Users Committee Virtual Meeting Agenda

Friday, September 24th 10-11AM MDT

(Times are Mountain Standard Time)

10:00 - 10:10 Open Meeting (Kevin and Josh)

10:10 - 10:40 Director's Report (Mohan Ramamurthy)

10:40 - 10:55 Around-the-table Reports (Committee)

10:55 - 11:00 All other business (Kevin)

Friday, October 8th noon to 1PM MDT

(Times are Mountain Standard Time)

Noon - 12:05 PM Convene and outstanding items

12:05 - 12:25 PM AI/ML Discussion (Rio)

12:25 - 12:55 PM NOAA Update (Anne Myckow) (Time TBC)

12:55 - 1:00 Close Meeting

Friday, October 22nd 10AM to noon MDT

10:00 - 10:05 AM Convene and outstanding items

10:05 - 10:15 Action Items

10:15 - 10:25 AM Engagement Update (Jeff Weber)

10:25 - 11:00 AM Status Reports (Committee)

11:00 - 11:05 AM Break

11:05 - 11:35 AM Community Survey Results (Committee)

11:35 - 11:50 AM Blue Skies (Committee)

11:50 - noon All other business and close session (Kevin Goebbert)

Friday, November 5th 10AM to noon

(Times are Mountain Daylight Time)

10:00 - 10:05 Convene and outstanding items

10:05 - 10:20 McIDAS X Transition Timeline (UPC Staff)

10:20 - 10:35 Staff Demo 1 (Science Gateway?) (Julien Chatang TBC)

10:35 - 10:50 Next Triennial (2022) (Committee and Joshua Young)

10:50 - 11:00 Training Update (Nicole Corbin)

11:00 - 11:05 Break

11:05 - 11:20 Around-the-table Reports (Committee)

11:20 - 11:40 Committee Recommendations for LDM/IDD CONDUIT (Tom/Committee)

11:40 - 11:55 Survey Results Final Suggestions (Committee)

11:55 - Noon AOB and close session

11:55 - 12:00 Close Meeting

Friday, November 12th 10AM to 11AM (HOLD FOR SURVEY AS NEEDED)

10:00 - 10:05

10:05 - 10:20

10:20 - 10:35

10:35 - 10:50

10:50 - 11:00 AM

Status Report: Users Committee Actions

April 2021- October 2021

Unidata Program Center Staff

Actions from the Previous Meeting

Action 1

Committee members to review their ability to participate in in 2022 Unidata Users Workshop planning, by the May meeting sessions (Committee)

Result

At the May 2021 meeting, this task was deferred until fall in order to give precedence to working on the Community Survey.

Action 2

Finalize and distribute summary of Fall 2020 student panel discussion (Joshua Young)

Result

The summary was published in the Spring of 2021 and is available here.

Action 3

Look into status of putting GFS output back into CONDUIT (Anne Myckow)

Result

Anne gave the committee an update on the status of the GFS, GEFS, and CONDUIT in general at the October 8, 2021 meeting.

Prepared *October 2021*

Status Report: AI/ML

April 2021- October 2021

Rio McMahon

Areas for Committee Feedback

- What are the current use cases for AI/ML in your work?
- What are current pain points for AI/ML in your work? This includes everything from data collection, processing, model training/validation and model deployment.
- What sort of short, medium, and long term AI/ML goals would you like to see Unidata pursue?

Activities Since the Last Status Report

- Rio McMahon joined Unidata as a software engineer focusing on AI/ML.

Activities Ongoing/In-Progress

- Began exploratory work into machine learning model output visualization in AWIPS. The current goal for this work is to have a containerized solution (likely via docker) that allows for simple deployment of machine learning models into AWIPS. This will involve deploying an EDEX server in a container which can communicate with another container hosting a TensorFlow (machine learning framework) server. Intra-container communication will utilize the pygcdm API for transferring netCDF files. This work includes:
    - Working with the AWIPS team to update existing docker containers to act as EDEX ingest/request servers.
    - Working with the THREDDS team to develop a netcdf-java gCDM interoperable python API (pygcdm) for transferring netCDF files via gRPC.
    - Internal UCAR presentations to solicit preliminary feedback and to find users for an initial beta test.
    - Abstracts have been submitted to AGU/AMS Poster sessions on this work.
- Step 1 proposal submitted for NASA Advanced Information Systems Technology (AIST) grant. The focus is on “New Observing Strategies” (NOS) and the proposal is in collaboration with JPL. The proposal focuses on creation of a “data pipeline” which:
    - Amalgamates data from several data sources (including THREDDS).
    - Feeds the data into a machine learning model which can detect oceanic anomalies (harmful algae blooms, etc).
    - Displays the machine learning model output on a visualization platform to inform operators of in-situ observation equipment of potential events of interest.

  Communication between the different project components will occur via gCDM (including the original netcdf-java API and the pygcdm python API).

Future Activities

- Review community survey AI/ML results to understand community needs as they pertain to AI/ML. This will help shape the short, medium and long term efforts of the AI/ML team at Unidata.
- Work with internal (and eventually external) collaborators to develop and refine the machine learning visualization within AWIPS. A long term goal of this project is to help support research activities involving “trustworthiness in AI” by allowing machine learning practitioners to easily deploy models into an operational-emulating environment (AWIPS).
- Continue to pursue additional funding awards to support the AI/ML program at Unidata.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Managing Geoscience Data**
   Developing tools that allow for extant geoscience data processing pipelines to be adapted for use in AI/ML (containerization of EDEX servers)

2. **Providing Useful Tools**
   AI/ML tools currently being developed are designed to support scientists exploring research questions involving trustworthiness in AI. Further, tools (pygcdm) are also being developed to support underlying networking infrastructure for established Unidata products such as the netCDF file format and the netcdf-java libraries.

3. **Supporting People**
   At this point, the AI/ML team is soliciting feedback in the form of the Unidata community survey and other avenues (presentations, conferences, etc). Once more feedback has been collected on what the Unidata community needs with regards to AI/ML, a more comprehensive community support strategy will be developed and implemented.

Prepared *October 2021*

Status Report: AWIPS

April 2021- October 2021

Tiffany Meyer, Shay Carter

Areas for Committee Feedback

We will be adding Himawari data to AWIPS in the near future, does the community have interest in seeing it added to the IDD or other services?

Activities Since the Last Status Report

AWIPS

Unidata's Jetstream production EDEX server continues to serve real-time weather and geographic data to CAVE clients and the python-awips data access framework API. By offloading the processing of one very large data feed (NEXRAD3) to a separate EDEX Ingest Node, the current implementation of edex-cloud is now capable of processing and serving more data than ever before. The distributed architectural concepts of AWIPS allow us to scale EDEX in the cloud to account for the size of incoming data feeds. With this distributed architecture, we were able to assist Texas A&M in setting up their very own EDEX server. The server they built is actually larger and consists of more machines than the current Unidata one. We were able to work closely with them this summer when we released our latest version of AWIPS, which had some significant updates for EDEX. We are currently looking into adding an additional machine for our own EDEX system, to help handle the increase in satellite data we've been processing and some new additional data we are planning on adding.

Since the last status report, we were able to finally push out our first official release of AWIPS since Michael James passed away (version 18.2.1). This update included many additions to our EDEX server that we had deployed locally at Unidata, that are now available for anyone running their own EDEX, additional datasets, as well as significant menu changes and added functionality on the CAVE side of things.

A link to all of our AWIPS release notes can be found here. With this being a direct link to the major 18.2.1 release. To summarize:

- A variety of new datasets were made available (MRMS and GOES Satellite)
- The GFS 0.25 degree model data was replaced with the 1.0 degree because it was overwhelming our server

- All Satellite, MRMS, and Model menus were moved from being configured in CAVE to EDEX which gives us the ability to quickly update those menus for all users
- Impact based warning tags were added
- Import GIS functionality was added back into CAVE
- The screen capture functionality on the Mac version of CAVE was fixed
- Our downloads for both the Mac and Windows applications are officially signed (and officially notarized for Mac as well)

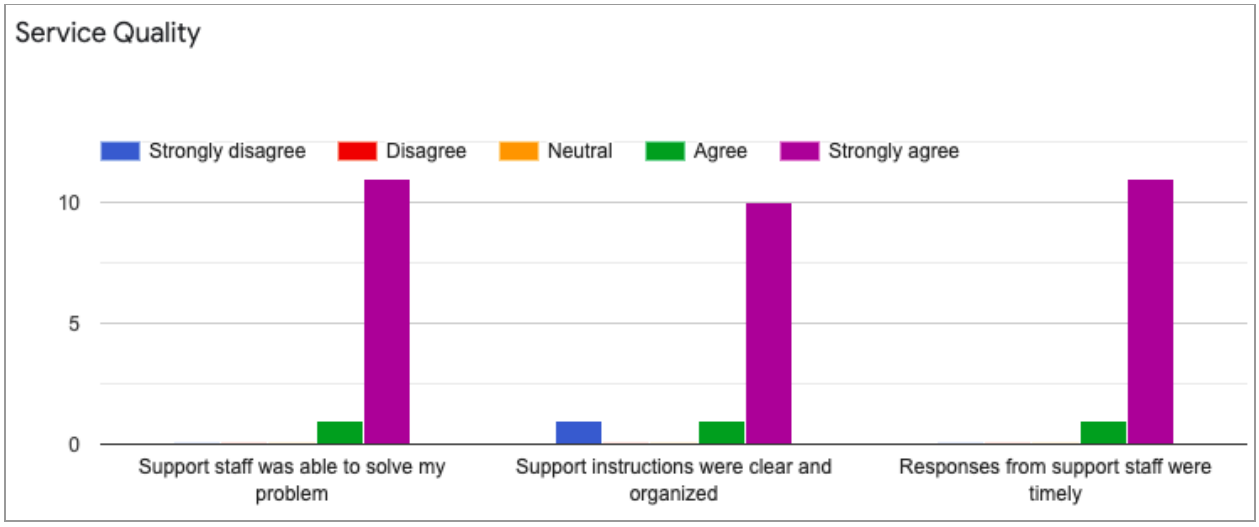
In addition to this major release, we had also put out a minor release specifically for the Mac version of CAVE to fix a bug and change the default functionality of the initial connection window. We are also currently working on another minor release for the Linux version of AWIPS, which consists almost exclusively of updates for the EDEX side of the software. These new EDEX upgrades include: modification of the modelsounding plugin to support changes made to the file format back in May of 2021, a units fix for absolute vorticity products, modification to EDEX authentication to do with usernames containing spaces, correcting a radar station name, and adding additional GOES fire products for both GOES East and West. With the knowledge of Dr. Bruning's gridded GLM processing nearing its end of life, we have worked with the National Weather Service to obtain a feed of the gridded GLM data for both GOES East and West along with the 5-minute summation files. Additionally we are working processing all of the Himawari satellite products. Unidata is looking to redistribute the gridded GLM data via the IDD and as mentioned above, we're seeking input from the community if Himawari data would be useful if added to the IDD.

A significant portion of our documentation both for CAVE and EDEX, and for 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. With the help of Nicole Corbin (our Educational Designer), we created and released an updated, detailed video for the installation of CAVE (and all its necessary prerequisite software) on Windows machines. In May of 2021, Microsoft intentionally removed packages we had previously been instructing users to download, so we had to figure out how to use the remaining options to complete a successful CAVE install on Windows.

Our blog series, AWIPS Tips, has successfully been running every other week for 13 consecutive posts, spanning six months. So far, we've had eight posts related to CAVE, two posts related to python-awips, and one post each for EDEX, general announcements, and release announcements. We plan to continue the blog series for the foreseeable future and have over a dozen 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. This work has also been made possible with contributions from Nicole Corbin.

