that were identified in our beta reporting form, linked above. The team intends to release updates to the beta as we are able to incorporate previous Unidata-added functionality into CAVE and EDEX. We are hoping to have a production version of 20 by the end of the year.

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.

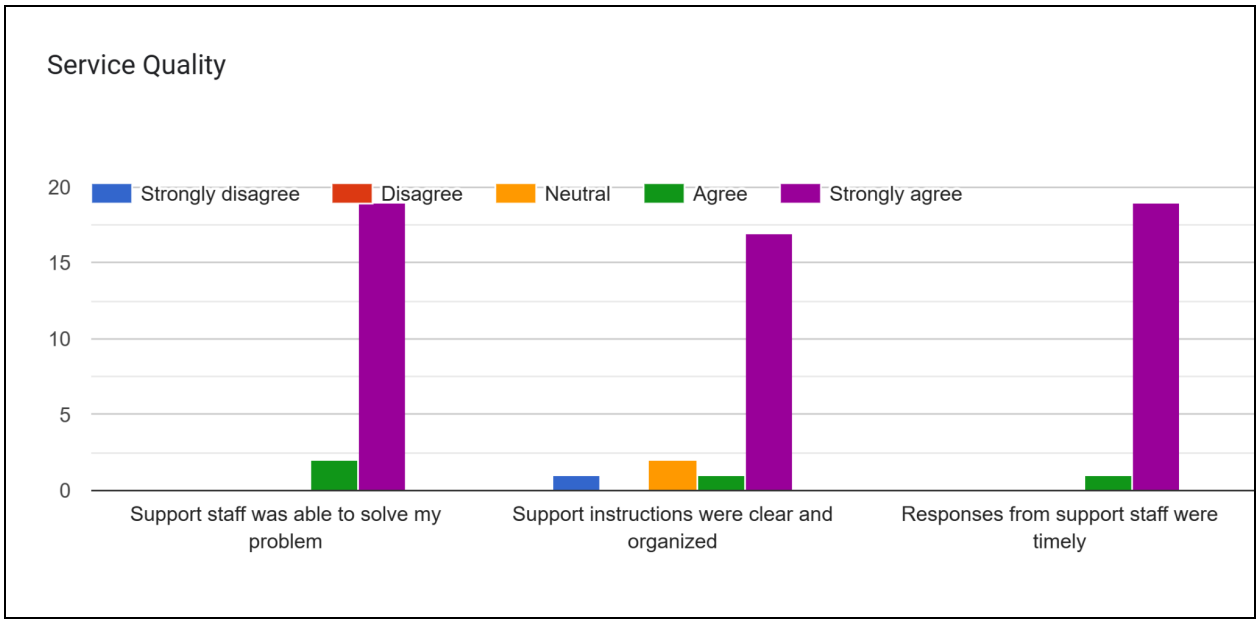
A significant portion of our documentation both for CAVE, EDEX, and python-awips has been modified for easier understanding and comprehension. We are continuing to update our python-awips example notebooks to follow our new template that contains a helpful table of contents, with consistent subsections across the various example topics.

Our blog series, AWIPS Tips, has successfully been running every other week for almost two years now. As of March 22nd, 2023, we will have released 52 blog entries (two consecutive years worth)! 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).

Our asynchronous CAVE training course has been live since October 2021. We have had over 100 users sign up and take our course. 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.

Along the same lines, our second training module, *Learn Python-AWIPS*, has been live since September 2022. *Learn Python-AWIPS* is tailored for users with entry level Python knowledge, and zero python-awips knowledge. The course walks learners through the process of using python-awips to connect to an EDEX, interrogating the server to find what data is available, constructing and tailoring a refined data query, and displaying the results of the data in a user-friendly plot. After completing the course, the learner will have a fully functioning Jupyter Notebook example from their work following along in the course. 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 have had a handful of users take this course since we released it and 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:



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

"Your AWIPS Team is awesome. I'm getting responses back usually in less than two to three hours. More responsive than the NWS AWIPS Program Office!!"
"Tiffany was able to correctly answer and have the problem fixed in a timely fashion!"
"They are incredibly knowledgeable and approachable. They always go the extra mile to help. They are appreciated very much."

At AMS 2023, the AWIPS team presented in a new Unidata session titled "Community Driven: Unidata Projects Enhancing Geoscience Teaching and Research". The talk, "Unidata AWIPS Hosted in the Cloud" covered an overview of AWIPS capabilities, the differences between our version and the NWS version, use of the Jetstream2 cloud platform, and our current development milestones and outlooks. The team also had the chance to demonstrate CAVE and AWIPS, and interact with users at the Unidata booth during the AMS Student Career Fair.

The AWIPS team is planning on submitting to host a workshop at AMS 2024. At this time, we are still deciding if we'd like to host a "short course", or a student workshop. The idea is the workshop would focus primarily on CAVE, but could potentially touch on python-awips as well.

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 a beta for version 20 (20.3.2-0.1). This update was

made available for direct install on CentOS via new RPMs, and in a Virtual Machine which can be run on Linux or Windows. We also have a publicly available version 20.3.2-0.1 beta EDEX server, hosted in our cloud allocation.

We are still actively developing the Unidata AWIPS beta release version 20 of AWIPS. We are anticipating upcoming releases of the beta to support directly installing CAVE on both MacOS and Windows. Additionally, we plan to make the EDEX installation available as we transition towards a production version of 20. Our plan is to continue releasing a few versions of the beta until we are comfortable supporting a production release, hopefully by the end of the year.

Activities Ongoing/In-Progress

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

- The AWIPS team is looking into more effective bundling and distribution of CAVE dependencies for both Windows and MacOS.
- The AWIPS team has been testing our new Jetstream2 platform with our production and development EDEX servers.
- 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 a new v20 AWIPS, with the hopes to have a production release by the end of the year.

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 release (version 20). This includes updating our python-awips package to work with our new beta version 20 of AWIPS. We are also looking forward to the opportunity of hosting an in-person workshop at AMS 2024. The team is actively participating in conferences, workshops, and virtual message boards (blogs) to expand our user base.

Metrics

Downloads November 1, 2022 - February 28, 2023

AWIPS downloads: 1,666

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 *March 2023*

Status Report: Science Gateway and Cloud Computing Activities

November 2022- March 2023

Shay Carter, Julien Chastang, Nicole Corbin, Ana Espinoza, Ward Fisher, Ryan May, Tiffany Meyer, Jen Oxelson, Mohan Ramamurthy, Jeff Weber, Tom Yoksas

Executive Summary

In the science gateway arena, we've experienced a strong increase in use demonstrated by the three workshops we helped at AMS, coupled with the nine universities we are supporting with specialized JupyterHub resources this semester. Moreover, we've had advancements on the technical side that have facilitated teaching. In particular, JupyterHub users now have the ability to access shared data directories across all user environments, a vital capability for instructors who wish to disseminate large case study data to students. In addition, we have several examples of WRF NWP now running on science gateway components discussed below. This acceleration in use and capabilities of the science gateway would not have been possible without the addition of science gateway staff over the last year.

Questions for Immediate Committee Feedback

1. As we try to bring components of the science gateway together into a meaningful presentation on the web (see "Science Gateway Reimagined" section), what components do committee members think are missing (e.g., ways of sharing material, discussion forums, web forms to ask for resources, etc.).
2. Are there any committee members that would like early access to an experimental GPU enabled JupyterHub server?

Activities Since the Last Status Report

American Meteorological Society 2023 Meeting

Workshops

Unidata staff supported three workshops with customized PyAOS JupyterHub servers: MetPy Short Course, Student Python Workshop, and the Colorado State University LROSE radar meteorology workshop. In all, Unidata had 141 participants using Jetstream2 resources at AMS 2023.

Presentation

Unidata delivered a presentation entitled "Unidata Science Gateway: Past, Present, and Future Plans" detailing our collaboration and use of the Jetstream and Jetstream2 projects since 2015.

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.

GPU Exploration on Jetstream2

Unidata staff have started experimenting with Jetstream2 GPU hardware. In particular, we have deployed a GPU-enabled JupyterHub equipped with TensorFlow and PyTorch software, and are evaluating its capabilities specifically with regard to educational settings. We are also currently working on testing out Andrea Zonca's GPU version of deploying a "JupyterHub with Kubernetes" cluster.

Unidata Science Gateway Re-Imagined

The Unidata Science Gateway unites Unidata technologies with Jetstream2 and open-source software such as Project Jupyter to provide the academic and research community an environment they can use to access, analyze, and visualize real-time and case study Earth Systems Science data. While the Unidata Science Gateway has been moderately successful in reaching our audience, we would like to improve and expand our web presence, building a portal that allows users to more easily access educational, computing, and data resources. We aim to revamp our current gateway interface to become a more dynamic hub for education, data access, and research. We presented our vision and mission statement in a poster presentation at the Gateways 2022 conference. Unidata staff are currently in the design phase of this project.

JupyterHub Servers for Classroom Instruction Spring Semester 2023

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. The Jetstream2 allocation is also an integral part of the Unidata Science Gateway project. Since the launch of Jetstream2 in 2022, we have deployed PyAOS (Python for the Atmospheric and Oceanic Sciences) JupyterHub installations at ten universities. Unidata tailors these servers to the requirements of the instructors so they can accomplish their Earth Systems Science teaching objectives. For the spring semester of 2023, 122 students at eight academic institutions have used Unidata JupyterHub servers running on Jetstream2.

University of Hawaii LROSE Workshop

Unidata staff provided JupyterHub resources enabled with LROSE radar meteorology software for a workshop led by Professor Mike Bell at Colorado State University. 14 workshop attendees employed this LROSE JupyterHub server. We also provided advice to the LROSE group to obtain their own Jetstream grant so that they may launch these types of servers on their own allocation.

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. Notably, we worked with Andrea to provide our users with a shared directory that can be accessed from all user JupyterLab environments. This has been a long-standing feature request that we were excited to finally deliver. For the spring semester, we have deployed this capability for JupyterHub servers at Millersville University, Virginia Tech, and OU. This allows instructors to much more easily render accessible large data sets in a data proximate manner to students' notebook environments.

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.

Unidata Docker Container Improvements

tomcat-docker

This Docker container is the parent container to both ramadda-docker and thredds-docker. It is therefore important it remains current and up-to-date especially with respect to security improvements that may be happening in the official parent tomcat Docker container.
To that end, we have improved GitHub actions to ensure this container remains automatically up-to-date with respect to upstream changes. The official upstream tomcat container supports various combinations of tomcat version, JDK or JRE version and implementation, and even the base OS image version. Therefore, we have enabled users to easily request for a

specific combination of the above to be similarly automatically updated and tracked by the same GitHub actions workflow, a feature that has already seen some use.

ramadda-docker

No recent changes.

thredds-docker

Documentation improvements with respect to available image tags for this container will hopefully alleviate confusion.

ldm-docker

Update to incorporate LDM version 6.14.5.