On a related note, we've implemented a new EDEX server which is specifically for the CAVE training that is being developed by Nicole Corbin (and is detailed in the Community Services Report). The training will be available online and is an asynchronous course that can be used by professors for introducing their students to CAVE. The new EDEX server currently has radar data from a severe weather event from May 31, 2013. The new training-edex server can be a starting point for us to investigate what it takes to host a full on case study server.

The AWIPS team still has an active support evaluation survey that is advertised in our support email signatures. Since the last status report we have had four new evaluation entries, which is slightly under our average of one per month. 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 open-ended feedback from the support evaluations includes the following:
"Unidata support has always been phenomenal, great to see newer staff keeping that work ethic!"
"Support team was very helpful! Responded quickly."
"The excellent support is appreciated!"

One last piece of work related to EDEX servers: the AWIPS team has spent a significant amount of time working with the Machine Learning (ML) division at Unidata to help integrate a ML pipeline and dataset into AWIPS. We hope the outcome of these efforts will be a docker container that is easy to install and will allow our community to test out different ML algorithms and easily see them in CAVE. This is still currently work in progress.

Conferences for both AGU and AMS are coming up quickly. We have applied to present at AMS 2022 about Unidata's AWIPS, its current status and a look toward the future. Additionally, the Unidata's Machine Learning expert, Rio McMahon, has submitted abstract

entries to both AMS and AGU regarding his EDEX docker work, that both Unidata AWIPS team members (Tiffany Meyer and Shay Carter) are authors of.

Software Releases

SInce the last status report we put out our first official AWIPS release for all three operating systems (Linux, Windows, and MacOS) -- version 18.2.1-1. We also made a very minor release shortly after for MacOS -- version 18.2.1-2. We also were successful in officially notarizing and signing all our software packages for the Mac, and will continue to do so with future releases.

We have made a number of minor changes to the Linux version of AWIPS (primarily on the EDEX side of things) and plan to put out a minor release for version 18.2.1-2 by the end of the month (October 2021).
In the more extended future, we aim to merge our Unidata AWIPS back with one of Raytheon's more recent builds (most likely version 20.3) which includes transitioning from Python 2 to 3 and Java 8 to openJDK 11.

Activities Ongoing/In-Progress

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

- The AWIPS team has been testing and implementing changes to improve and optimize our cloud EDEX server.
- The AWIPS team is exploring the possibility of adding additional data and increasing the archive time of some existing data.
- 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 has worked through the build process for the Linux, Windows, and MacOS distributions of CAVE and is ready to make deployments as soon as the code is ready.
- The AWIPS team has started a new bi-weekly blog series called AWIPS Tips that began on April 7th, and is intended to highlight useful functionality and fundamentals for CAVE, EDEX and python-awips.
- The AWIPS team has been working closely with our Educational Designer to develop an asynchronous training set which is focused on CAVE, that can be used by professors and their students.
- The AWIPS team is getting close to putting out an additional minor release that includes many EDEX improvements.

Future Activities

Future plans are constantly evolving to meet the needs of our users. The AWIPS team is in the process of adding new data that has been requested from our community. We are currently waiting on new GOES WEST derived products that are set to be distributed later this month (September 30, 2021), after that, the goal is to package and release a minor version of AWIPS for Linux -- version 18.2.1-2. After that, our asynchronous CAVE training will be live and we plan to contact our community about trying it out in their classrooms and labs. In the more distant future, we plan on beginning the transition of Unidata's AWIPS distribution away from version 18, onto version 20.

Metrics

Downloads April 1, 2021 - September 30, 2021

AWIPS downloads: 3,777

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, and encouraging community members to assist each other through Unidata-managed forums for GEMPAK support.

Prepared *October 2021*

Status Report: Science Gateway and Cloud Computing Activities

April 2021- October 2021

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Has the COVID-19 pandemic and the shift to online learning increased your reliance on cloud computing technologies? How can Unidata better assist you in this area? What is missing in your toolkit to teach effectively during the pandemic?
2. Do you need a Unidata hosted JupyterHub for your classroom or workshop?
3. What new cloud technologies are our community members using and investigating on their own initiatives?
4. What cloud computing environments or platforms are our community members using? Commercial? E.g., Amazon. NSF? E.g., Jetstream.

Activities Since the Last Status Report

NSF Jetstream Cloud Grant Renewal for 2021-2022

Our NSF Jetstream Cloud allocation must be renewed annually with NSF XSEDE (Extreme Science and Engineering Discovery Environment) to ensure continuity of Unidata services (EDEX, THREDDS, JupyterHub servers, etc.) running on Jetstream. Our submission was successfully awarded for 2021-2022:

- ~4,000,000 Service Units (SUs) to run virtual machines on Jetstream.
- 40TB of disk storage.
- XSEDE, Extended Collaborative Support Service (ECSS) to continue JupyterHub cluster support with Andrea Zonca (San Diego Supercomputing Center).
- SGCI (Science Gateway Community Institute) Support.

JupyterHub Servers for Online Instruction During COVID-19 Crisis and 2021 Summer/Fall Semester

Unidata JupyterHub activities continue to advance since the last status report. These JupyterHubs are deployed In collaboration with XSEDE, ECSS and the Jetstream group at Indiana University (IU).

We have supported a number of semester-long classes, and workshops with JupyterHub

servers hosted on the Unidata Science Gateway. The JupyterHub servers are tailored to the instructor's objectives with pre-configured PyAOS (Python for the Atmospheric and Oceanic Sciences) environments, classroom material and data. Demand for Unidata JupyterHub servers has increased since the arrival of the COVID-19 pandemic and the transition to online learning. We are more than happy to assist instructors in this area, and would like to help in whatever way we can with these resources.

For the summer and fall semester of 2021, Unidata assisted (or will help) four universities, one workshop, and three groups of SOARS students with pre-configured PyAOS JupyterHub servers. We have helped about 108 students over this time period (see table below in the metrics section).

Helping SOARS Protégés Program for Summer of 2021

For the summer of 2021, we assisted the SOARS program with three separate JupyterHubs: one instructional JupyterHub for the summer cohort. Two customized JupyterHubs for summer internship projects. For the latter two, we worked closely with the students and their mentors to ensure they had all the tools and data they required to complete their projects.

University of Oklahoma with Ben Schenkel

Unidata collaborated with Ben 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.

Disk Volume Problems Associated with JupyterHub Servers

Kubernetes JupyterHub servers operating on Jetstream occasionally experience disk volume problems where a student cannot login to their classroom JupyterHub server. We have some **good news** to report in this area that has been vexing us for a long time (see previous status reports). A GitHub user indicated a much faster workaround than the one we have been employing in the past. The underlying problem still persists, but coping with it is not as tedious as previously described.

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 Jetstream 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. This server needs to be regularly updated as the Jupyter, JupyterHub, and JupyterLab ecosystems rapidly evolve.

Docker Containerization of Unidata Technology

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.

Progress has been made on the following

- Tomcat Docker container continues to be updated as new versions of Apache Tomcat become available. We try to do this quickly in the event of a Tomcat security update.
- We work closely with Jen Oxelson and the systems administration group to ensure that our containers are configured as securely as possible.
- Collaborated with Sean Arms for the deployment of TDS/TDM production and beta releases as new versions of the TDS/TDM are released.
- DockerHub has discontinued autobuilds to maintain containers up-to-date with respect to Dockerfile changes and upstream builds. As a result we are using Jetstream to keep these images up-to-date by employing that platform for building the docker images and pushing them to DockerHub. (This development is in progress.)
- We are automating the update of running containers to ensure those remain up-to-date with respect to security enhancements. (This development is in progress.)

Product Generation for IDD

For the past five years, Unidata generated products for the IDD, FNEXRAD and UNIWISC data streams have been created by a VM hosted in the Amazon cloud. This product generation has been proceeding smoothly with almost no intervention from Unidata staff.

AWIPS EDEX in Jetstream Cloud

Unidata continues to provide an EDEX data server on the Jetstream cloud, serving real-time AWIPS data to CAVE clients and through 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 Jetstream to develop

cloud-deployable AWIPS instances, both as imaged virtual machines (VMI) available to users of Atmosphere and OpenStack, and as Docker containers available on DockerHub and deployable with the science gateway toolset.

EDEX is designed with a distributed architecture, so different components can be run across separate virtual machines (VMs) if needed, to improve efficiency. Our current design makes use of two VMs: one large instance to process most of the data and run all of the EDEX services, and another smaller instance to handle radar data specifically.

For the past year, we have successfully created and maintained a duplicate set of VMs to mirror our production EDEX environment. These backup VMs have served as a testing ground for implementing new changes, as well as a backup for when our production server is unavailable.

This coming year, we will add an additional VM to each of our EDEX clusters, dedicated to processing the ever growing GOES satellite products to reduce processing latencies. Additionally, we plan on potentially adding one more VM to our EDEX environment to help handle the large influx of CAVE connections we periodically encounter (for a total of two new VMs -- one for each of the production and backup systems). This scenario can happen in severe weather conditions, or when numerous students in a university class connect to our server simultaneously.

Our community has expressed interest in a case study server. With the upcoming release of our CAVE E-learning course, we've added a small VM that serves up an archived radar case for a severe weather event from May 31, 2013 to go with the training. The new training-edex server can be a starting point for us to investigate what it takes to host a full on case study server.

Nexrad AWS THREDDS Server on Jetstream Cloud

As part of the NOAA Big Data Project, Unidata maintains a THREDDS data server on the Jetstream 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.

Jetstream Security

We work with the Unidata system administrator staff to ensure that our web-facing technologies and virtual machines on Jetstream 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.

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.

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

- M. K. Ramamurthy and J. Chastang. The use of the unidata science gateway as a resource for facilitating education and research during covid-19. In *17th IEEE eScience 2021*, Innsbruck, Austria, Sept. 20-23 2021.

New Activities

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

Forthcoming Virtual Conference Attendance

- American Geophysical Union Fall Meeting 2021
- American Meteorological Society 2022
    - Abstract submitted to AMS DEI conference: “Unidata Partners With UCAR SOARS Program to Help Protégés and Their Mentors With Atmospheric Science Internships”

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

XSEDE ECSS JupyterHub Collaboration

We plan to continue our collaboration with Andrea Zonca (XSEDE ECSS, San Diego Supercomputing Center) for deploying JupyterHub clusters on Jetstream. We continue to provide Andrea with feedback as he releases new versions of the software. As the ECSS project appears to be winding down, Andrea is looking for a new source of funding to continue this vital collaboration.

Jetstream2

As described in the fall 2020 Cloud Computing Activities status report, the NSF Jetstream2 project, a follow-on to the current Jetstream facility, has been awarded funding. Jetstream2 hardware is being assembled and testing is underway. It should be available for use later this

year.

NSF Supplemental Funding for Unidata Science Gateway

We are extremely fortunate to have received NSF supplemental funding with the aim of adding a new team member to the Unidata Science Gateway program. We are in the process of assembling the search committee and composing a position description which we can advertise soon.

Relevant Metrics

Summer and Fall 2021 JupyterHub Servers

	Nº of Users	**Point of Contact**	**Notes**
Summer 2021			
UCAR SOARS Internship	20	Keith Maull, UCAR/UCP	
SOARS B	3	Antara Banerjee - NOAA Affiliate	One student, two instructors
SOARS C	1	Keith Maull, UCAR/UCP	Allysa Dallmann, SOARS student
Fall 2021			
OU	20	Shawn Riley, Ben Shenkel OU School Meteorology	JupyterHub started summer 2021
University of Louisville	6	Professor Jason Naylor	
U of North Dakota	3	Dr. Aaron Kennedy Assoc Prof Dept of Atmos Sciences U of North Dakota	
U of North Dakota 2	15	Dr. David Delene Prof Dept of Atmos Sciences U of North Dakota	
Southern Arkansas University	34	Keith Maull (UCAR/NCAR Library)	
Fall 2021 Python Workshop	6	Drew and Nicole	

Github Statistics

	Watches	Stars	Forks	Open Issues	Closed Issues	Open PRs	Closed PRs
sience-gateway	4	15	8	9	152	0	466
tomcat-docker	7	48	48	2	35	0	65
thredds-docker	9	23	21	3	105	0	147
ramadda-docker	2	0	1	1	10	0	21
ldm-docker	6	11	11	1	33	0	56
tdm-docker	3	3	6	1	9	0	16

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 Jetstream 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 Jetstream for the benefit of the Unidata community. Unidata also makes the AWS Nexrad archive data accessible through the TDS Nexrad server running on Jetstream at no cost to the community. These data can be accessed in a data-proximate manner with a JupyterHub running on Jetstream 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 *October 2021*

Status Report: Community Services

April 2021- October 2021

Nicole Corbin, Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We are requesting your feedback on the following topics:

Do your needs from the Unidata Program Center change during this unique 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:

- New: Downloadable Resources from COMET MetEd University Course Support
- Toward Standardized Digital Representations of Units of Measurement
- Meet Unidata's 2021 Summer Interns
- Unidata Program Center Welcomes Rio McMahon
- 2021 Community Equipment Awards
- MCFETCH Provides Free Satellite Archive Data Access for the Unidata Academic Community
- What's New in IDV 6.0
- Summer 2021 Unidata Interns Wrap Up Their Projects
- Offer: Unidata Science Gateway JupyterHub Resources Available for Fall 2021 Courses
- Unidata Welcomes New Committee Members
- THREDDS Data Server version 5.0 Released
- Community Survey: Tell Unidata What You Think!
- Software release information
- Many AWIPSTips 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:

- Continue to engage with underserved populations and institutions as part of Unidata's outreach efforts to groups such as Rising Voices, SACNAS and AISES
- Continue to serve on the CUAHSI HIS and DEI standing 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
- We continue to actively support the NCAR/SOARS program.
- Actively participate in Super Science Saturday.
- 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.

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.

Learning and Outreach

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

Progress has been made on the following:

- Delivering a second offering of the Introduction to MetPy workshop (open enrollment)
- Offering a prerequisite web course to prepare learners for the workshop, includes machine setup, a self-assessment of prerequisite knowledge, and practice lessons in matplotlib, cartopy, and siphon/xarray.
    - Introduction to MetPy Prerequisite Course
- Creating an introductory asynchronous web course titled Learn AWIPS CAVE, scheduled for release mid-October.
    - Created with undergraduate meteorology students in mind, but applicable to meteorologists of any education level
    - Separated into modules for spaced learning
    - Instructors/supervisors can request quiz results
    - Learn AWIPS CAVE Learning Objectives
- Continuing to share tips and best practices for using AWIPS through the AWIPS Tips blog series

Dependencies, challenges, problems, and risks include:

- With the increase in scalable/reusable training materials offered for MetPy and AWIPS, we are interested in increasing participation and learning more about potential barriers to use. Of particular interest is participation in virtual workshops and using ready-to-use materials in the classroom/lab.

Ongoing Activities

We plan to continue the following activities:

- Engagement with science or cyber communities at large
- NAWIPS migration to AWIPS, including the overall AWIPS project
- Ongoing development of news articles for publication through News@Unidata
- Continue to support and contribute to governing committees
- Seminars
- Outreach
- Inclusion and equity
- Engagement with professional societies
- Support for cloud-related projects
- Further development of the Data Management Resource Center
- Support the pursuit of funding
- Site visits as the budget and pandemic allow
- Engage other UCAR/NCAR divisions regarding Unidata software use i.e. CESM/IDV
- Active participation in CUAHSI HIS (Hydrologic Information System) and DEI
- Active participation in the Hydroshare Advisory Committee

New Activities

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

- Expanded emphasis on engagement with MSIs
- Expanded effort organizing and supporting community seminars/working sessions
- Provide additional support for instructors/supervisors using or planning to use Learn AWIPS CAVE, especially in the form of ready to use lab worksheets and grading rubrics.

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

- Make structural changes to broaden participation in Unidata community engagement
- Continue to engage the hydrologic community regarding WRF-Hydro/IDV interactions and the National Water Center's efforts

- Continue to engage the arctic community to find opportunities for collaboration
- Seek additional opportunities to engage and listen to the community
- Research options for creating a learning management system for hosting and reporting metrics for learning offerings.

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

- Support the providing additional cloud-related training

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:

- 42,201 unique pageviews (46,703 in previous period)
- 7.8% of total unique pageviews (8.1% in previous period)

Top community pages

1. All blog pages
   33815 unique pageviews (32397 in previous period)
   80% of total community pageviews (69% in previous period)
2. www.unidata.ucar.edu/community
   3728 unique pageviews (8592 in previous period)
   9% of total community pageviews (18% in previous period)
3. www.unidata.ucar.edu/about
   2546 unique pageviews (3101 in previous period)
   6% of total community pageviews (7% in previous period)
4. www.unidata.ucar.edu/events
   1600 unique pageviews (1927 in previous period)
   4% of total community pageviews (4% in previous period)

Social media statistics, October 5, 2021

1. # of Twitter followers: 1672 (up from 1597)
2. # of Facebook followers: 879 (up from 874)
3. # of YouTube subscribers: 2126 (up from 1760)

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.

Prepared *April 2021*

Status Report: Unidata Community Equipment Awards

Sponsored by the National Science Foundation

April 2021- October 2021

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Possible theme(s) for the 2022 Unidata Community Equipment Awards;
2. Please consider volunteering to serve on the 2022 Review Panel;
3. Suggestions from previous panel members on how to improve the program

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.

A Request for Proposals was sent out on December 1, 2020 with a March 26, 2021 submission deadline. The Review Panel will meet virtually on April 14 to finalize their reviews and to make funding recommendations.

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 *October 2021*

Status Report: GOES-R/S, NOAAPort and Other Satellite Imagery

April 2021- October 2021

Mike Schmidt, Tom Yoksas

Questions for Committee Members

- What image coverages, spatial and temporal resolutions and possibly projections should be considered for addition to the **UNIWISC** IDD feed?

- What kind(s) of data access methods are most desired/usable for the community?

  We currently provide access via the IDD (push), THREDDS Data Server (pull), McIDAS ADDE (pull) and AWIPS EDEX (pull).

- Other questions?

Activities Since the Last Status Report

- We are actively working towards establishing additional GOES-R downlinks in locations that have good views of the southern skies

  Comment(s):

  An effort to establish a new GOES-R 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 has been done for the entire site. Exactly when an environmental impact assessment will be done is currently unknown.

  By any measure, GOES-16/17 imagery continues to be **very** popular in the community!

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-17 GRB products.

  We are feeding from SSEC's GOES-16/17 fanout servers, and they are feeding from the ingest machine that we operate. Sharing of the feed streams has allowed by SSEC and Unidata to minimize effects of solar and terrestrial interference.

- Ingest GOES ReBroadcast (GRB) streams from GOES-16 and GOES-17 in real-time

  Since repointing the 4.5m satellite dish at the NCAR Mesa Lab from GOES-16 (75 W to GOES-17 (137.2 W), the Terrestrial Interference (TI) that had been hampering GOES-16 data ingest activities has decreased to the point that our ingest quality rivals what UW/SSEC experiences on their 6.3m dishes.

  Background:

  In the fall of 2017 we began experiencing significant TI in the GOES-16 signal being received by our 4.5m satellite dish at the NCAR Mesa Lab. An outcome of the collaborations we had with Quorum Communications (the manufacturer of the electronics we use in our GOES-R/S ingest installations) was our moving of the GOES-16 ingest to a 3.8 m satellite dish located at the UCAR FL-2 location. The relocation of GOES-16 ingest required that an additional signal cable be pulled from the satellite dish that was repurposed from GOES GVAR ingest into the 2nd floor FL-2 NCAR/RAL computer room where our ingest electronics are located. The cost of this work was contributed by the UCAR/NCAR networking group.

  While the actual source of the TI being experienced at the Mesa Lab could not be pinpointed, the best guess at the time was that the noise being experienced was coming from power lines that lie south and downhill from the Mesa Lab. All of those power lines have since been replaced (by Xcel Energy), so there is a possibility that the TI is now gone. We will be conducting an experiment of re-pointing the dish from GOES-West back to GOES-East to see if the power lines were, in fact, the source of the TI. If the results of the test are positive, meaning the TI is gone, we will leave the dish pointing at GOES-East and resurrect the original plan of installing a GOES-West downlink on the western satellite pad at the Mesa Lab. This experiment will be conducted in conjunction with UWisc/SSEC to minimize potential data loss to the Unidata community.

  In the spring, we were given a 3.7m satellite dish that was being excessed by a private company that was relocating their operations. We will use this dish to establish a GOES-R downlink at the NCAR Marshall field site. Even more recently, we were given an additional 3.8m satellite dish that was being replaced by another company in Centennial, CO (on the southern side of Denver). Pending the testing described above, this dish may be installed on the western satellite pad at the Mesa Lab and used for GOES-West data ingestion.

- Continue to distribute GOES-16 and GOES-17 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

- Work with NOAA/GSL to get a feed of their **RRFS** (Rapid Refresh Forecast System) when it has matured enough for the products to be distributed to end users for evaluation.

Future Activities

CSPP GEO Gridded Geostationary Lightning Mapper (Gridded GLM)

On March 21, 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 have recently 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 currently being created by Eric Bruning of Texas Tech University will be implemented in the coming weeks.

Himawari Imagery and Level 2 Products

We have also recently 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.

NOAAPort SBN

In the spring, the NWS informed COMET that they would no longer be supporting the NOAAPort downlink facility that is installed near the front entrance of Foothills Lab 2. It is still unclear if the NWS will want to remove the dish or to transfer ownership to UCAR. It is interesting, however, that a tuned cavity filter was installed on the dish within the past week (reportedly the government is funding the installation of filters at all NOAAPort and GRB downlink sites in order to mitigate potential/expected interference from new cell phone services.

We are trying to keep abreast of a possible expansion of the NOAAPort SBN that would, if implemented, increase available bandwidth twofold. As one might imagine, progress on this and other fronts has been slowed by reorganization of priorities during the COVID-19 epidemic.

IDD NIMAGE and UNIWISC Datastreams

As noted earlier, both the **NIMAGE** and **UNIWISC** datastreams have been revamped to include GOES-16/17 imagery and products. We will add more products if asked to do so by the governing committees. More recently, the **FNEXRAD** datastream was enhance by the addition of MRMS products we receive in an LDM feed from NOAA/NCEP. The volume of data available in the **NIMAGE**, **UNIWISC,** and **FNEXRAD** datastreams can be seen in:

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

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

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

VALUE-ADDED Products

Texas Tech University (Eric Bruning) has been creating value-added Level 2 products created from Geostationary Lightning Mapper (GLM) images as a precursor for similar products potentially being added to NOAAPort. We have been distributing these Level 2 products in the **NIMAGE** IDD datastream. Creation of these GLM Level 2 products is scheduled stop when the Unidata equipment grant that funded the effort expires.