Jetstream2 Allocation Usage and Time Extension

Unidata's Jetstream2 allocation request has been active and receiving use since April 1, 2022. Since then, Unidata's SU usage has occurred at a rate which agrees with our projected usage for the original proposal's year-long timeline (see the Relevant Metrics section below). However until September 7, 2022, Jetstream2 was in the "early operations" phase and was not actively "billing" projects for the service units (SUs) which virtual machines typically consume. These five "free" months of Jetstream2 usage has resulted in an excess of SUs that are unlikely to be used by the allocation's original expiry date of March 31, 2023. As such, Unidata staff have submitted, and been granted, a six month time extension for the project. We supplied a "Status Report" to the ACCESS (formerly XSEDE) program as part of that request. The new end date for the allocation is September 30, 2023.

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.

JupyterHub Demonstration Server

Unidata continues to enhance the Unidata JupyterHub demonstration server which receives occasional use though we tend to try to steer our users to dedicated semester or project long JupyterHubs. Due to the limited uptime of these project-based servers, they are easier to maintain.

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. Last summer the AWIPS team worked closely with other Unidata staff members (namely Julien Chastang, Ana Espinoza, and Mike Schmidt) to successfully transition all our EDEX machines from Jetstream1 to Jetstream2.

EDEX is designed with a distributed architecture, 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 3 sets of servers as described above: two sets are running our v18 software (production version of AWIPS), and one set is running our new beta v20 software. The live backup allows us to be able to patch and maintain our servers and have a fail-safe when something goes wrong with the current production system. The new v20 version allows us to support our beta while simultaneously supporting our v18 software and not disrupting classrooms/labs that may not have the opportunity to update in the middle of a semester.

In our new allocation for Jetstream2, we secured access to an even more powerful machine (a “large instance” virtual machine) that we used as a test platform for developing our v20 EDEX server. Unfortunately, because of the current architecture of EDEX, even this “large instance” was not able to sustain the ingest and decode workflow of the entire dataset volume we serve. We may do more testing in the future when we have more resources to dedicate to the task.

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.

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

- Chastang, J., Corbin, N., Davis, E., Espinoza, B., & Vance, T. (2022). Unidata Science Gateway Reimagined: Unifying Access to Educational and Research Resources. https://doi.org/10. 5281/zenodo.7201472
- Chastang, J., Espinoza, A., & Ramamurthy, M. (2023). Unidata Science Gateway: Past, Present, and Future Plans. https://doi.org/10.6084/m9.figshare.22043642
- Espinoza, B., Chastang, J., Weber, J., Dye, D., & Romine, P. (2022). Democratizing Access to Atmospheric Modeling with WRF employing NSF Cloud Computing Resources.
- Meyer, T. C., Carter, S., & Corbin, N. (2023). Unidata AWIPS Hosted in the Cloud . Retrieved from https://ams.confex.com/ams/103ANNUAL/meetingapp.cgi/Paper/419377
- Weber, J., Dye, D., & Romine, P. (2023). A network system for environmental monitoring, data sovereignty and data governance on tribal lands. Retrieved from

https://ams.confex.com/ ams/103ANNUAL/meetingapp.cgi/Paper/419480

New Activities

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

Forthcoming Conference Attendance

- 2023 European Geophysical Union Spring Meeting
- 2023 ESIP Summer Meeting

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

JupyterHub Collaboration Andrea Zonca

We plan to continue our collaboration with Andrea Zonca (San Diego Supercomputing Center) for deploying JupyterHub clusters on Jetstream2 and exploring new technologies in this area such as Dask and GPU programming. We continue to provide Andrea with feedback as he releases new versions of the software. We are happy to report that Andrea Zonca is now being funded in part by the Jetstream2 project. This funding will allow this vital collaboration to continue.

Relevant Metrics

Summer/Fall 2022 JupyterHub Servers

Since spring of 2020, Unidata has provided access to JupyterHub scientific computing resources to 1212 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.

	Nº of Users	Point of conact
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 Hawaii WS	15	Prof Mike Bell (CSU)
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	10	Prof Jason Naylor
University of Wisconsin	0	Pete Pokrandt
Virginia Tech University	12	Prof Craig Ramseyer

Jetstream2 Allocation Usage Overview

In addition to service units (SUs) used for running various kinds of virtual machines, “regular” CPU instances, GPU instances, and “large memory” 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. The metrics presented below are as of March 15, 2023.

SU Metrics

Type	Usage (SUs)	Total (SUs)	Percent Usage*
“Regular” CPU	3335255.732	5098320.0	65 %
GPU	75887.77	600000.0	13 %
Large Memory	359926.383	400000.0	90 %

Resource Metrics

Compute			
Type	**Used**	**Total**	**Percent Usage***
Instances	93	150	62 %
vCPUs	1046	4035	26 %
RAM	4.1 TB	15.8 TB	26 %

Storage			
Type	**Used**	**Total**	**Percent Usage***
Volumes	266	400	67 %
Volume Snapshots	0	50	0 %
Volume Storage	30.4 TB	39.1 TB	78 %

Network			
Type	**Used**	**Total**	**Percent Usage***
Floating IPs	41	310	13 %
Security Group	72	100	72 %
Security Group Rules	213	300	71 %
Networks	17	100	17 %
Ports	151	250	60 %
Routers	3	15	20 %

* 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	4	16 (+1)	11 (-1)	5 (-7)	166 (+10)	6 (+6)	596 (+46)
tomcat-docker	10 (+1)	59 (+5)	65 (+1)	2	40(+2)	0	72(+1)
thredds-docker	15(+2)	25	25(+1)	5	110(+1)	0	159(+1)
ramadda-docker	2	0	2	1	10	0	24
ldm-docker	8(+2)	15(+2)	13	5(+2)	36	0	61(+2)
tdm-docker	4(+1)	4(+1)	7	1	9	0	18

* 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.

Prepared *March 2023*

Status Report: Community Services

November 2022- March 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
- Outreach to the Earth System Science community through participation at AGU, AMS, SACNAS, and AIHEC conferences
- Expansion of learning materials, resources, and workshops
- Strategizing actions to make our educational services more discoverable on the web as well as promoting the variety of educational services we are able to provide
- We have also explored external funding to support an innovative low-code learning experience in AI/ML
- Providing support for the planning and implementation of the 2023 Unidata Users Workshop
- Strategic Planning in preparation for for the 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:

- Unidata Staff at AGU Fall 2022 Meeting
- Ryan Abernathey Receives 2022 DeSouza Award
- Call for Proposals: Unidata 2023 Community Equipment Awards
- Unidata Summer Student Internships Available!
- Unidata Staff at AMS 2023 Meeting
- Offer: Unidata Science Gateway JupyterHub Resources Available for Spring 2023 Courses
- 2023 DeSouza Award Nominations
- AMS 2023 Conference Highlights from the Unidata Staff

- Radio Occultation Data from COSMIC Available in the IDD
- Recommended Books for Machine Learning
- JupyterHub Expands Learning Opportunities at Valparaiso University
- Convergence Science in Action: I-GUIDE Summer School 2023
- Register Now for the 2023 Unidata Users Workshop
- Software release information
- Many AWIPS Tips and MetPy Mondays episodes
- Community job postings
- Community meetings and other announcements

Dependencies, challenges, problems, and risks include:

- Finding community members willing to contribute stories (or story ideas) for the blog is an ongoing challenge. We're starting to make progress working with committee members, but there is more to do.

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 a new Unidata strategic plan for the next 5 years
- The community services team has supported the Users Committee's work in planning the June 2023 Unidata Users Workshop
- Organized a successful Unidata "track" in the EIPT conference at the 2023 AMS Annual Meeting
- Participated in and judged student posters at the 2023 AIHEC Conference and co-led site visit to SIPI with Sovereign Network partners, NEON, and Dean of SIPI's Earth Science Program and Department Chair, Dr. Milford Muskett
- Continue to work with AIHEC, NEON, UCAR/NCAR staff on NSF funded MSI engagement and NSF funded Data Sovereignty Project
- Met towers were installed at SIPI and NTU, instrumentation is underway and progress made to run WRF over 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"
- 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) and supporting partners at 2023 SACNAS Conference
- 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:

- Submitted NSF proposal "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)
    - 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.
- Prototyping the Science Gateway, Reimagined website, including a comprehensive learning catalog for all Unidata's educational offerings. See Cloud Computing Activities for more details.
- Rebranded and promoted the Python Readiness for Earth Systems Sciences learning series, originally piloted at Colorado State University: https://elearning.unidata.ucar.edu/examples/PythonReadiness
- Development of a new netCDF microlearning module on Unidata eLearning is in progress. Expected release in late Spring 2023.
- Delivered learning events at AMS 2023:
    - MetPy for your Data: Analyzing Meteorological Observations in Python-Member
    - Student Conference: Python Workshop
- In late October, 2022, Steve Emmerson, Tom Yoksas, and Stonie Cooper made a presentation on the LDM and its use in the Google offices in Boulder

- From 2023-02-22 through 2023-02-24, Steve Emmerson, Tom Yoksas, Stonie Cooper, and Mike Zuranski conducted an LDM workshop for NOAA/NWS personnel in Alaska, Hawaii, and Guam. Approximately 40 people attended.

Dependencies, challenges, problems, and risks include:

- Assignment and allocation of resources to build the new Science Gateway website

Learning and Outreach Ongoing Activities

We plan to continue the following activities:

- Support for governing committee activities (convening and coordinating discussions, planning for the upcoming Users Workshop, etc.)
- Ongoing development of news articles and blogs for publication through News@Unidata
- Seeking partnerships to build and deliver community educational resources and continue to expand Unidata's educational services
- Continue to engage with underserved populations and institutions
- Engagement with professional societies
- Support the pursuit of funding
- Engage other UCAR/NCAR divisions regarding Unidata software use
- Ongoing work to transition Unidata's web site to UCAR-mandated system
- Active participation in the Hydroshare Advisory Committee (CUAHSI)
- Continue working with AIHEC, NEON, 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:

- A Train-The-Trainer session for Colorado State University is pending for Summer 2023. This provides graduate students teaching resources and facilitation training, enabling them to deliver their own Python instructor-led training.
- Host three summer interns with interests in scientific software development/engineering and education
- Participate in May 2023 ESIIL Summit
- Support coordination, facilitation, two presentations, coaching, and delivery of June 2023 Unidata Users Workshop
- Co-lead June 2023 AIHEC TCU Data Workshop
- Work with students and faculty from NTU and SIPI for Summer 2023 DOE/SAIL and NOAA/SPLASH internships

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

- AWIPS instructor-led training, tentatively to be piloted as a short course at AMS 2024
- Lead a data management and visualization professional development workshop for UCAR interns
- Facilitate a MetPy Instructor-Led training in August 2023 with Howard University, Morgan State University, and the University of Maryland Baltimore County
- Pending funding, development of a series of AI/ML educational resources
- Hold a workshop in October 2023 for AIHEC and Sovereign Network participants
- Engage other underrepresented communities
- Expand sovereign network to other TCU's and AIHEC members
- Prepare for follow up proposal w NTU, SIPI, TOCC, and others

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

Relevant Metrics

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

All community pages

Most recent six months:

- 40,323 unique pageviews (35,619 in previous period)
- 11.4% of total unique pageviews (9.3% in previous period)

Top community pages

1. All blog pages
   25126 unique pageviews (25784 in previous period)
   60% of total community pageviews (61% in previous period)
2. www.unidata.ucar.edu/community
   8781 unique pageviews (4543 in previous period)
   21% of total community pageviews (11% in previous period)
3. www.unidata.ucar.edu/about
   2965 unique pageviews (3196 in previous period)
   7% of total community pageviews (7% in previous period)
4. www.unidata.ucar.edu/events
   2983 unique pageviews (1601 in previous period)

7% of total community pageviews (4% in previous period)

Social media statistics, March 15, 2023

1. # of Twitter followers: 1978 (up from 1901)
2. # of Facebook followers: 899 (up from 889)
3. # of YouTube subscribers: 3160 (up from 2853)
4. #of LinkedIn followers: 89 (up from 70)

Unidata eLearning statistics, March 7, 2023

1. Total unique users: 130
2. Enrolled users in Learn AWIPS CAVE: 113
3. Enrolled users in Learn Python-AIPS: 12

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 web site, electronic newsletters, and social media to keep community members informed about enhanced data services, software tools, and cyberinfrastructure.