In the spring, the **NIMAGE** feed was enhanced by the addition of three RGB products being created at CSU/CIRA. The repurposing of **NIMAGE** from a feed that only contained satellite imagery that derived from NOAAPort to one in which new, innovative Level 2 satellite products has seemingly been accepted by the community.

We welcome contributions of additional value-added products by TTU and other sites.

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-16/17 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!

  The volume of GOES-16 and GOES-17 GRB products (13 GB/hour average and 20 GB/hour maximum; that this is the second most voluminous IDD feed!) can be seen in the real-time statistics plot from our GOES-R/S ingest machines:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?SATELLITE+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-16 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 *October, 2021*

Status Report: Internet Data Distribution

April 2021- October 2021

Steve Emmerson, Mustapha Iles, Mike Schmidt, Jeff Weber, Tom Yoksas

Questions for Committee Members

- Do you have suggestions regarding content of data streams like CONDUIT, FNEXRAD, NIMAGE, UNIWISC, NLDN Lightning, SATELLITE, etc.?

  We (UPC, the Unidata community and UAlbany for the NLDN component of LIGHTNING) have control of the content of these data streams, so their contents are open for suggestions.

  NB: We now have access to Himiwari imagery and Level 2 products on AWS S3, so we could add some of these products to the IDD (probably in the SATELLITE feed), if the committee thinks that this would be useful AND manageable by sites.

Activities Since the Last Status Report

Internet Data Distribution (IDD)

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

```
20211007

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume 119358.700 M bytes/hour
Average hourly volume  74526.521 M bytes/hour

Average products per hour     509799 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)   number/hour
CONDUIT                 15205.338    [ 20.403%]    51146.268    89535.800
SATELLITE               14869.071    [ 19.951%]    20346.543     6565.956
NEXRAD2                 10720.251    [ 14.384%]    19668.797   113109.089
NIMAGE                   9378.150    [ 12.584%]    13941.999     7690.933
NGRID                    7312.037    [  9.811%]    12105.337    67971.733
HDS                      4800.405    [  6.441%]    12270.567    30426.267
FNEXRAD                  4549.867    [  6.105%]     5041.324     9852.467
NEXRAD3                  3058.478    [  4.104%]     7014.444   115908.533
```

```
FNMOC                  1752.945    [  2.352%]     6124.308     6190.400
UNIWISC                 938.094    [  1.259%]     2028.621     1872.978
GEM                     715.343    [  0.960%]     4250.579     4473.800
FSL2                    701.342    [  0.941%]     1811.204     1406.311
NOTHER                  367.817    [  0.494%]     5569.146      560.689
IDS|DDPLUS              142.175    [  0.191%]     1820.098    53815.733
LIGHTNING                15.161    [  0.020%]      440.881      417.756
GPS                       0.047    [  0.000%]        1.112        0.422
```

New Data Distribution:

IDD CONDUIT feed:

The GFS model was upgraded from v15.3 to 16.0 last spring. The increase in data volume and number of products in CONDUIT was significant enough to warrant Unidata upgrading its ingest machine to be able to handle the increased volume. The same action was not taken by NCEP on the virtual machines that they use to create the CONDUIT datastream. Over the past several months, delivery of the full set of CONDUIT products in a timely manner has been affected to the point that full sets of products are not being sent, and long latencies are being experienced for a large fraction of the products. The last half day of the recently held virtual LDM training workshop that we hosted for NCEP was spent troubleshooting this situation, and it was determined that the sizes of the LDM queues on the two CONDUIT source machines was too small to hold data for a sufficient length of time, in the worst case that we saw, the minimum residency time of products in one of the LDM queues was 93 seconds. When NCEP will be able to upgrade the RAM in these virtual machines, or redo the virtual machines completely is currently unknown.

IDD FNEXRAD, NIMAGE and UNIWISC feeds:

We continue to create the content for the FNEXRAD (NEXRAD Level III national composites), NIMAGE (GOES-16 and -17 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-16 and -17 images converted to McIDAS AREA format for ease of use in legacy systems like GEMPAK).

MRMS imagery/products we are receiving from an NCEP LDM feed and that is in GRIB2 format was added to the IDD FNEXRAD feed after testing of the products was completed. These products are being generally distributed in the IDD.

IDD NIMAGE feed:

The NIMAGE feed was enhanced by the addition of three products being created by CSU/CIRA: GeoColor, DebraDust and CloudSnow. GOES-East CONUS and -West PACUS coverages are available for the GeoColor and CloudSnow products while the DebraDust product is available

in a GOES-East CONUS coverage. All three of these are RGB products - displays are created using different wavelength channels to drive the Red, Green and Blue portions of a composite display. The GeoColor product is quite interesting and useful, so users are encouraged to take a look!

Support for displaying these RGB images has been added to Unidata AWIPS, and finishing touches on adding support to Unidata McIDAS-X were just checked into the code repository at SSEC.

Experimental HRRR feed:

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 couple/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. During the summer, a week long power down at the NWSC forced us to create new instance of idd.unidata.ucar.edu at the NCAR Mesa Lab. Transition from the NWSC instance of idd.unidata.ucar.edu to the Mesa Lab instance was done seamlessly by a DNS change. Very few (2 or 3) support inquiries were received as a result of the change. After the work was complete at NWSC, idd.unidata.ucar.edu was returned to the cluster there, and transition was once again handled by a DNS change. We received no support inquiries from the switch back to the NWSC instance of idd.unidata.ucar.edu.

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 (GOES-17 at the NCAR Mesa Lab and GOES-16 at UCAR Foothills Lab 2). The data volume seen in the **NIMAGE** entry represents GOES-16/17 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-16/17 image products that are in PNG compressed McIDAS AREA format that is suitable for use in GEMPAK, the IDV, McIDAS-V, and McIDAS-X.

Challenges, problems, and risks:

More sites 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 took over the data distribution of GPS radio occultation solutions from COSMIC. COSMIC will still gather incoming GPS data and create the solutions, but due to hardware constraints COSMIC has requested Unidata to provide distribution from our top level IDD relay clusters (idd.unidata.ucar.edu and iddb.unidata.ucar.edu) to the community. The solutions (Precipitable Water Vapor and Total Electron Content-Ionosphere) are in netCDF format and are available in the GPS feedtype.

- 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.

  Unidata has been assisting LSU/Climate (formerly LSU/SRCC) with the maintenance of their NOAAPort ingest capability for several years. Activities have included providing a spare LNB to bring their NOVRA S300N receiver to Boulder for testing, configuration, power supply replacement and routine monitoring of their data and distribution. Recently, the signal (as recorded by a Novra S300N in the LSU/Climate data center) has experienced severe problems. The LSU/Climate folks have been trying to identify a competent satellite tech on the LSU campus to help with troubleshooting the problem (which is most likely corrosion in connectors in the signal path).

During this and the end of the previous reporting periods, considerable effort has been expended to streamline our NOAAPort ingest systems and assist other sites (UW/SSEC, NOAA/GSL, NOAA/SPC) in troubleshooting problems being experienced in their systems. Most recently, the NOAAPort ingest capability in the LDM has been enhanced to allow the user to specify the size of the ingest buffers used. This capability, while requested by NOAA/Raytheon has increased our understanding of the need for additional input stream buffering, and the idea has been transferred to our GRB ingest efforts where significant improvements have been experienced.

- The NOAAPort-derived data streams (**HDS**, **IDS|DDPLUS**, **NGRID**, **NIMAGE**, **NEXRAD3** and **NOTHER**) are redundantly injected into the IDD at four geographically separate locations: UCAR/Unidata, UW/SSEC, LSU/Climate, and Allisonhouse.com. The **NOTHER** data stream contains GOES-16 and GOES-17 tiles that need to be stitched together to make full image scenes usable to end-user applications. Unidata is using Ryan May's ldm-alchemy package (available in the Unidata section of Github) to create full ABI L2 images that are then relayed in the **NIMAGE** datastream which was revamped earlier in the year since the content of the NOAAPort-received **NIMAGE** products dropped to essentially zero when GOES-15 was put into standby storage. Even though GOES-15 will periodically be taken out of storage for use in western Pacific hurricane monitoring, the GINI image products that were once created from its scans will no longer be produced by NOAA.

- We are now looking for a site that could augment/replace LSU/Climate as a fourth contributor to NOAAPort-derived content in the IDD. Discussions with FoxTV 13 in Tampa, FL have recently begun, but the prospect of them agreeing to feed us via a quid pro quo arrangement is in doubt.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.13.16. A new LDM distribution is being readied for release very soon.

Relevant Metrics

- Approximately **577** machines at **209** 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 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.

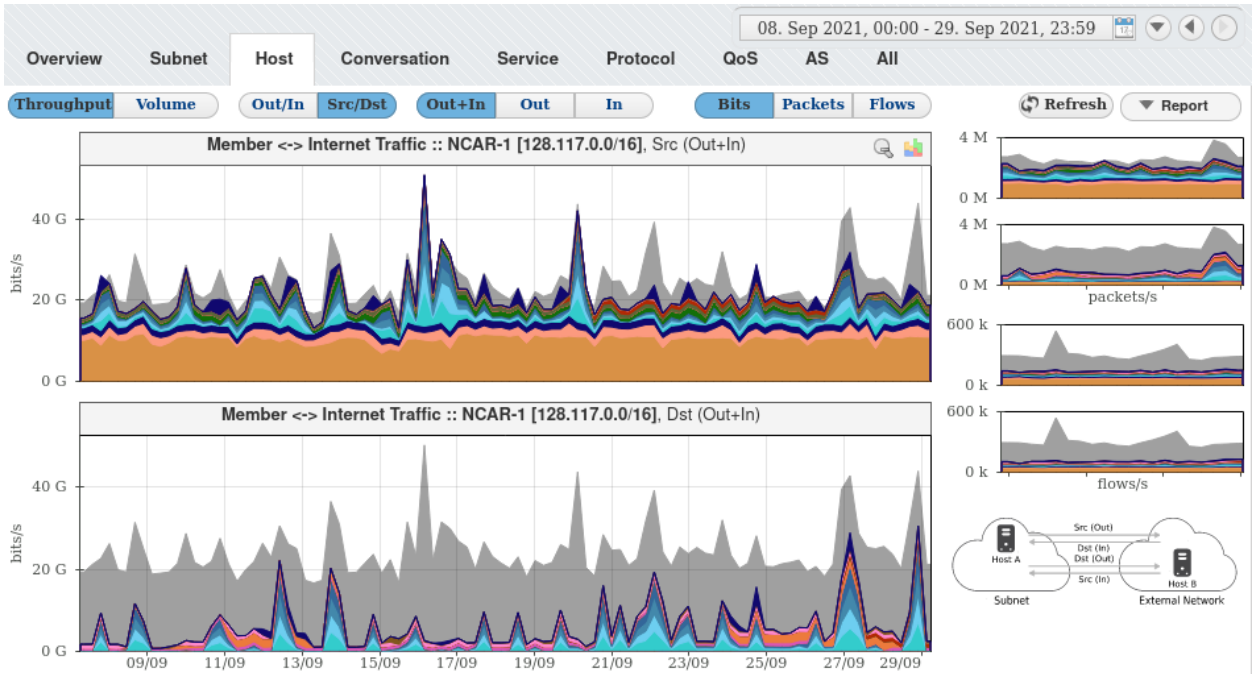
  We know that there are a number of sites that are participating in the IDD, but are not reporting real-time statistics back to us. Reporting of real-time statistics is not and never has been mandatory.

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 relays, **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 1205 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.