   We participate in UCAR/NCAR and NSF projects for 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

November 2022 - March 2023

Mike Zuranski, Tom Yoksas

Executive Summary

Welcome to the first edition of the Data Services Status Report!

Work continues on the RTSTATS revamp project. Once that reaches a stable point work will begin on IDD cataloging, the hope being to have results to showcase for the Fall committee meetings. Additional data has been added to the IDD, and efforts towards building a better LDM/NP monitoring solution are underway. We are also looking for ways to improve communications with data providers, users and partners.

Questions for Immediate Committee Feedback

Regarding data services, what methods of community communication do you prefer; are the existing email lists adequate? Would consolidating some of the data lists help or hurt?

Would additional venues for communication, Discourse for example, help exchange information or would that be too much?

It appears that the text product driving our nws-changes list has been discontinued for some time. Would you desire a replacement, or can this list be discontinued?

Activities Since the Data Engineer was hired

RTSTATS Revamp

Our Real-time Statistics (RTSTATS) pages and ingest code were created decades ago and are in need of an update. Daryl Herzmann made a solid effort here during his tenure on the Users Committee, his project continues to run at rtstatstest.unidata.ucar.edu.

I have begun work on my own RTSTATS revamp, which currently uses Daryl's ingest code and an interactive front-end interface made with Dash. While only a proof of concept of this is running now, feedback on the interface thus far has been very positive. It has become clear, however, that I will not be able to keep using Daryl's code, and I will need to rewrite the ingest side as well.

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:

- None

Improved Monitoring with the TIG stack

We are investigating the use of the TIG stack (Telegraf, InfluxDB, Grafana) for monitoring both system health metrics as well as LDM, NOAAPort and other data metrics. Telegraf can collect a wide variety of data, to be stored in an InfluxDB database, with Grafana as an online interactive dashboard. Grafana also includes alerting capabilities so staff can be notified of anomalous situations. While system metrics will only be available to UPC staff, the goal is to make certain data metrics public as we believe there's value in our users being able to see the health/status of the IDD through these tools.

Progress has been made on the following:

- Initial testing of the TIG stack on UPC environments
- Software deployment to select internal machines
- Scripting to collect LDM and NOAAPort metrics

Dependencies, challenges, problems, and risks include:

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

Add Data to the IDD

We recently partnered with the UCAR COSMIC Program to provide new SPIRE Radio Occultation data on the IDD. These include four products (three netCDF and one BUFR format) 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

New Activities

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

- IDD Cataloging (it's finally coming!)
- Begin offering a "data listing" page, to help find external sources of data
- Begin investigating a new support/ticketing system

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
- Revamp the data pages on the Unidata website

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

- 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.

A Note Regarding Support

With the departure of Tom Yoksas imminent, Mike Zuranski is taking over many of Tom's support duties. This includes responding to inquiries to data & dataflow support inboxes, as well as becoming the de facto point of contact for users, data providers, stakeholders and other partners. To say that Tom has done a lot in these areas is quite the understatement, and Mike is not a 1:1 replacement for Tom either, so there may be some minor differences or speed bumps during this transition. Legendary service continues to be our priority, of course, so please reach out to Mike Zuranski with any questions, concerns or other feedback.

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.

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 *March, 2023*

Status Report: Data Standards and Technical Outreach

November 2022- March 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; and
2) the development of the Zarr version 3 specification.

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.

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:

- Planning for the 2023 CF Workshop has begun.
- 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 hope 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.
- 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 2022*

Status Report: Unidata Community Equipment Awards

Sponsored by the National Science Foundation

November 2022- March 2023

Executive Summary

The review panel for the 2023 Community Equipment Awards will convene immediately after the March 2023 Users Committee meeting.

Questions for Immediate Committee Feedback

None at this time.

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.

The Users Committee met in early November 2022 to discuss possible adjustments to the wording of the Equipment Awards Call for Proposals to clarify what types of equipment are fundable through the program.

Relevant Metrics

Since taking over the management and administration of the Equipment Awards program in 2003 on behalf of the NSF, Unidata has made 104 awards totaling over $1,300,000.

Prepared *March 2023*

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

November 2022- March 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 Himiwari 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 Himiwari 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 FNEXRAD, NEXRAD3 and NIMAGE Datastreams

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

The **NEXRAD3** feed content was changed with the replacement of "high resolution" products with "super res" products. Since, we have obtained access to all of the Level 3 products being created at NEXRAD sites including some that have never been distributed in NOAAPort and others that were dropped from the mix provided in NOAAPort. The question posed at the beginning of this status report is if any of these products should be added back into the mix that is distributed in the IDD **NEXRAD3** feed?

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., NCAR twice per year power downs; 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.

Prepared *March, 2023*

Status Report: Internet Data Distribution(IDD)

November 2022- March 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 now has access to NEXRAD Level 3 products, some of which are not being distributed in the NOAAPort SBN and others that have been removed from the SBN. Does the committee think that some of these products should be added back in the IDD? A list of all NEXRAD Level 3 products can be found in:

https://docs.google.com/spreadsheets/d/1fECewX3Lp6sDAY5o-CFP_mx8wyLyRdtFaNVv-95qXW8/edit?usp=sharing

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:

```
20230313

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume 112587.060 M bytes/hour
Average hourly volume  72801.843 M bytes/hour

Average products per hour     488484 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)   number/hour
CONDUIT                 16296.993    [ 22.385%]    46647.115    94108.812
SATELLITE               15240.204    [ 20.934%]    20409.388     6639.250
```

```
NIMAGE                 9333.192    [ 12.820%]    13054.727     7656.146
NGRID                  7335.496    [ 10.076%]    12386.507    67391.417
NEXRAD2                7168.650    [  9.847%]     8619.310   103065.521
HDS                    5007.272    [  6.878%]    10195.029    31178.750
FNEXRAD                3936.429    [  5.407%]     4187.288     9120.833
EXP                    3828.463    [  5.259%]     7054.491    23738.104
NEXRAD3                2589.320    [  3.557%]     3150.517    86843.750
UNIWISC                1011.400    [  1.389%]     1155.513      963.479
FSL2                    697.952    [  0.959%]     2154.534     1473.396
NOTHER                  265.265    [  0.364%]      759.417       46.625
IDS|DDPLUS               89.251    [  0.123%]      101.508    55845.938
LIGHTNING                 1.956    [  0.003%]        5.344      411.521
```

Data Distribution:

IDD CONDUIT feed:

CONDUIT contents have not changed since the upgrade of GFS to version 16.

IDD FNEXRAD 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.

We are still waiting for access to the full suite of NEXRAD Level III products to be opened. When this access is established, we will, upon guidance from the Users Committee, resume redistribution of some products that were removed from the NOAAPort SBN.

Experimental HRRR feed to eventually be replaced by RRRS:

Unidata used to receive experimental High Resolution Rapid Refresh (**HRRR**) grids (both 2D and 3D fields) in an LDM/IDD feed from NOAA/GSL and feed these products to a small number of university sites on **hrrr.unidata.ucar.edu** (which is also known as **lead.unidata.ucar.edu**). Once the HRRR data went operational, NOAA/GSL stopped creating experimental **HRRR** output. The experimental **HRRR** is, however, being replaced by the **RRFS** (Rapid Refresh Forecast System) in NOAA/GSL. We have requested a feed of these data, but we have been told that the **RRFS** is still a few months away from being available.

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 February 8, 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

- NCEP (operational) HRRR fields and forecasts times were added to the IDD CONDUIT datastream.

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. More on the most recent of these activities can be found in the LDM status report.

- 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 v6.14.5. A new LDM distribution is being tested on Unidata ingest machines, and should be ready for release *soon*.

Relevant Metrics

- Approximately **533** machines at **188** 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.

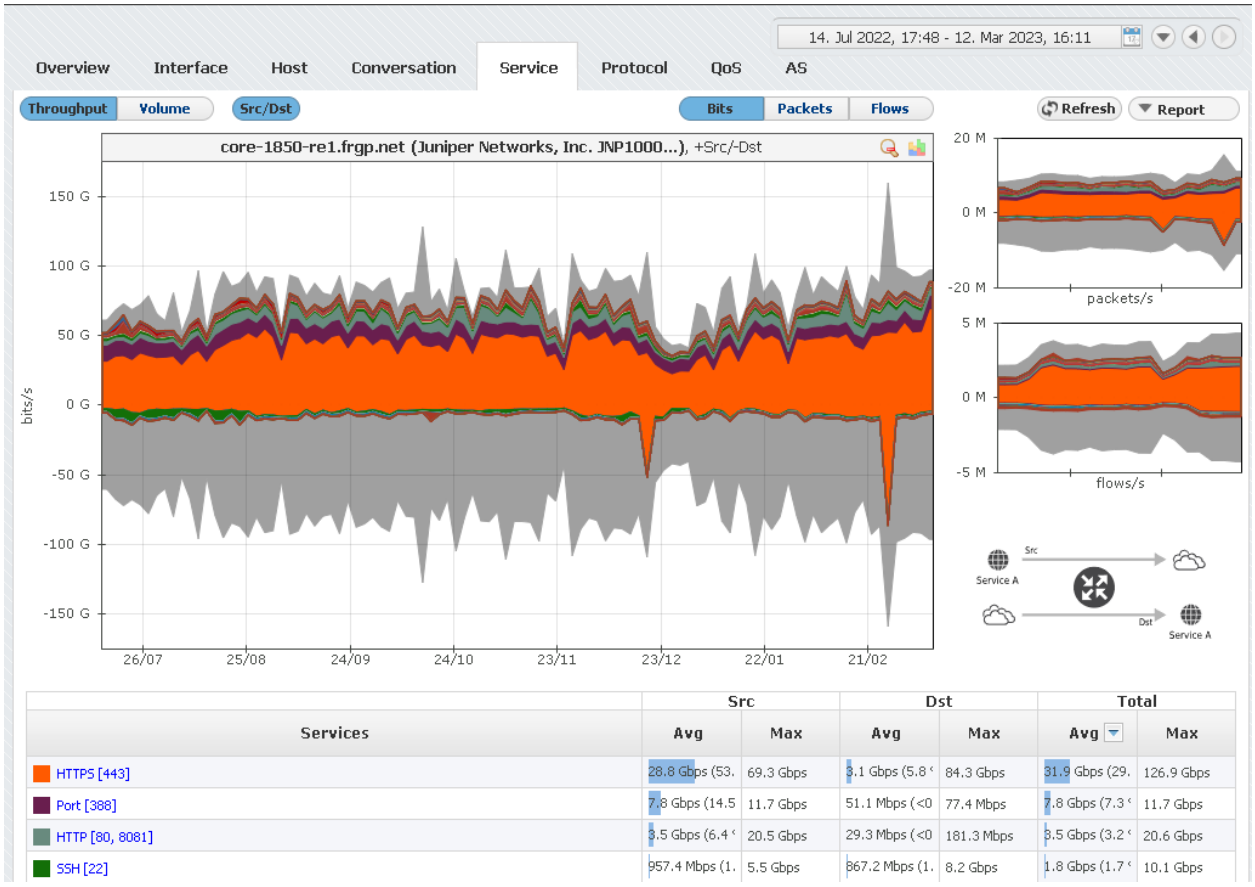
  Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service. We have learned about sites not sending real-time statistics during these kinds of support activities, and a number of times the impediment to sending in stats is firewall configurations at the 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 July 14, 2022 through March 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 7.8 Gbps (~84.24 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 Ave Max	Dst Ave Max	Total 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

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 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

Prepared *March, 2023*

Status Report: IDV with RAMADDA

November 2022- March 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 noticed that many advanced features of the IDV, such as formula and trajectory, have not been widely used in the community and many data servers that the IDV can directly access are less knowledgeable to the end users, and we would like to provide help to classes, research groups and project teams. I wonder if committee members can help to establish such connections.