Hosts	Src (Out+In) Avg	Src (Out+In) Max	Dst (Out+In) Avg	Dst (Out+In) Max	Total (Out+In) Avg	Total (Out+In) Max
idd.unidata.ucar.edu	5.7 Gbps (33.	11.7 Gbps	41.4 Mbps (0.	96.8 Mbps	5.7 Gbps (17.	11.8 Gbps
rda.ucar.edu	1.6 Gbps (9.3	3.6 Gbps	12.0 Mbps (<1	22.1 Mbps	1.6 Gbps (4.7	3.6 Gbps
iddb.unidata.ucar.edu	1.2 Gbps (7.2	1.9 Gbps	9.3 Mbps (<0.	14.8 Mbps	1.2 Gbps (3.6	1.9 Gbps
gridftp03.ucar.edu	530.8 Mbps (3	9.0 Gbps	339.3 Mbps (3	6.7 Gbps	870.1 Mbps (3	9.0 Gbps
gridftp04.ucar.edu	520.1 Mbps (3	7.8 Gbps	342.9 Mbps (3	6.7 Gbps	863.0 Mbps (3	7.8 Gbps
gridftp02.ucar.edu	517.2 Mbps (3	9.9 Gbps	338.0 Mbps (3	6.8 Gbps	855.2 Mbps (3	9.9 Gbps
gridftp01.ucar.edu	500.6 Mbps (3	8.0 Gbps	340.1 Mbps (3	6.8 Gbps	840.7 Mbps (3	8.0 Gbps
tds.ucar.edu	766.6 Mbps (4	2.8 Gbps	7.0 Mbps (<0.	16.4 Mbps	773.6 Mbps (3	2.8 Gbps

Over the period from September 8 through September 29, 2021 (there was a long

period previous to September 8 where network traffic information is unavailable) the average volume of LDM/IDD data flowing from the UCAR/NCAR network averaged around 7.6 Gbps (~82.1 TB/day), and peak rates reached 15 Gbps (which would be ~162 TB/day if the rate was sustained (which it is definitely **not**)).

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

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

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 *October, 2021*

Status Report: IDV with RAMADDA

April 2021- October 2021

Yuan Ho, Julien Chastang

Areas for Committee Feedback

We have no questions at this time.

Activities Since the Last Status Report

IDV Release

IDV 6.0u1 was released on July 29 of 2021.

IDV System Changes

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the stable release version (6.0u1) is the 5.4.2-SNAPSHOT . There have been many new features and bug fixes in that range. The complete release notes for these versions can be found here.

__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

__Dark Mode Appearance__

In macOS and Linux users can choose to adopt a light, dark, or IDV regular appearance. The dark appearance, known as Dark Mode, implements an interface style that many apps already adopt. Users can edit the preference and can choose their interface appearance based on ambient lighting conditions.

__New Volume Rendering__

The volume rendering capability native to IDV produces a semitransparent volume representation. The new volume rendering methodologies in IDV implement a on-the-fly resample to a Linear3DSet. By using the Stack 2D algorithm, any field can be transformed into a binary representation. This feature is still under development and testing.

__Cross Seam Subsetting__

The cross seam subsetting in the grid coverage type works similar to the grid file subsetting where you can set the property for all fields in the Data Source or override the default for a particular field in the field selector subset panel. The only difference in the Coverage data spatial subset is that it will allow users to flip longitudes in a cyclic rect linear grid from 0/360 to -180/180 (or vice-versa) before performing the cross seam subsetting. This feature is still under development and testing.

__Text Html Display__

New enhanced Text/Html display available under Displays > Special for displaying straight text or HTML, this display can also be added to the view window legend panel for descriptive text about the display.

__Grid Diagnostics Formulas__

The new GFS dataset collection has several parameters with multiple accumulation hours in the time axis for each time step and the previous logic in the IDV picks the first one in the list, unfortunately, always the longest one. So the new logic will try to pick the shortest accumulation hour for each time step, and what we got are 3 and 6 hours for

GFS one degree or 0.25 degree, we can use the subtraction to get the 3-hour time step accumulation. This change has been added as a derived variable named: TimeStepAccumulatedPrecip. In general, this formula can be applied to any mixed interval parameters.

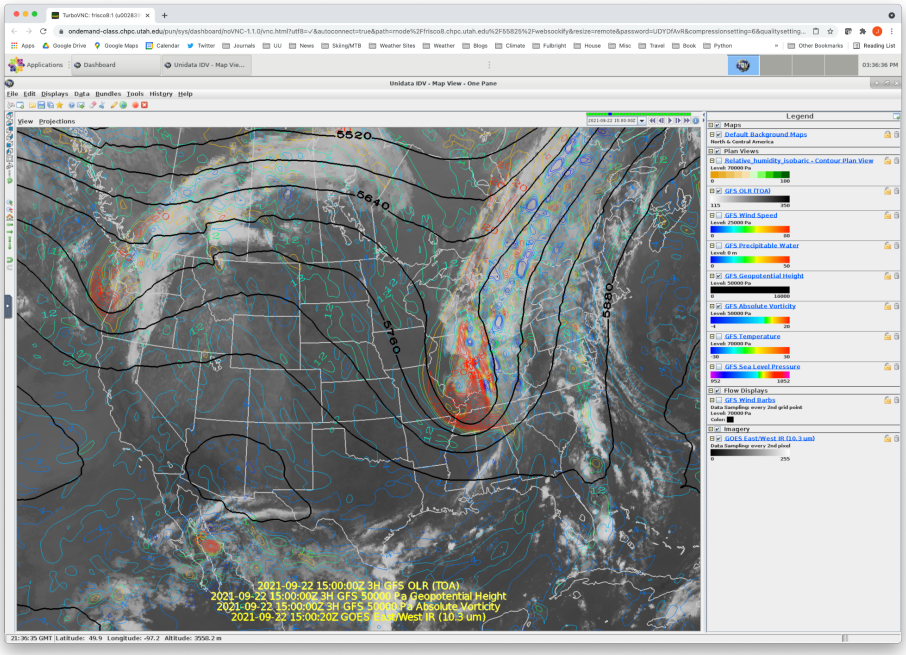
__Latest Version of VisAD__

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

IDV Community Support

In the transition to a remote-learning system as a result of the COVID-19 pandemic, help local MSI university to borrow several refurbished MacBook computers for the students to be able to run Unidata's Integrated Data Viewer (IDV) at home. I provided a remote IDV training class to a group of radar class students from the University of Millersville. Yuan also prepares several remote IDV training classes for the coming school year.

Another interesting community usage of the IDV is a browser based IDV by the University of Utah. IDV running on a linux server, displayed and controlled through browser with "openondemand". With a combination of the existing plugins, the HPC center can launch a browser based IDV for those students taking labs remotely:



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

__2021 AGU Fall Meeting__

- Impact of GNSS radio occultation data on the prediction of convective systems associated with a Mei-Yu front
- Machine Learning Predictive Model of Ice Supersaturated Regions (ISSRs) with Advanced 3D Visualization and Analysis

__2022 AMS Annual Meeting__

- Cross boundary subset of 2D/3D Grid coverage dataset and visualization

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 finally migrate away from Oracle Java 8u51 and towards a more modern version of OPenJDK Java. This switch will necessitate altering the IDV building and distribution workflow to work with OpenJDK.

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 October, 2021

Status Report: Information Technology

April 2021- October 2021

Mike Schmidt, Matt Perna, & Jennifer Oxelson

Major Activities

****Remote working**** -- with the entire Unidata staff working remote, it's been a successful test of the infrastructure including bandwidth through the gateway systems and remote supporting staff with software and hardware issues and well as security and update management. Fortunately, we now have all three people listed above allowed and certified to be onsite if the circumstances warrant.

****Network upgrades**** -- as UCAR upgrades their backbone infrastructure to 100Gb/s links, we will continue to upgrade our data movers (IDD cluster nodes, data aggregators, motherlode clones) from 2 x 1Gbp/s bonded ethernet to 10Gb/s as necessary.

****Security**** -- we continue efforts to keep services and systems secure which takes consistent attention and occasional herculean efforts (to patch everything all at once). Unidata staff have moved as a group to use Duo two factor authentication. Initial access to most Unidata and UCAR resources requires some form of two factor authentication.

****LDM 7 node**** -- we maintain a LDM7 test node at the Front Range GigaPOP (FRGP) just off downtown Denver in co-location with the major backbone networks supporting FRPG 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

Prepared October, *2021*

Status Report: LDM

May 2021- October 2021

Steve Emmerson, Tom Yoksas, Mike Schmidt, Mustapha Iles, Yuanlong Tan (University of Virginia)

Activities Since the Last Status Report

LDM

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:
  - Added support systems for that don't have rpcgen(1) installed
  - Added support for systems that use chronyd(8) instead of ntpdate(8) (e.g., RHEL 8)
  - Improved robustness on SunOS systems
- ldmadmin(1):
  - Improved robustness
  - Improved documentation
  - Added "showsettings" command to assist troubleshooting by support
- pqcreate(1):
  - Increased assumed product-size from 51 kB to 140 kB when the number of slots is defaulted. This value will be wrong, but should be closer.
- pqact(1):
  - Improved logging
  - Improved robustness when a single entry monopolizes its pqact(1) process for 6 hours.
  - Modified the "-strip" option to stop deleting characters whose unsigned value is greater than 127. This allows textual products in an extended ASCII character set (e.g., ISO 8859-1).
- scour(1):
  - Replaced scour(1) script with multi-threaded C program that scours much faster than the script
  - Added "-d" option to scour empty directories
- Documentation:
  - Updated the instructions on using metrics to monitor the LDM
  - Remove obsolete feedtype webpages
- NOAAPort:
  - Made updating the GEMPAK tables (ldmadmin updategempaktables) more robust
  - Added many new parameters to GEMPAK tables
  - Added option "-R <buf_size>" to noaaportIngester(1) to set the size of the

        receive buffer in order to reduce the number of missed packets. The operating system must allow the new size.
    - Made the assignment to the NEXRAD feedtype more robust
- Miscellaneous: Removed lint identified by clang(1) and Coverity Scan.
- Support:
    - Answered many questions from Universities, NOAA, US Military, and corporations
    - Troubleshot several sites that were having problems that were, overwhelmingly, network-related
    - A 14 hour, virtual workshop on the LDM was held for NOAA/NWS/NCEP/NCO/IDP. (They paid for it.) It went well.
- Released versions 6.13.14 through 6.13.17

Dependencies, challenges, problems, and risks include:

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).

Multicast LDM (aka LDM-7)

The multicast LDM project was separately funded by NSF's Computer and Information Science and Engineering directorate. The goal was to reduce the outgoing bandwidth requirement of the LDM -- yet retain the current level of reliability -- by converting it into a hybrid system that combines use of a new, semi-reliable multicast protocol -- developed jointly with the University of Virginia -- with the time-tested unicast capability of the current LDM-6.

This project ended September 30th, 2021.

Dependencies, challenges, problems, and risks include:

- The amount of manual intervention required to implement and maintain the multipoint VLAN used by LDM-7 renders this hybrid solution impractical for all but very large entities that distribute data via a dedicated, centrally-administered network.

Ongoing Activities

We plan to continue the following activities:

- Support and maintain the LDM

Relevant Metrics

- Data on LDM downloads
- The LDM system at the Unidata Program Center powers the Unidata IDD (Internet

Data Distribution) system. Metrics on that program can be found in the IDD status report.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

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

3. **Supporting People**
   By answering support questions, writing documentation, and conducting workshops.

Prepared *October 2021*

Status Report: McIDAS

April 2021- October 2021

Tom Yoksas

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Are there any features that users would like to be added to Unidata McIDAS-X and/or ldm-mcidas?

Activities Since the Last Status Report

Aside from routine updates/bugfixes to existing code and tables, the main area of activity recently has been the incorporation of Unidata-developed code into the next UW/SSEC McIDAS-X/XCD release, v2021. The main development area that we have been involved in is enhancing ADDE serving to better handle the various forms of ABI satellite imagery and Level 2 products.

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 1.7 TB/day)
- McIDAS-X is used to convert GOES-16 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 v2020 release featured the following:

- SSEC added support for McIDAS-X on RHEL 8 systems, ended support of McIDAS-X on RHEL 6 and Windows 7 systems, and added preliminary support of McIDAS-X on Windows 10 systems.
- Updated GOES-R Series ABI servers to list calculated resolutions in **IMGLIST** FORM=BAND and FORM=ALL output, and to support GOES-17 Fusion imagery because of modifications to allow more flexible file naming formats.
- Updated **GEO** command with improved logic when using Himawari imagery, and when merging imagery from multiple satellites and the domain contains the dateline.
- Improved performance of **RGBDISP** command, especially when it is run multiple times in a script.

- Updated VIIR servers and calibration module to correctly set the RAW value of Band 18 (M13) pixels to zero in bowtie deletion and bad or missing line regions.

Ongoing Activities

We plan to continue the following activities:

- SSEC McIDAS Advisory Committee (MAC)

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

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

- Interest in McIDAS by non-core users

  The UPC 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 existing and new community members

New Activities

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

Relevant Metrics

- Data delivered by the Unidata McIDAS ADDE servers exceeds 1 TB/day. The great majority of the data being served is imagery from GOES-16 followed by imagery from GOES-17.

- 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.

Planned for the next major addition to this package will be the development of a "decoder" for GRB delivered Geostationary Lightning Mapper (GLM) data. This development is aimed at greatly increasing the speed at which displays of the GLM data can be made in McIDAS-X, the IDV and McIDAS-V.

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 (GOES-17) data via McIDAS ADDE services on three publicly accessible servers (lead.unidata.ucar.edu, adde.ucar.edu (aka atm.ucar.edu) and adde.ssec.wisc.edu) has been averaging on the order of 48 TB/month for the past 14 months.

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-15 and GOES-14 imagery when available. GOES-15 and GOES-14 were put into standby mode on March 2, 2020. GOES-14 remains in its standby location (104W) and will be turned on for periodic testing as needed. GOES-15 supplemental operations began on Sunday, August 9, 2020 at 0000 UTC and continues to provide surveillance during Pacific hurricane seasons.

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

New Activities

Establish a testbed for generating Level 2 products from GOES-16/17 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 *October, 2021*

Status Report: netCDF

April 2021- October 2021

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Are there other cloud-based block storage formats/locations (TileDB, Azure, etc) that are actively in use? What is the next venue for investigation once we have our Zarr support in place?
2. What aspects of the modern AI/ML workflow might be improved by changes to the netCDF technical infrastructures.

Activities Since the Last Status Report

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 **192** open issues for netCDF-C, **71** open issues for netCDF-Fortran, and **44** 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 release of ncZarr (netCDF with native Zarr support) has been improved as of netCDF-C version 4.8.1.
- 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.

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 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

NetCDF/Zarr Integration

The netCDF team has released the first public version of netCDF-C which provides Zarr I/O compatibility, dubbed 'ncZarr'. This work has been highly anticipated, and well received, by the broader netCDF and Zarr communities.

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

- Release the next version of netCDF with Zarr+Xarray support (ncZarr).
- 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

Static Analysis Metrics

There are currently about 226,892 lines of code (up from 202,428 lines of code) in the netCDF C library source. The Coverity estimate for defect density (the number of defects per thousand lines of code) in the netCDF C library source has slightly decreased to **0.68,** where it was **0.68** six months ago. According to Coverity static analysis of over 250 million lines of open source projects that use their analysis tools, the average defect density with 100,000 to 500,000 lines of code is **0.50**.

Google 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:

- **1,020,000** for netCDF-3
- **2,110,000** for netCDF-4
- **921,000** for HDF5
- **113,000** for GRIB2

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

- **461** for netCDF-3
- **1,060** for netCDF-4
- **18,800** for HDF5
- **1,440** for GRIB2

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 *April, 2021*

Status Report: Outreach to Underserved Communities

April 2021- October 2021

Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We continue to request your feedback on the following topics:

Are you currently collaborating with an MSI?

Are there MSI's geographically close to you that you have not engaged with?

Are there other underrepresented communities you would like to engage with?

Activities Since the Last Status Report

- Proposal to NSF solicitation CISE 21-533, Titled:
  A Sovereign Network System for Environmental Monitoring, Data and Information Exchange, and Collaboration among Tribal Colleges and Universities
  Was fully funded
  -Instrumentation (towers, soil moisture sensors, air quality sensors) will be mostly deployed tihs fall with some in the spring
  -Workshops involving LDM configuration, IDV visualizations, RAMADDA install and set-up are being planned at SIPI, NTU, and UCAR. These will be scheduled soon and will likely be virtual this fall, hopefully going to in person Spring/Summer 2022.
  -Jeff Weber, Co-PI, may do site visits to SIPI and NTU this fall if pandemic allows.
- Institutions involved with Unidata are: Southwestern Indian Polytechnic Institute, Navajo Technical University, and Tohono O'odham Community College
- We are also engaging with 2 field campaigns (SAIL (DOE) and SPLASH (NOAA)) to examine the potential for collaboration. SAIL and SPLASH are monitoring the snow/precip process as it relates to discharge into the Colorado River. TOCC and NTU are in the Colorado River basin and will be collecting met tower data including a full radiation suite of sensors, precip, soil moisture, and air quality. This will help us demonstrate the abilty to share some data, while holding other data sovereign to the tribe, college, or institute.

Engaged with Rising Voices

- Involved with Indigenous Peoples Climate Change Working Group **(IPCCWG)**
- Involved with Indigenous-**FEWSS**. Indigenous Food, Energy & Water Security and Sovereignty

Internships

- Active engagement in the SOARS program
    - Will participate on selection committee for 2022

Exploring the possibility to work with SAIL and SPLASH this summer from TOCC, NTU, or SIPI as part of the Unidata Intern Program

Progress has been made on the following:

- Designing structural changes (e.g. modifications to how equipment awards, internships, workshops, and committee placements are announced and selected)

Dependencies, challenges, problems, and risks include:

- The only known dependency is regarding funding and time both of which have been dedicated to this effort. Since this is a new project, other dependencies or risks have not been identified at this time.

Ongoing Activities

We plan to continue the following activities:

- SACNAS and Rising Voices engagement

New Activities

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

- Continue to develop outreach stakeholder list for broadening distribution of opportunities (equipment awards, internships, etc) and engage remotely
- SIPI and NTU site visits

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

- Implement changes to the process of how Unidata opportunities are announced and awarded
- Plan exhibition or other activities at subsequent appropriate conferences
- Identify relevant metrics (contacts, partners identified, meetings attended)
- Identify sustaining partnerships for the next five years
- Engage other underrepresented communities

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

- This effort is an ongoing commitment for the next award period (5 years); however, during this first year we are piloting efforts and then will apply lessons learned for the next 4 years.

Relevant Metrics

Relevant metrics should be discussed and decided for reports going forwarded

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Providing Useful Tools**
   Better understanding these communities and characterizing their needs will allow us to develop more fit for purpose tools that can and will be adopted

2. **Supporting People**
   Unidata has always served the broad geoscience community; however we are making a concerted effort to expand our reach to underrepresented individuals and organizations as an emphasis of our new award

---

Prepared *April, 2021*

Status Report: Python

April 2021- October 2021

Ryan May, Drew Camron, Sean Arms, Julien Chastang, Nicole Corbin

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What are the pain points you encounter when trying to use Python for your work? When using MetPy and/or Siphon?

Activities Since the Last Status Report

Python Training Efforts

Python training efforts continue to be an important part of the Python portfolio. We continue to be successful in identifying opportunities to offer training within our resource constraints. Not only do these generate significant goodwill and grow our audience, but they are a significant source of information to inform our library development. The spin-up of Unidata's in-house instructional design staff has created benefits for the organization and effectiveness of synchronous virtual offerings, as well as important insight for stronger design of asynchronous learning materials both internal and on Unidata collaborations. Our eye is towards leveraging the success of our burgeoning virtual workshop presence to further engage our community to the best of our ability. This includes tailoring the packaging of our virtual offering to be more portable and reusable, identifying needs for content for synchronous workshops, and considerations for altogether new educational avenues moving forward. We also look to refine our resources to be more accessible and directly integrate them into the needs of new institutions and organizations as yet unreached by us.

To support this, we are nearer to completing the unification of Unidata's Python training materials using sustainable technologies in-line with our other Python projects, including an automatically tested and deployed Sphinx-based website with Unidata-consistent theming. Lessons learned from our educational collaborations and the fast-moving Jupyter ecosystem are being leveraged to fully take advantage of the presence of our online materials to best fit the needs of our community.

Progress has been made on the following:

- **Unidata** hosted a synchronous half-day virtual *Intro to MetPy* workshop, benefitting from Unidata's JupyterHub on Science Gateway resources, and featuring a tailor-made asynchronous pre-workshop course created by Unidata's instructional designer
- Unidata continues to be a primary collaborator on *Project Pythia*, which is beginning to establish itself as an authoritative presence on Python learning in geoscience via Jupyter
- The AMS Intermediate Short Course on MetPy collaboration with Kevin Goebbert,

delayed until the 2022 AMS Annual Meeting, is scheduled to occur in-person at the conference.
- Unidata is scheduled to host a synchronous, in-person training session at the 2022 AMS Annual Meeting Student Conference
- John Leeman continues to lead the “MetPy Mondays” effort.

MetPy

Development continues to be driven by requirements for our dedicated awards (in addition to bug reports and pull requests from community members). MetPy 1.1 was released with a variety of fixes and enhancements including

- Many improvements to the structure and reliability of the METAR parser, including support for returning more fields
- PlotGeometry class added to the simplified plotting interface, which makes it straightforward to plot geometric products, like SPC outlooks or NHC forecast cones. This was created by Connor Cozad, one of Unidata’s summer interns.
- Various improvements to the simplified plotting interface
- Support for parsing GEMPAK data files. This was created by community member Nathan Wendt from SPC.
- Showalter index calculation contributed by NCAR
- Dropped support for Python <3.7 and Numpy <1.17

Moving forward, 1.2 is planned for release late Fall and will include a variety of improvements in the simplified plotting interface and elsewhere, spearheaded by community members transitioning from GEMPAK. This release will also feature enhanced support for calculating derivatives in projected coordinates (the so-called “map factor” issue). Beyond that, a prototype of the “automated solver” has been created. Developments beyond that will include finally landing support for parsing WPC’s surface analysis bulletins and plotting fronts. We will also begin the performance work that was featured in the project’s recent CSSI award.

It should also be noted that we have continued to see an uptick in community engagement. GitHub issue reports, discussions, and most importantly, Pull Requests, have seen an increase overall of ~20% year-over-year, along with continued growth in other support forums like StackOverflow and Unidata’s E-support system. Being able to continue to continue to service the community’s needs, along with all of the rest of the team’s obligations, will require continued creative solutions and strong engagement in public forums. This allows other members of the user community to continue their help and involvement.