Activities Since the Last Status Report

IDV Releases

The IDV 6.1 was released on August 1 of 2022.

The IDV 6.1u2 was released on August 20 of 2022.

The IDV 6.1u2 was released on September 8 of 2022.

IDV System Changes

__Upgrading netCDF-Java to OpenJDK 11__

l am working on upgrading to openJDK 11, the current version IDV can be built with either openJDK 8 or openJDK11, but it can only run with openJDK 8. If it is run with JDK11, the HiDPI problem will appear and all lines (used to be solid lines) become dashed lines and the mouse cursor not point to the correct geo location. The IDV is not correctly scaled to the monitor they are in. This problem will eventually require

upgrading the Java 3D to 1.7 version.

__VisAD upgrading Java 3D 1.7 version__

VisAD project: Working with other developers to migrate VisAD library to be compatible with the Java 3D 1.7 library, the release engineering process of Java 3D library in the OpenJDK 11 is different compared to the OpenJDK 8 that includes the lib/ext structure. We are exploring the best practice to add the Java 3D library in the upgrading processes.

__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 employes 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

__USGS Web Map Server Support__

It has been a while since IDV can access Web Map Servers of USGS and some other government agencies. The problem is the library conflicts causing the “handshake exception” when the IDV tries to access these Map servers. After many efforts to isolate the specific libraries causing the exception, we successfully upgraded those libraries to re-enable the IDV to access a rich set of USGS WMS servers.

__Non Geo-Located data display__

In general the IDV is used to display and analyze many types of geo-located dataset. With the new requests of the community, we added the support of displaying non geo-located time series dataset. Furthermore, many statistical analyses of the current IDV formulas such as: area average, level average, max, min, percentile, and level sum, the outputs of these analyses can

be generated as non geo-located data and exported as CSV or netCDF data formats. This new feature allows users to take advantage of the IDV accesses of many data servers and prepare the data for purposes of machine learning and other scientific usages.

__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.

KIOSK IDV Project

In collaboration with UCAR Center for Science Education and Computational Information Systems Laboratory, the project developed an extended IDV package for a Real-Time Weather Museum Touchscreen. This new real-time weather museum touchscreen display will undergo further usability testing to eventually join other weather and climate exhibits at NCAR's Mesa Lab in Boulder, CO, and at the NCAR-Wyoming Supercomputing Center Visitor Center in Cheyenne, WY.

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).

IDV and RAMADDA Training, Conference Attendance and Presence

__2022 AGU Fall Meeting__

- Yuan participated in the 2022 AGU fall meeting.

__2023 AMS Annual Meeting__

- Advances in Urban Flash Flood Model Forecasting System

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.

Prepared May, 2022

Status Report: Information Technology

November 2022- March 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

- Progress has been made with the migration of Unidata web content to new hosts to accommodate the UCAR webmod2 efforts.
- 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 competed and/or anticipated are;
  - centralized DNS (InfoBlox, complete)
  - centralized backups (delayed, in progress)
  - centralized telephony 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 March, *2023*

Status Report: LDM

November 2022- March 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 elaborate on 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:

- Installation:
    - Adding support for non-root installation of LDM.
    - Creating binary-only packages for client-side installation of LDM.
- ldmd(1):
    - Improved daemonization of top-level LDM server.
- NOAAPort:
    - Response and support of NOAA migration of carrier satellites, including migrating Unidata assets to point at new service.
    - Updated GEMPAK tables.
- rtstats(1):
    - Made error reporting more robust.
    - Re-architected and re-factored some code.
    - Eliminated small memory-leak.
    - Improved man(1) page.
- Logging:
    - Added some missing log messages and eliminated some superfluous ones.
- DevOps:

        - Adjusted code and build procedures to accommodate the use of GitHub actions for continuous-integration and continuous-deployment.
        - Created GitHub action script for continuous integration and continuous-deployment.
- Miscellaneous:
    - Ported some pqinsert(1) and rpc(3) code to MacOS.
- Documentation:
    - Documented that the systemd(8) parameter “RemoveIPC” must be “no” to prevent deletion of the upstream LDM database when any LDM user’s crontab(1) job terminates.
- Released versions 6.14.1, 6.14.2, 6.14.3, 6.14.4, and 6.14.5.
- Support:
    - Answered many questions from Universities, NOAA, US Military, and corporations.
    - Diagnosed a problem with NWS’s temporary “Terrestrial SBN” replacement for NOAAPort (it uses the LDM to distribute all NOAAPort data to WFOs while their dishes are being repointed to Galaxy 31).
    - Gave a virtual workshop on the LDM to NOAA/NWS Alaska, Hawaii, and Guam.

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).

Non-backward-compatible upgrades to operating systems can create problems. For example, Google notified us that their LDMs were failing to execute properly. Investigation showed that this was due to the shared-memory segment that holds the upstream LDM database being mysteriously deleted at random times. Further investigation revealed that the culprit was the system utility systemd(8) when the value of its “RemoveIPC” parameter was “yes” as this caused the shared-memory segment to be deleted whenever an LDM user’s crontab(1)-initiated job terminated.

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 will gradually assume 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 March *2023*

Status Report: McIDAS

November 2022- March 2023

Tom Yoksas, Mike Zuranski

Executive Summary

Unidata McIDAS-X/-XCD v2022 is in the last phase of being released (documentation modifications are the last thing remaining).

Questions for Immediate Committee Feedback

Unidata will be ceasing support for McIDAS-X use in the community. What efforts should be continued in order to ease the community transition away from McIDAS (e.g., continuing the mcidas-x email list, continuing to support downloads of the package with clear indication that new development will not be forthcoming, etc.)?

Activities Since the Last Status Report

Aside from routine updates/bug fixes to existing code and tables, the main area of activity recently has been the impending release of Unidata McIDAS v2022. The main development area that we have been involved in is enhancing the GOES-R servers.

Current Activities

- Support use of McIDAS-X/XCD in-house and in the community
- 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.

The Unidata v2022 release includes these SSEC v2022.1 release features:

- Added support for McIDAS-X on **macOS** systems with Apple silicon processors, and preliminary support for Red Hat Enterprise Linux 9.
- Added new calibration and navigation modules for the **new MTG (Meteosat Third Generation) ADDE servers.** The MTG FCI servers are being released as a preliminary release separate from McIDAS-X. These will be the first operational servers written in Python and will be officially released when they are fully tested with data from the satellite.

- The NEXRAD servers and McIDAS-XCD modules were updated to handle the new **NEXRAD super resolution WSR Level 3 Products** made available in February 2022.
- Updated **GOES-R Series ABI and GLM** servers with additional bug fixes and enhancements, including adding the NOAA STAR enhancements used in the GOES Image Viewer to -XRD, and added support of new variables in Level 2 Aerosol Detection Smoke and Dust fields, as well as in Level 2 & 3 NOAA Vlab GLM products.
- **Updated the VIIRS servers** to read VIIRS EDR files and to improve scaling for some bands.
- **Updated the Himawari servers** to correctly identify Himawari-9 images and to use a new format of HimawariCast data available from Eumetsat.
- **The GRIB servers and GRD* commands** were updated to list and display GRIB data with entire atmosphere levels (EATM) and additional sigma levels (SIGM).
- The **mcinet.sh** script was updated allow the McIDAS-X system service to be controlled by **systemd** or **xinetd**, and to better report the status of the remote ADDE servers.
- **Updated USCOUNTY.MAP, USSTATE.MAP and USZONE.MAP files** with the most recent National Weather Service GIS-AWIPS Shapefile database updates.
- **Updated the McIDAS-XRD Package** included in the McIDAS-X installation files to include one new command (DAYNITE, which combines a visible and infrared image at the terminator) and seven new data files (STAR-ABI_*.ET, which are the standard enhancements used in the GOES Image Viewer).

Ongoing Activities

We plan to continue the following activities:

- Interest in McIDAS by non-core users

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

- Continued support of internal use of McIDAS-X/-XCD

New Activities

- None

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.

- McIDAS-X/-XCD 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.

Planned Activities

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.

New Activities

Establish a testbed for generating Level 2 products from GOES-East/West imagery and select model output. The intention is to be able to test vetted algorithms submitted by community members for a long enough period for the algorithms to be fully tested.

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 *March, 2023*

Status Report: netCDF

November 2022- March 2023

Ward Fisher , *Dennis Heimbigner* , *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 M1/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 **239** open issues for netCDF-C, **97** open issues for netCDF-Fortran, and **51** 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:

- The netCDF and netCDF-Java teams 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.
- The most recent release of netCDF-C, version 4.9.2, adds support for netCDF4 files based on HDF5 1.14.0.
- 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. 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.
- Introduction of 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 esupport, etc. With 1.5 FTE assigned to the project, the workload is significant.
- 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 architecture) are requiring additional overhead to ensure compatibility.

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).
- 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

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:

- **1,490,000** for netCDF-3
- **1,790,000** for netCDF-4
- **2,950** for ncZarr
- **2,560,000** for HDF5
- **144,000** for GRIB2
- **1,900,000** for ZARR

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

- **443** for netCDF-3
- **1,290** for netCDF-4
- **35** for ncZarr
- **40,000** for netCDF
- **22,700** for HDF5
- **1,680** for GRIB2
- **8,010** 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, 2022*

Status Report: Python

November 2022- March 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. The training aspect continues to be busy with two workshops taught at the AMS Annual meeting. We also continue to plan additional courses along with supporting the 2023 Unidata Users Workshop. 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 1.4.0 feature release (largely based around "spherical calculus" support) and a recent 1.4.1 bugfix release. We are planning a 1.5 release at the end of April which will largely consist of work done in fulfillment of our NSF CSSI award. Participation within the broader community continues with anticipated attendance of SciPy and direct participation on the matplotlib, cartopy, and conda-forge projects. It has become increasingly difficult to dedicate time to these efforts given the full portfolio of responsibilities on the team.

Questions for Immediate Committee Feedback

Nothing at this time.

Activities Since the Last Status Report

Python Training Efforts

Continued collaboration of Unidata teams providing material development, instructional design practices, and Science Gateway infrastructure support have enabled multiple synchronous workshops at the turn of the year. After our Python Readiness course at CSU in the fall, we prepared two workshops for the 2023 AMS Annual Meeting: a day-long professional short course surveying unique weather data workflows in Python and a quarter-day intro to data basics in Python for the 2023 AMS Student Conference. We are supporting development for regional workshops and the Unidata Users Workshop this summer.

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 are supporting the upcoming Pythia Summer Hackathon in Boulder, June 2023. We are still transitioning our "Unidata Python Training" resource to its own Pythia Cookbook(s) with

support from regular MetPy developers- and community-call participants.

Progress has been made on the following:

- Unidata delivered two AMS workshops on Python, one full-day short course and one quarter-day workshop for the Student Conference
- Unidata continues to be a primary collaborator on Project Pythia, particularly through support of the growing Cookbook ecosystem. We are testing Cookbooks as the future medium for our “Python Training” example gallery
- 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 1.4.0 was released in late December 2022 (despite plans and efforts to release earlier) with a variety of fixes and enhancements including:

- Added SWEAT, CCL, and convective temperature sounding calculations (community contributions)
- Support for improved spatial derivative calculations on projected and lat/lon (spherical) grids
- Added ArrowPlot and RasterPlot, as well as other various enhancements, to the simplified plotting interface
- Added many more examples to MetPy’s gallery and in the documentation itself
- GINI support works as an xarray backend
- Better support for working with other libraries using Pint for unit-handling

The MetPy team continues to work to increase the release cadence for the project, nominally trying to make a release every other month. This is aimed to avoid having releases slide to get “one more thing” in the release, and instead more readily get developments in the hands of the user–whenever they choose to upgrade. Success in doing so has been mixed, with the most recent push to get improved derivative calculations (aka “spherical support”) into the 1.4 release. MetPy 1.4.1 was released in early March 2023 with a few minor code and documentation fixes.

Moving forward, 1.5 is planned for release at the end of April 2023. This is planned to include the long-promised support for plotting fronts and analysis from the WPC. 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.

To lower the barrier to project participation by community members, Unidata and regular participants in the existing “MetPy Dev Call” (biweekly on Google Meet, see MetPy GitHub discussion #2624) planned for the first “MetPy Community Call”. This community call is a synchronous meeting which shifts the discussion focus to open-floor demonstrations of successful (and unsuccessful) MetPy workflows, and removes the expectations that

participants only contribute through “development”. The first call was advertised at AMS 2023 and delivered 19 Jan 2023 to around eight attendees, including five new in the space. This call occurs approximately monthly.

Progress has been made on the following:

- MetPy 1.4.0 released late December 2022
- MetPy 1.4.1 released early March 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 2489 followers.
- 13 theses or peer-reviewed publications have cited or mentioned MetPy so far in 2023; this is in comparison to 53 in all of 2022, 57 in 2021, and 43 in 2020.

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 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

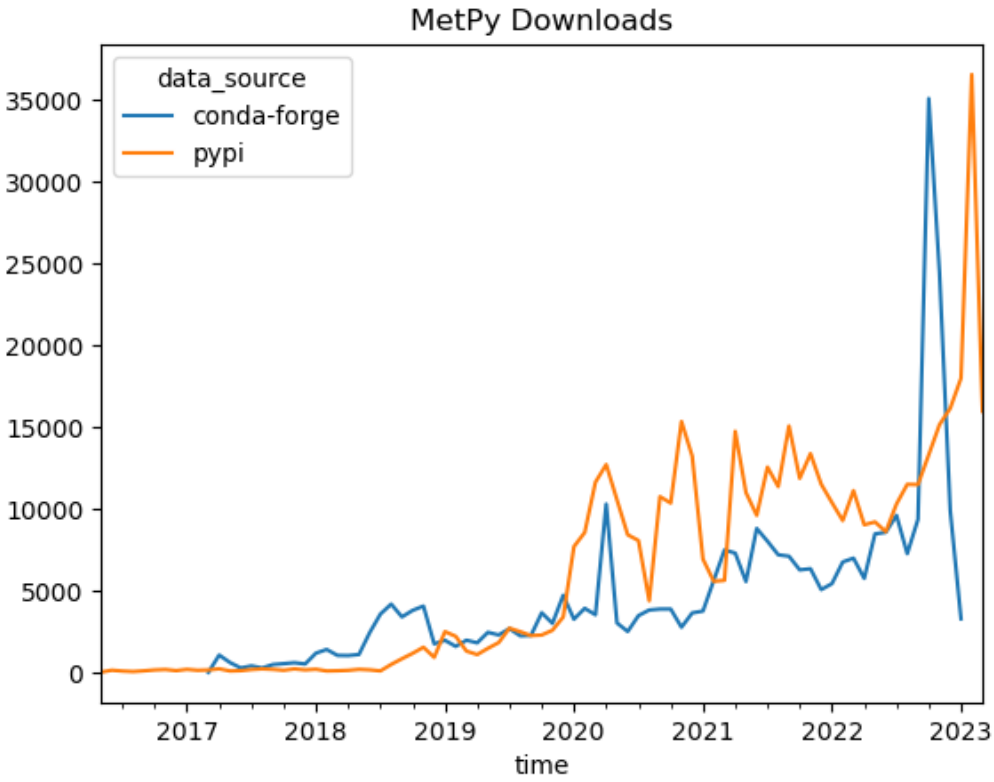
Over the next three months, we plan to organize or take part in the following:

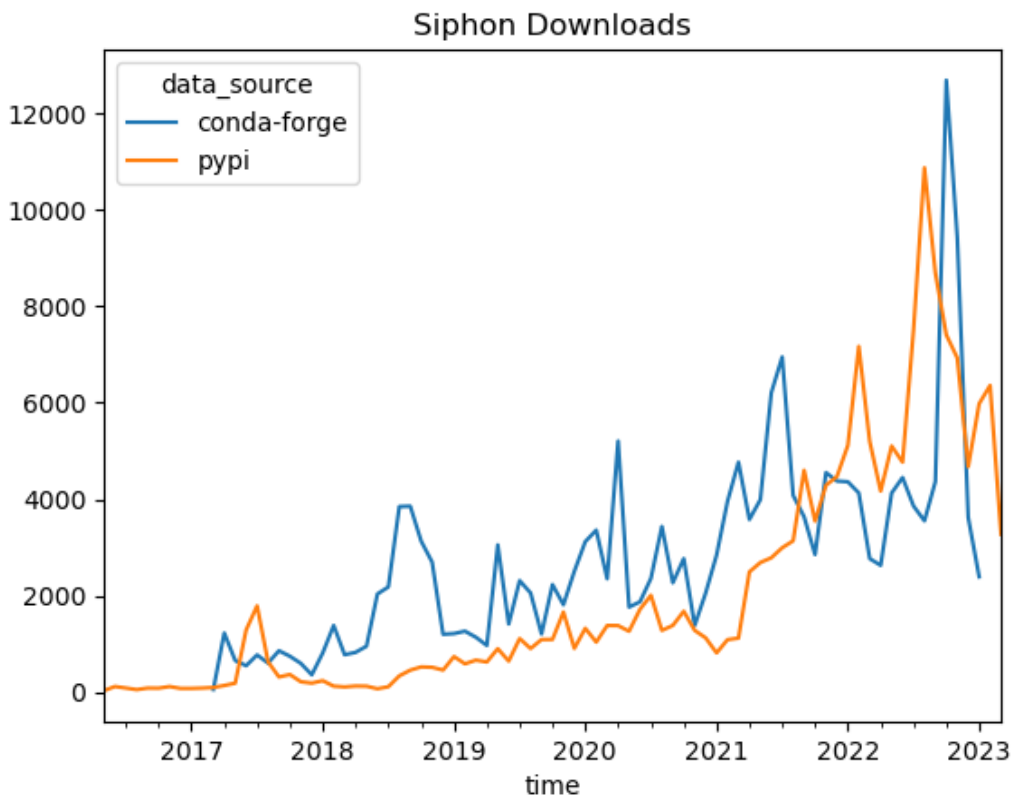
- Deploy the "Unidata Python Training" gallery and training materials as Project Pythia Cookbook(s)
- Release MetPy 1.5 with a variety of features, including frontal analysis plotting and tools to facilitate access to cloud-based stores of radar and satellite data
- Submit a short course proposal for the 2024 AMS Annual meeting
- Engage in continued support of Project Pythia and adjacent UCAR Python education efforts

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

- Present on updates to MetPy at the 2024 AMS Annual Meeting
- 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





MetPy

- 95% test coverage
- Watchers: 59
- According to GitHub, 338 repositories and 38 packages depend on MetPy
- Downloads for the releases made in the last year (Conda + PyPI):
    - 1.3.0: 26467
    - 1.3.1: 92753
    - 1.4.0: 35100
- Since 1 April 2022
    - Active Issues: 123 (54 created, 56 closed)
    - Active PRs: 218 (188 created, 196 closed)
    - External Issue Activity: 41 opened, 85 comments
    - External PR Activity: 13 opened, 45 comments
    - Unique external contributors: 32
    - Stars: 94 (1004 total)
    - Forks: 5 (337 total)
    - Commits: 482
- Since 1 March 2022
    - Active Issues: 178 (112 created, 88 closed)
    - Active PRs: 399 (378 created, 373 closed)
    - External Issue Activity: 82 opened, 124 comments
    - External PR Activity: 44 opened, 79 comments

    - Unique external contributors: 61
    - Stars: 159 (1004 total)
    - Forks: 7 (337 total)
    - Commits: 815

Siphon

- 98% test coverage
- Watchers: 14
- According to GitHub, 158 repositories and 15 packages depend on Siphon
- Since 1 April 2022
    - Active Issues: 4 (2 created, 1 closed)
    - Active PRs: 94 (85 created, 83 closed)
    - External Issue Activity: 2 opened, 7 comments
    - External PR Activity: 0 opened, 0 comments
    - Unique external contributors: 5
    - Stars: 16 (173 total)
    - Forks: 0 (61 total)
    - Commits: 93
- Since 1 October 2021
    - Active Issues: 12 (5 created, 4 closed)
    - Active PRs: 202 (183 created, 190 closed)
    - External Issue Activity: 3 opened, 14 comments
    - External PR Activity: 6 opened, 5 comments
    - Unique external contributors: 11
    - Stars: 25 (173 total)
    - Forks: 0 (61 total)
    - Commits: 254

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 2022*

Status Report: Support

November 2022- March 2023

Jennifer Oxelson, Tom Yoksas, Mike Zuranski, UPC Staff

Executive Summary

The currently-used eSupport package is long in the tooth. A small group of people within Unidata have volunteered to start vetting replacement packages.

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

Training

Unidata training/workshop information can be found in the Community status report.

Looking for a new in-house support package

- The currently-used eSupport package is long in the tooth. A small group of people within Unidata have volunteered to start vetting replacement packages.