Progress has been made on the following:

- MetPy 1.1 was released Summer 2021
- Work towards requirements of MetPy-related NSF awards
- Community awareness continues to grow, with the volume of engagement and mentions on social media growing; the MetPy twitter account has reached 2044. followers.
- We are working on a manuscript draft for an article on MetPy for the Bulletin of the AMS.
- Prototype “automated solver” has been created

Siphon and Data Processing

Siphon continues to grow and develop, though at a slower pace than MetPy; its development tends to be driven by obstacles to access of remote data. 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. Siphon has seen infrastructure and usability improvements over these months, and 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.

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.

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 a new version of the Python Training website using sustainable technologies
- Present a short course at AMS 2020 Annual Meeting on MetPy aimed at intermediate users
- Release MetPy 1.2 late Fall 2021
- Engage in AMS 2022 student conference
- Engage in the rollout of Project Pythia and adjacent UCAR Python education efforts

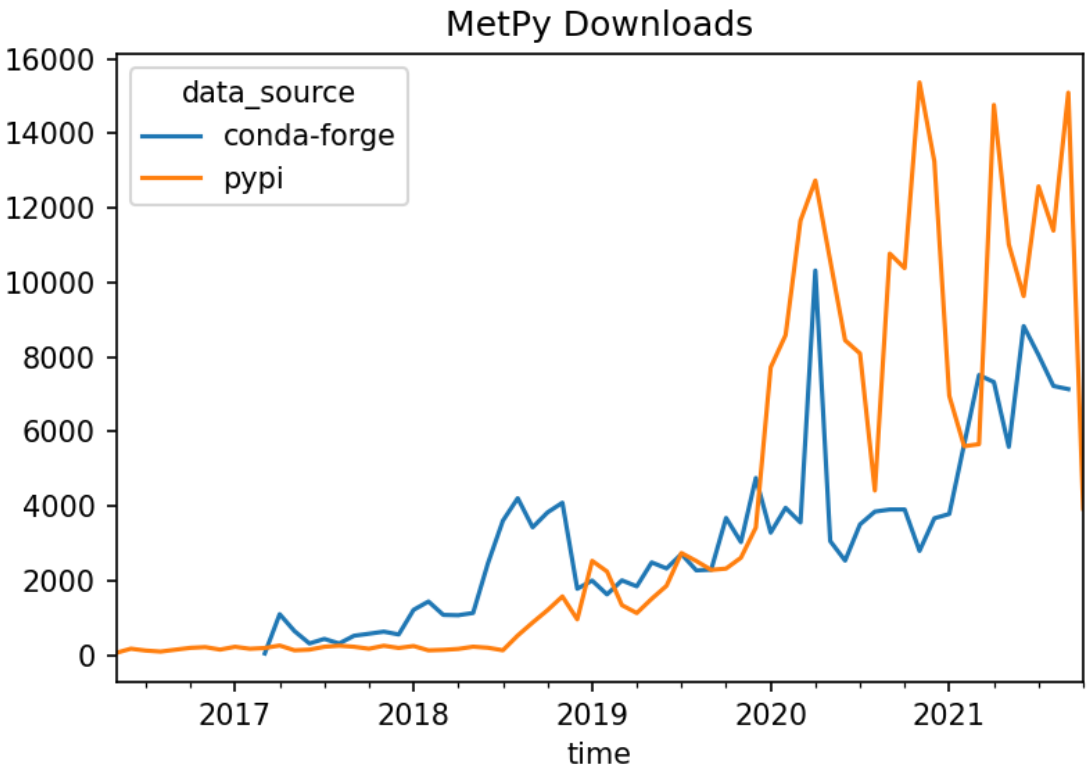
Over the next twelve months, we plan to organize or take part in the following:

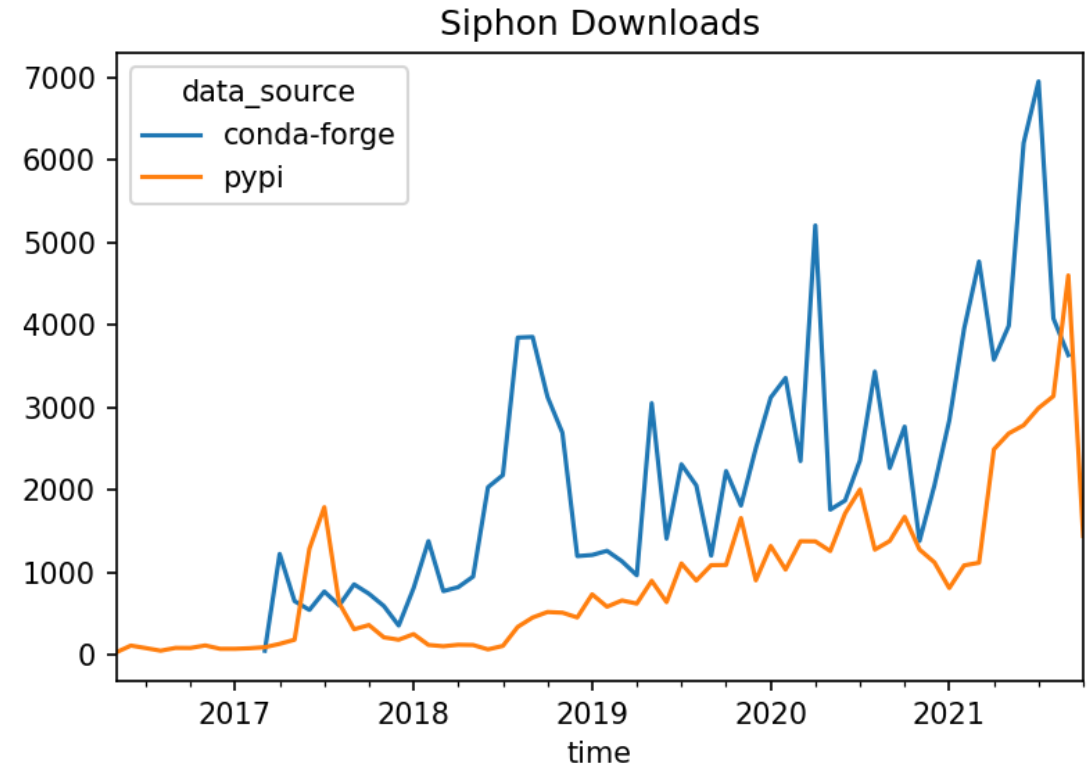
- Offer additional virtual MetPy workshops
- Separate non-TDS siphon capability into new MetPy remote functionality
- Evaluate ways to improve MetPy performance and scalability using tools like Dask and Numba

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

- Evaluate the possibility of extending siphon functionality to interface with the AWIPS-II EDEX server

Relevant Metrics


MetPy Downloads
data_source
conda-forge
pypi
16000
14000
12000
10000
8000
6000
4000
2000
0
2017
2018
2019
2020
2021
time


Siphon Downloads
data_source
conda-forge
pypi
7000
6000
5000
4000
3000
2000
1000
0
2017
2018
2019
2020
2021
time

MetPy

- 96% test coverage
- Watchers: 59
- According to GitHub, 197 repositories and 20 packages depend on MetPy
- Downloads for the releases made in the last year (Conda + PyPI):
    - 1.0: 43634
    - 1.0.1: 45291
    - 1.1: 21686
- Since 1 April 2021
    - Active Issues: 164 (92 created, 79 closed)
    - Active PRs: 254 (229 created, 228 closed)
    - External Issue Activity: 64 opened, 177 comments
    - External PR Activity: 58 opened, 96 comments
    - Unique external contributors: 62
    - Stars: 72 (789 total)
    - Forks: 2 (292 total)
    - Commits: 552
- Since 1 October 2020
    - Active Issues: 235 (159 created, 123 closed)
    - Active PRs: 431 (397 created, 387 closed)
    - External Issue Activity: 105 opened, 288 comments
    - External PR Activity: 84 opened, 151 comments
    - Unique external contributors: 94
    - Stars: 165 (789 total)
    - Forks: 5 (292 total)
    - Commits: 945

Siphon

- 98% test coverage
- Watchers: 15
- According to GitHub, 85 repositories and 9 packages depend on MetPy
- Downloads for releases made in the last year (Conda + PyPI):
    - 0.9.0: 35452
- Since 1 April 2021
    - Active Issues: 4 (4 created, 0 closed)
    - Active PRs: 81 (76 created, 63 closed)
    - External Issue Activity: 1 opened, 3 comments
    - External PR Activity: 1 opened, 0 comments
    - Unique external contributors: 1
    - Stars: 14 (148 total)
    - Forks: 1 (56 total)
    - Commits: 38
- Since 1 October 2020
    - Active Issues: 24 (12 created, 6 closed)
    - Active PRs: 116 (112 created, 97 closed)

- External Issue Activity: 6 opened, 11 comments
- External PR Activity: 2 opened, 0 comments
- Unique external contributors: 9
- Stars: 23 (148 total)
- Forks: 2 (56 total)
- Commits: 120

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, 2021*

Status Report: Support

April 2021- October 2021

Jennifer Oxelson, Tom Yoksas, UPC Staff

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Is the support that we provide sufficient for the community's needs?

   If not, what else should we be doing?

Activities Since the Last Status Report

Training

- From 2018 until the end of March, 2020, the UPC was focusing its in-person training efforts on regional workshops and short courses. Since the end of March 2020, numerous virtual workshops have been held for Python (MetPy), AWIPS and LDM training
- Additional resources will be directed towards developing online training materials.
- Unidata recently hired a specialist in online training to help with modernization of our training materials

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 65600 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.

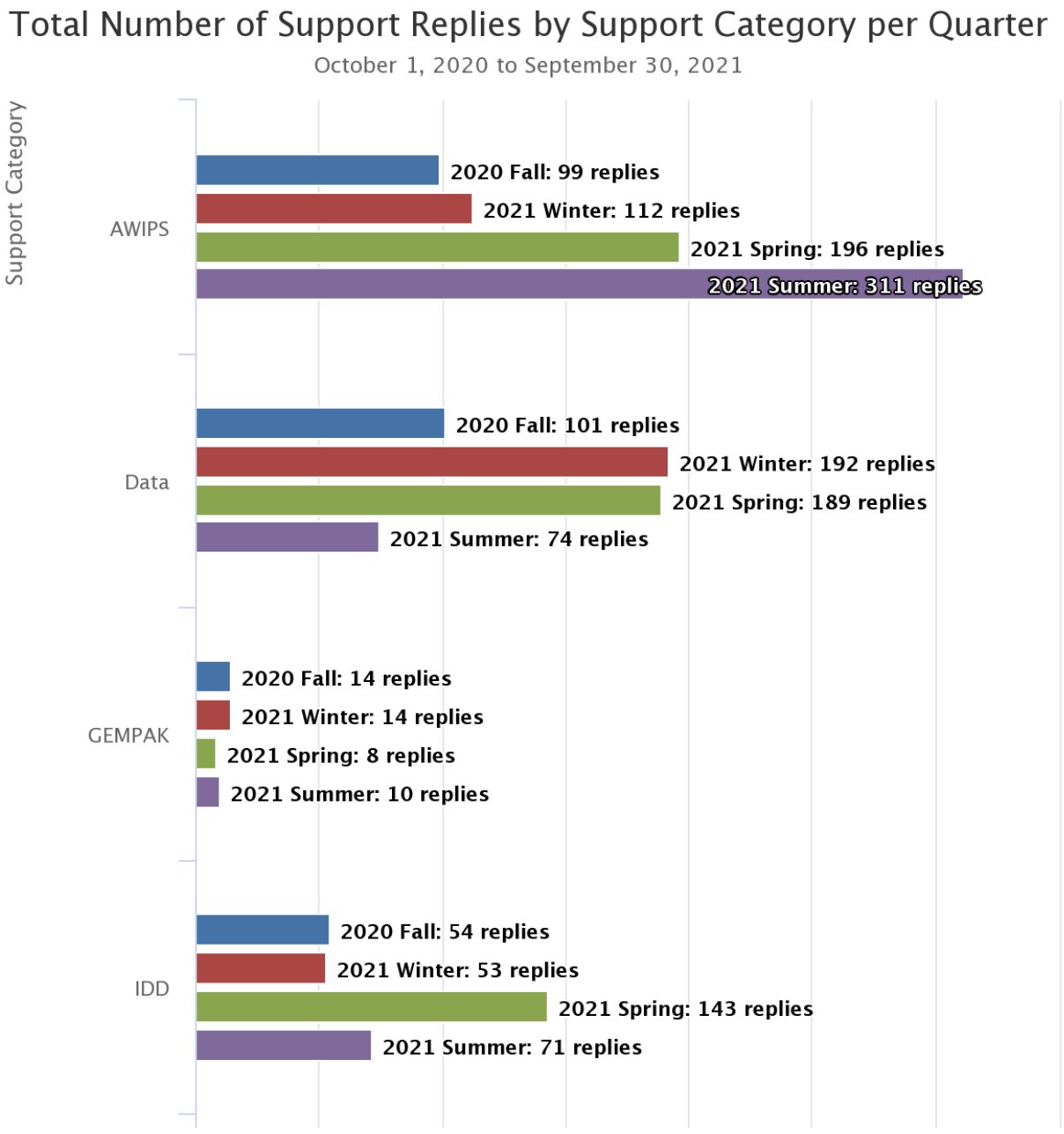
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