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

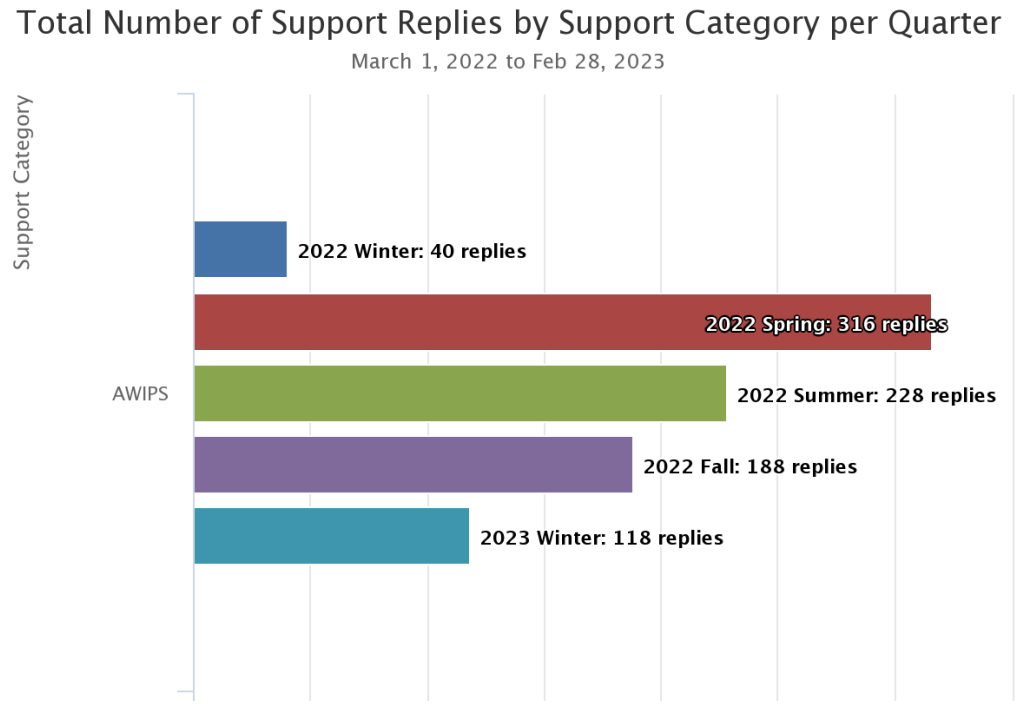
Since January 26, 2006 approximately 69516 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

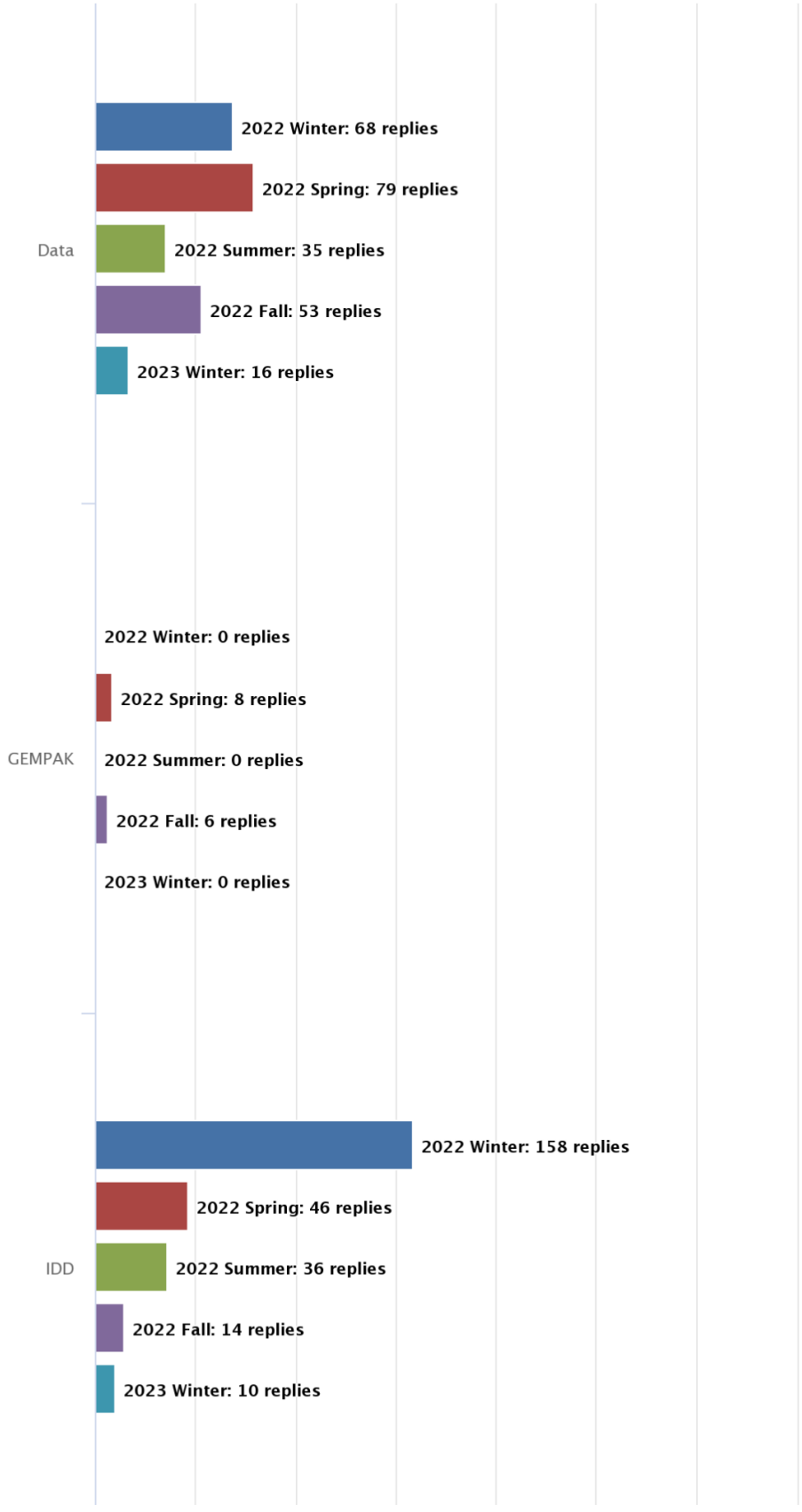
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

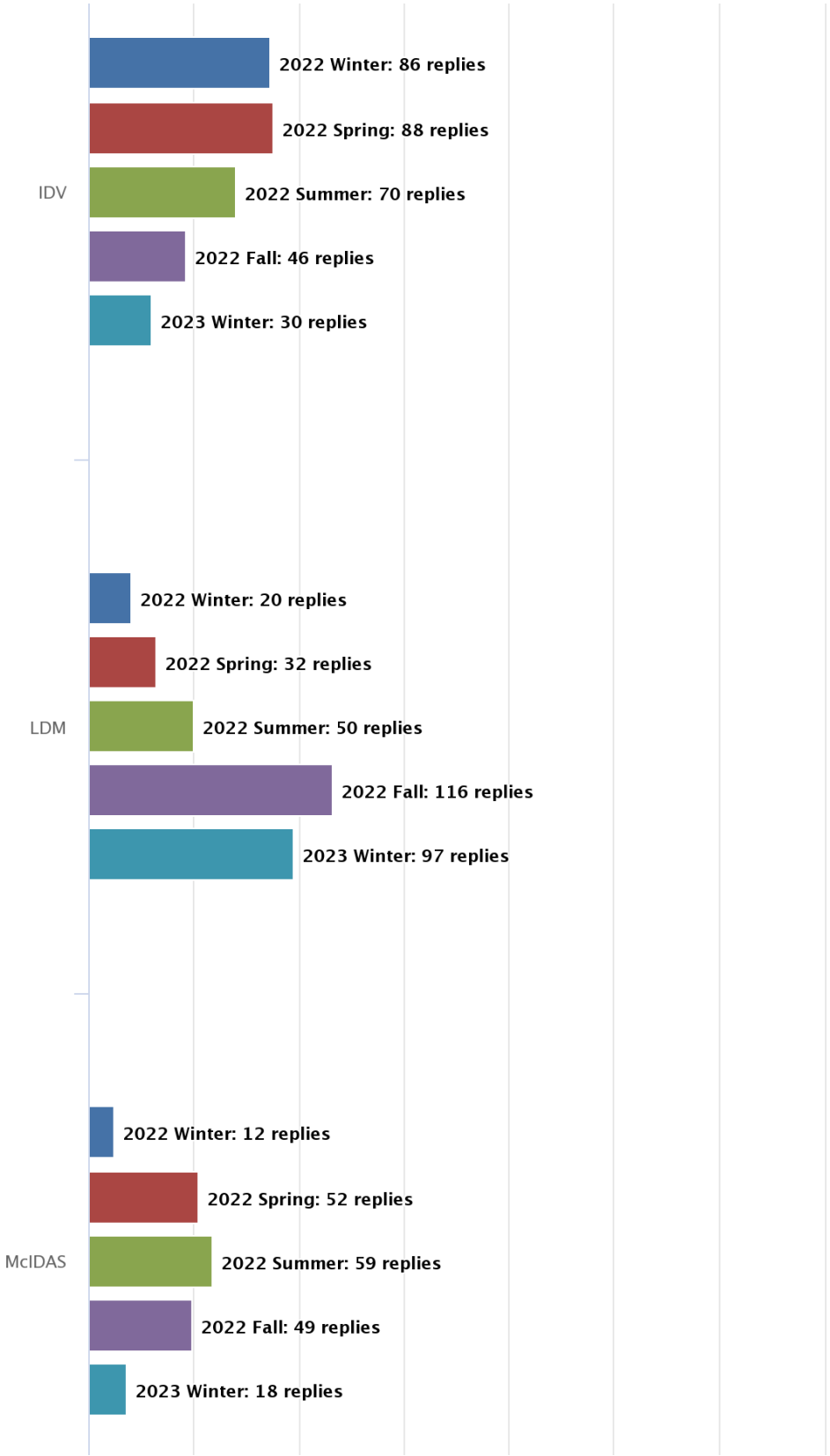
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 12 month period from **March 1, 2022 until February 28, 2023**.

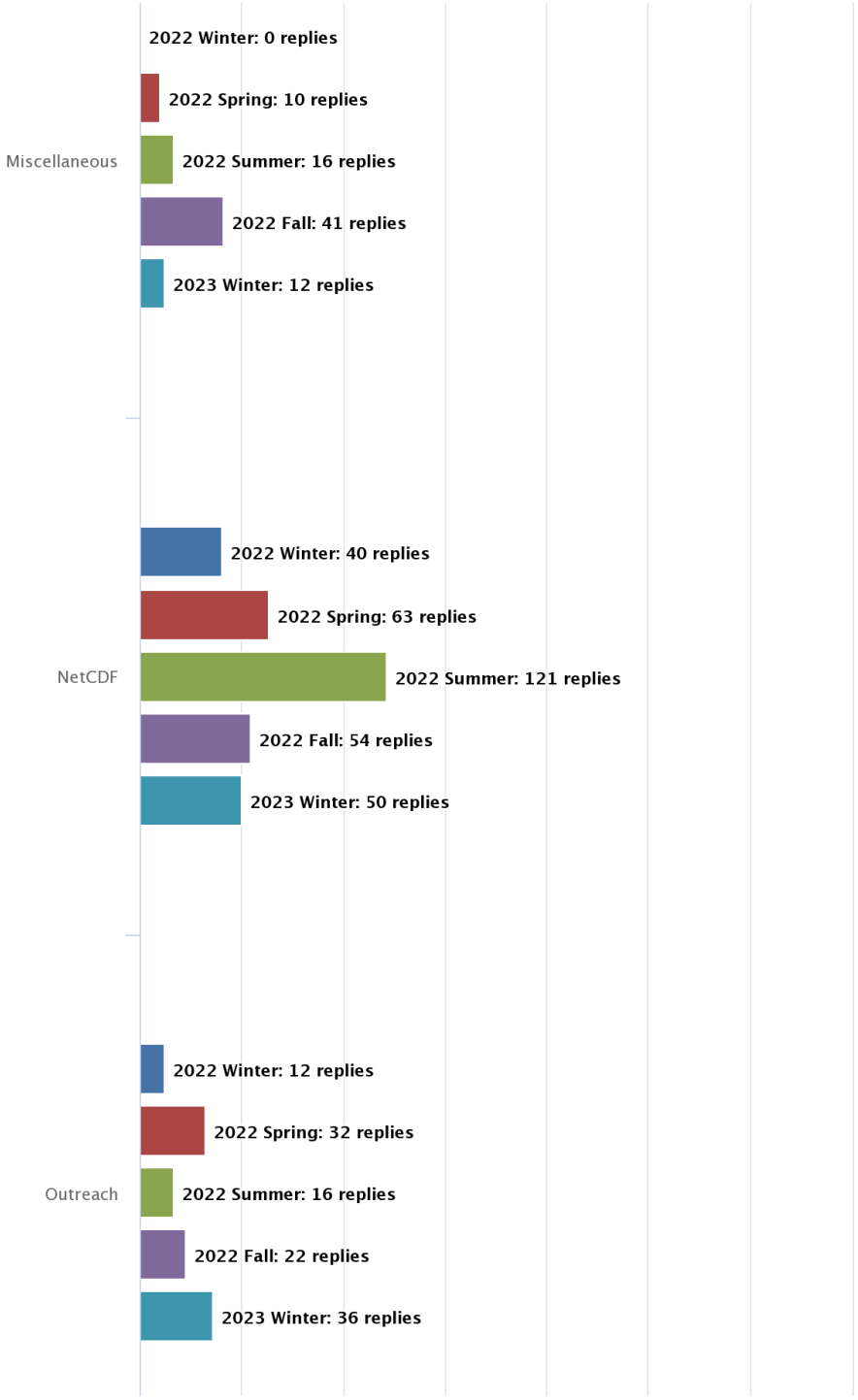
The quarters shown are defined as:

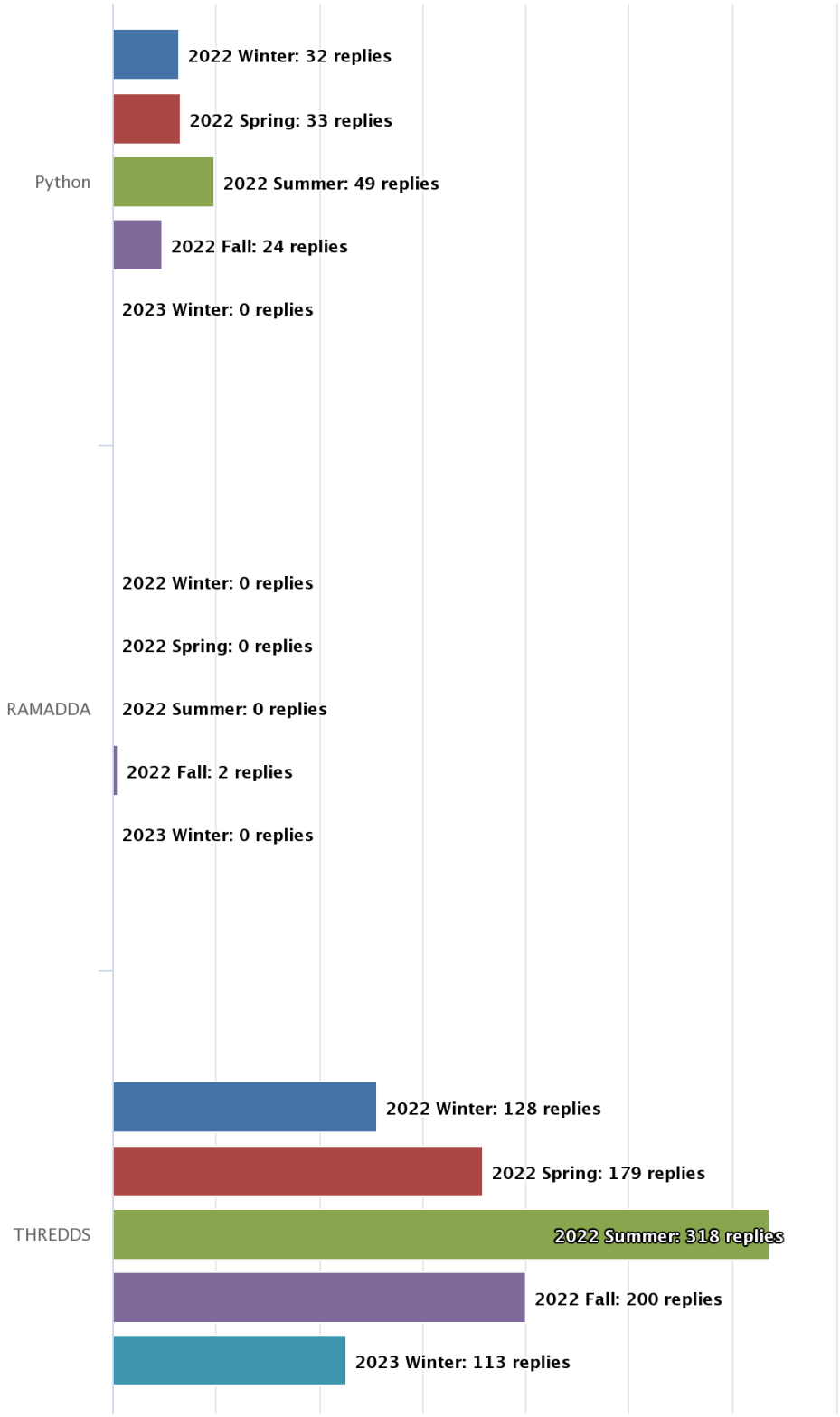
Winter:	**Spring:**	**Summer:**	**Fall:**
January, February, March	April, May, June	July, August, September	October, November, December

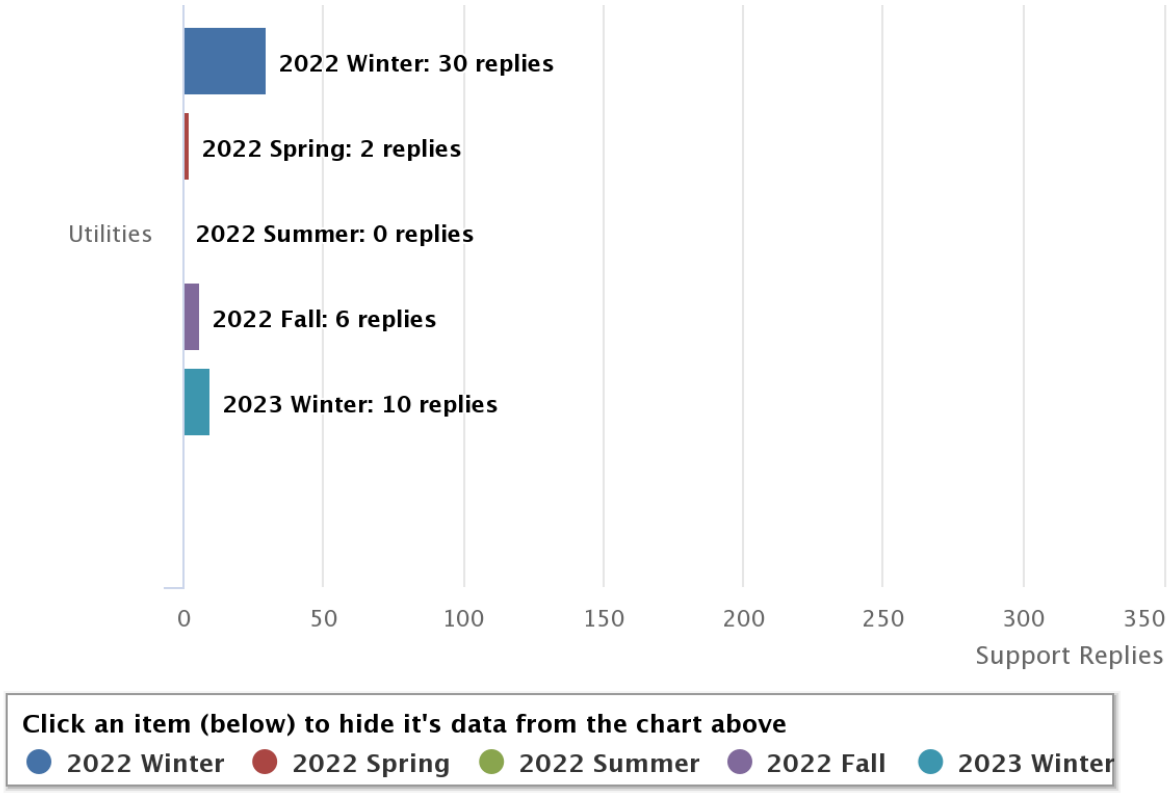



Data
2022 Winter: 68 replies
2022 Spring: 79 replies
2022 Summer: 35 replies
2022 Fall: 53 replies
2023 Winter: 16 replies
GEMPAK
2022 Winter: 0 replies
2022 Spring: 8 replies
2022 Summer: 0 replies
2022 Fall: 6 replies
2023 Winter: 0 replies
IDD
2022 Winter: 158 replies
2022 Spring: 46 replies
2022 Summer: 36 replies
2022 Fall: 14 replies
2023 Winter: 10 replies


IDV
2022 Winter: 86 replies
2022 Spring: 88 replies
2022 Summer: 70 replies
2022 Fall: 46 replies
2023 Winter: 30 replies
LDM
2022 Winter: 20 replies
2022 Spring: 32 replies
2022 Summer: 50 replies
2022 Fall: 116 replies
2023 Winter: 97 replies
McIDAS
2022 Winter: 12 replies
2022 Spring: 52 replies
2022 Summer: 59 replies
2022 Fall: 49 replies
2023 Winter: 18 replies


Miscellaneous
2022 Winter: 0 replies
2022 Spring: 10 replies
2022 Summer: 16 replies
2022 Fall: 41 replies
2023 Winter: 12 replies
NetCDF
2022 Winter: 40 replies
2022 Spring: 63 replies
2022 Summer: 121 replies
2022 Fall: 54 replies
2023 Winter: 50 replies
Outreach
2022 Winter: 12 replies
2022 Spring: 32 replies
2022 Summer: 16 replies
2022 Fall: 22 replies
2023 Winter: 36 replies


Python
2022 Winter: 32 replies
2022 Spring: 33 replies
2022 Summer: 49 replies
2022 Fall: 24 replies
2023 Winter: 0 replies
RAMADDA
2022 Winter: 0 replies
2022 Spring: 0 replies
2022 Summer: 0 replies
2022 Fall: 2 replies
2023 Winter: 0 replies
THREDDS
2022 Winter: 128 replies
2022 Spring: 179 replies
2022 Summer: 318 replies
2022 Fall: 200 replies
2023 Winter: 113 replies



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.
- Support for the legacy visualization packages GEMPAK and McIDAS has decreased over the past several years, most likely due to GEMPAK users' investigations of use of AWIPS and Python/MetPy.
- Support for AWIPS has steadily increased and has exceeded that for GEMPAK over the past couple of years.
- Support for Python scripting using MetPy is 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, GEMPAK, and McIDAS requires significant resources, but is difficult to assess.
- Not all support records were indexable for this report. Given this, the above numbers are an **underestimate** of the actual support being provided by the UPC.

Additional User Support 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.

Status Report: THREDDS

November 2022- March 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

- Would any of your automated data access workflows be disrupted if the format of the THREDDS catalogs were to change (e.g. from XML to JSON)?
    - I.e., are you explicitly parsing catalogs or relying on a provided API like Siphon?

Activities Since the Last Status Report

NetCDF-Java

- NetCDF-Java remains on version 5.5.3 (as of 3/15/23).
- 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 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 focuses for the TDS currently are stability, performance, and cloud support.
- The THREDDS team has begun working on the TDS microservices, with progress being made on gRPC access (netcdf-Java), File Service (Python), and Catalog Service (Java).
- Plans have been made to develop a performance and benchmarking test suite for the TDS services.

Ongoing Activities

Server management

- We have implemented a new continuous deployment process for the Unidata THREDDS servers:
    - thredds.ucar.edu now 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

Maintenance

- Maintain thredds.ucar.edu and keep up with the addition of new datasets to the IDD.
- 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.

Development

- Zarr and NCZarr support
    - Continued work to implement write support for Zarr and NCZarr.
    - Continue development of the new filters module and add support for requested common filters, starting with ZStandard.
    - Add support for Zarr and NCZarr in the TDS.
- DAP4 support
    - Continued work to support the DAP4 protocol in netCDF-Java and the TDS
- Cloud Storage
    - Expand S3 support in netCDF-Java and TDS to effectively mirror that of local storage.
    - Expand testing for S3 support.
- Performance and benchmarking
    - Create automated benchmarking and regression testing tools for both netCDF-Java and the TDS.
    - Improve performance for TDS S3 access, particularly for large aggregations, to prevent potential server timeouts.
- TDS microservices
    - Continue planning a development for the next iteration of the TDS with a microservice-based architecture.

New Activities

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

- netCDF-Java
    - Improve the performance of scale/offset transformations
    - Release version 5.5.3
    - Continue to participate in the Zarr development community
    - Finish porting the gCDM (gRPC for the Common Data Model) module to netCDF-Java.
    - Begin developing other gRPC endpoints for netCDF-Java
    - Plan for a netCDF-Java version 6 with a limited public API

- TDS
    - Release version 5.5 of the TDS.
    - Continue to help the user community upgrade their servers to TDS version 5.x.
    - Work with Thomas Martin on AI/ML pipeline using the THREDDS Notebook Service

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
- TDS
    - Develop performance and benchmarking tools for the 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 support 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.
- TDS
    - Continue development of standalone TDS services
    - Continue to explore object storage as it relates to the TDS.
    - 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

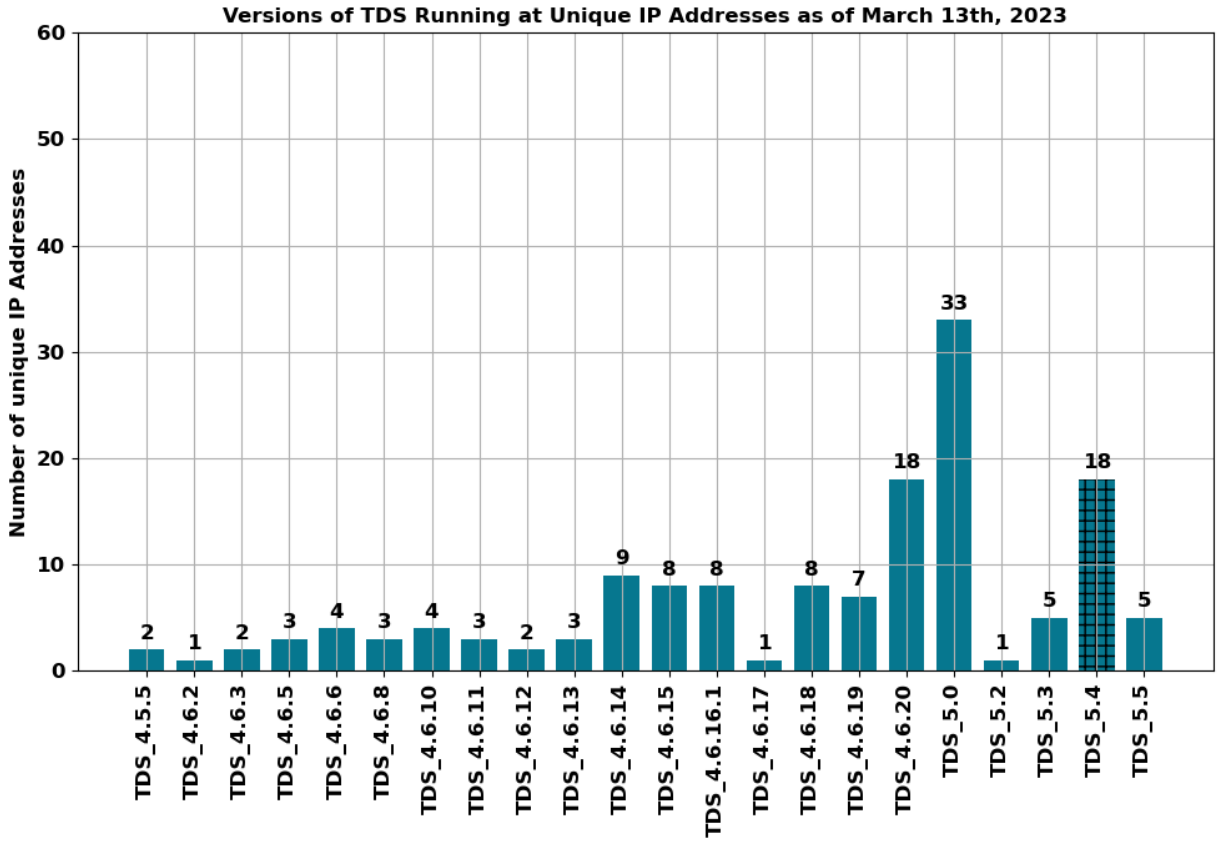
	2022-11 — 2023-02	2014-08 — 2023-02
TDS Startup (unique IP address count)	1759	41107

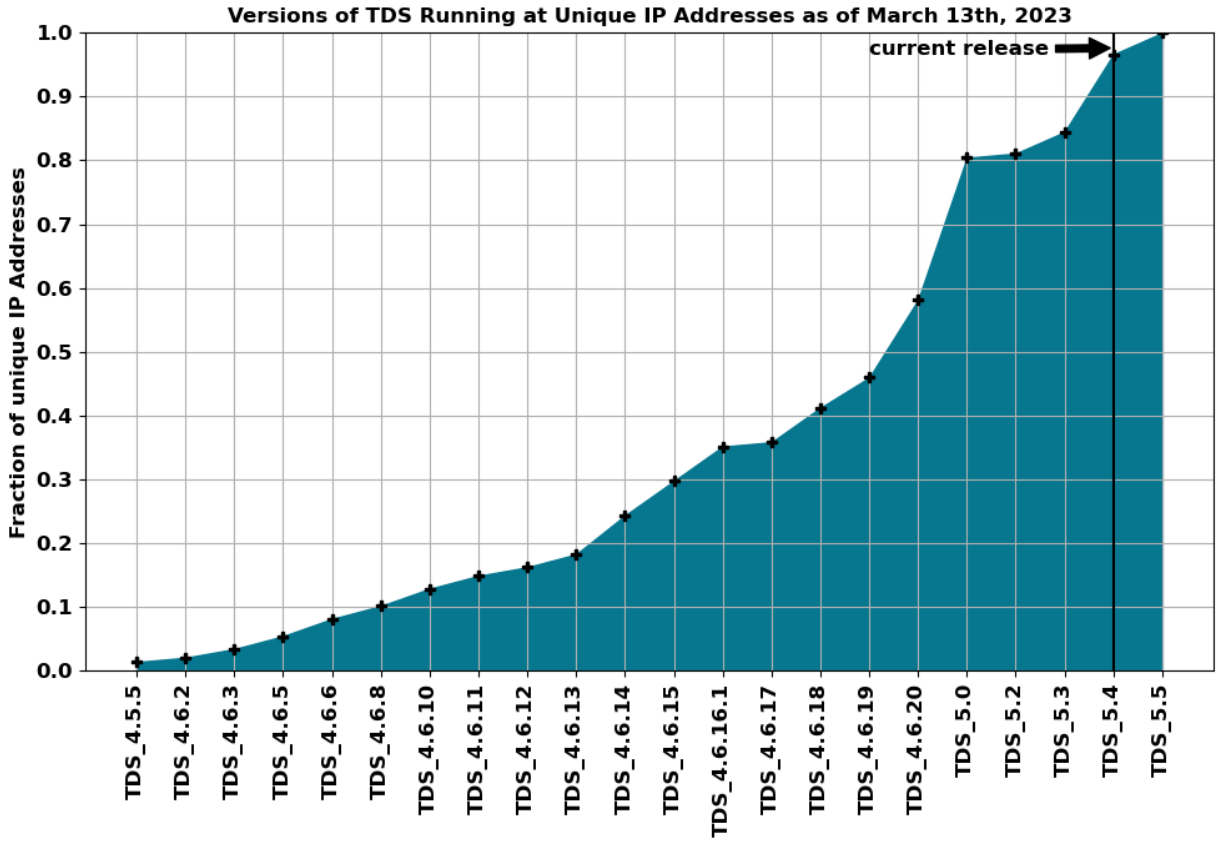
	Total Servers	Information page updated
Publicly Accessible[1] TDS count	148	76

Over the past 4 months, **1,759** unique IPs started up the TDS (Nov 2022 through Feb 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 **148** publically accessible TDSs running "in the wild". Of the **148** 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. The majority of the publicly accessible servers are running v4.6.13 or above . TDS v5.0 remains the dominant specific version running in the wild, despite several later releases.

[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 March 13th, 2023
Number of unique IP Addresses
0
10
20
30
40
50
60
2
1
2
3
4
3
4
3
2
3
9
8
8
1
8
7
18
33
1
5
18
5
TDS_4.5.5
TDS_4.6.2
TDS_4.6.3
TDS_4.6.5
TDS_4.6.6
TDS_4.6.8
TDS_4.6.10
TDS_4.6.11
TDS_4.6.12
TDS_4.6.13
TDS_4.6.14
TDS_4.6.15
TDS_4.6.16.1
TDS_4.6.17
TDS_4.6.18
TDS_4.6.19
TDS_4.6.20
TDS_5.0
TDS_5.2
TDS_5.3
TDS_5.4
TDS_5.5



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.

Prepared *October, 2022*