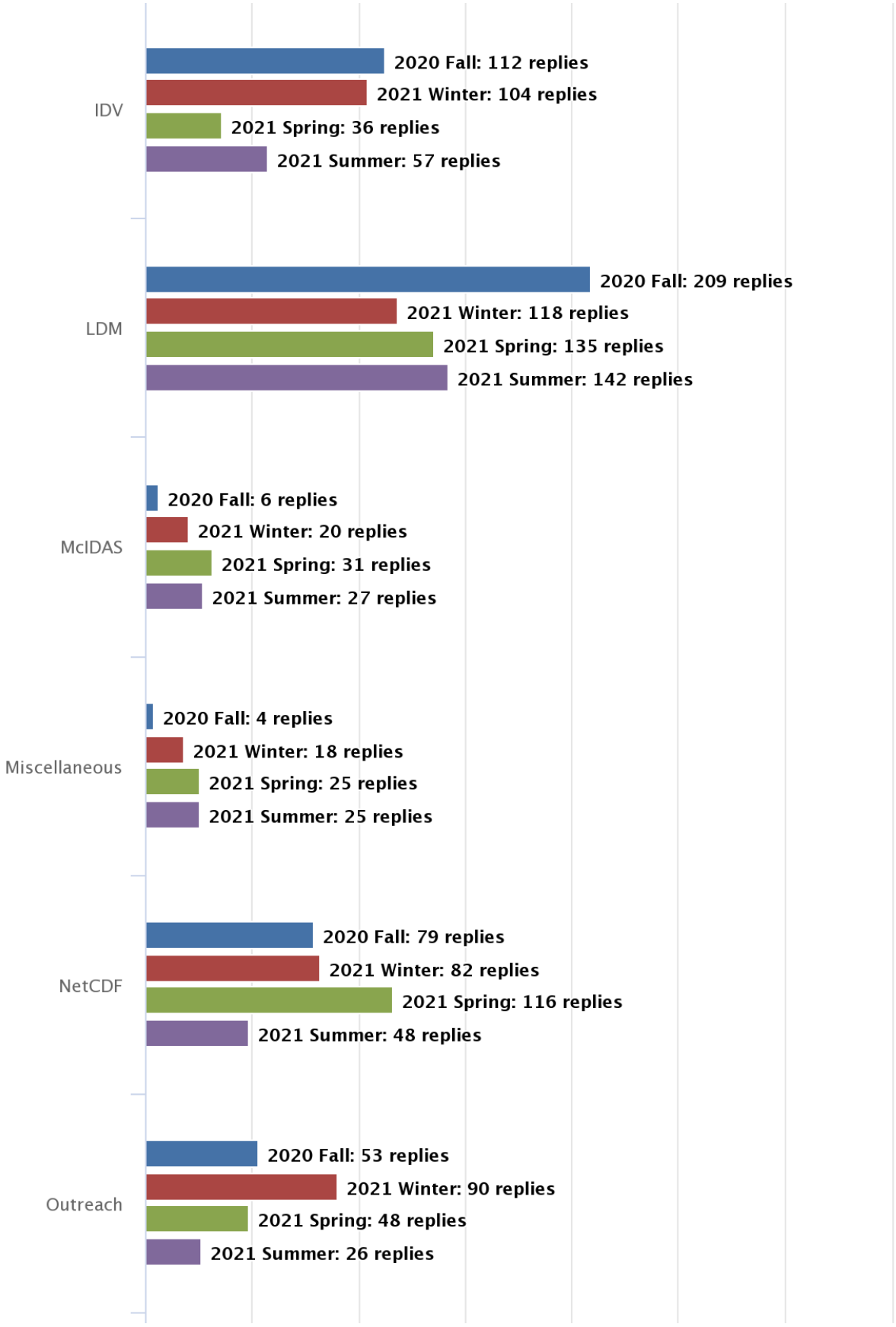
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 12 month period

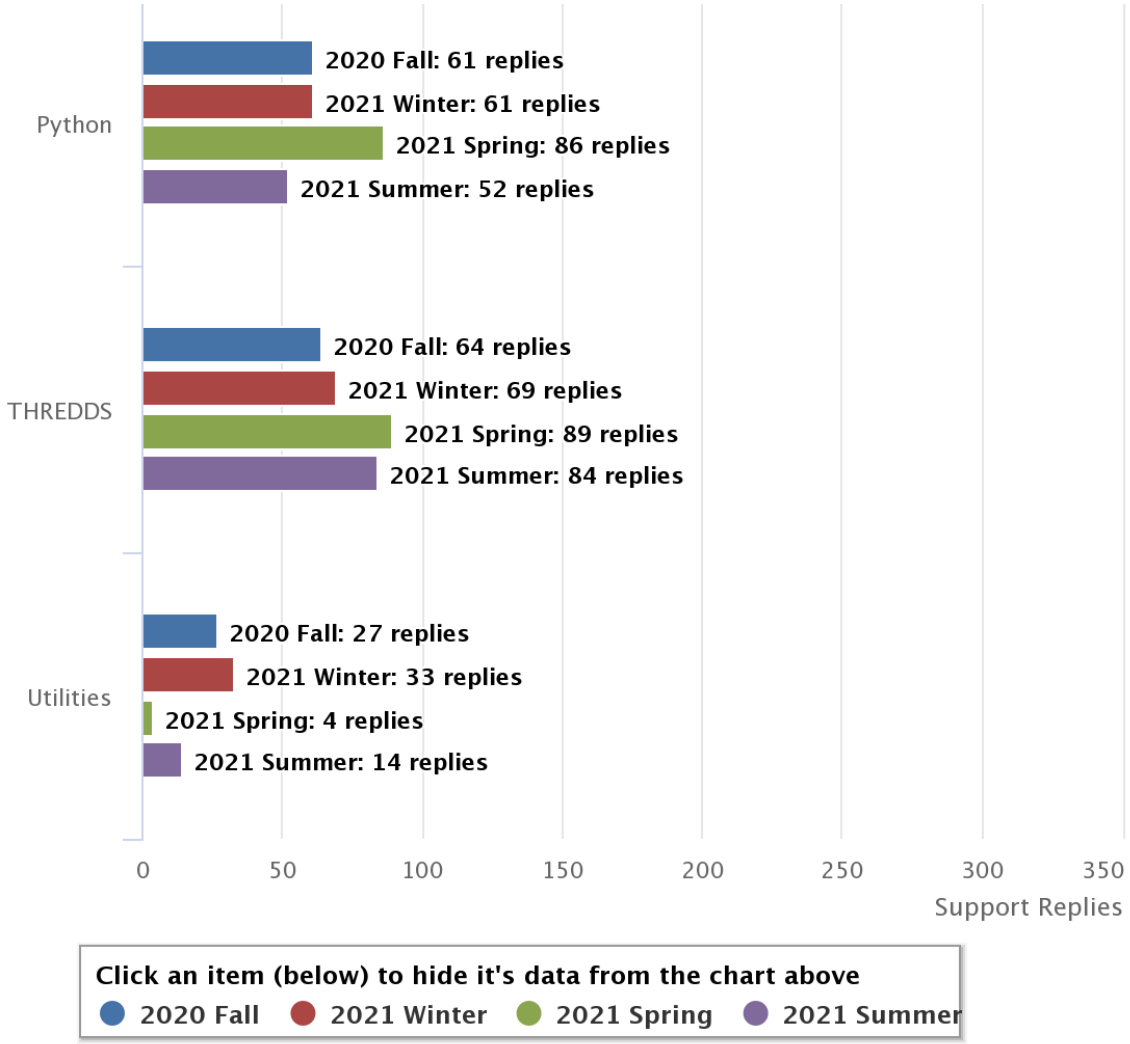
from October 1, 2020 until September 30, 2021.

The quarters shown are defined as:

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




IDV
2020 Fall: 112 replies
2021 Winter: 104 replies
2021 Spring: 36 replies
2021 Summer: 57 replies
LDM
2020 Fall: 209 replies
2021 Winter: 118 replies
2021 Spring: 135 replies
2021 Summer: 142 replies
McIDAS
2020 Fall: 6 replies
2021 Winter: 20 replies
2021 Spring: 31 replies
2021 Summer: 27 replies
Miscellaneous
2020 Fall: 4 replies
2021 Winter: 18 replies
2021 Spring: 25 replies
2021 Summer: 25 replies
NetCDF
2020 Fall: 79 replies
2021 Winter: 82 replies
2021 Spring: 116 replies
2021 Summer: 48 replies
Outreach
2020 Fall: 53 replies
2021 Winter: 90 replies
2021 Spring: 48 replies
2021 Summer: 26 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
Miscellaneous	Administration, Development, Plaza, Staging Folder, Support, Support eSupport, Support Miscellaneous, Support Platforms, Support Plaza, Student Interns, Systems
NetCDF	Support LibCF, Support netCDF
Outreach	Outreach, Polcomm, Science Gateway, Support Egrants, 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 by 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 levels 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.

Prepared *October, 2021*

Status Report: THREDDS

April 2021- October 2021

Sean Arms, Hailey Johnson, Jennifer Oxelson, Ryan May, Ethan Davis, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. How can we help you and your students? We can do much more than Java programming - we love Python too! Our team comes from a variety of academic backgrounds as well, including Meteorology (boundary-, surface-, and canopy-layer, in-situ observations, radar), Computer Science, Oceanography, Chemistry, and Physics!
2. Have you ever used one of the Web Start viewers from the TDS? If so, be aware that as of Java Web Start has been deprecated as of Java 9, and has been removed in Java 11 onward. While TDS v5 does not include Java Web Start based viewers, it does include the updated Godiva3 client, as well as a new Jupyter Notebook viewer service. Will this impact your education and research efforts?
3. Do you have thoughts on the migration of the TDS to microservices? What do you foresee as the greatest challenges and benefits associated with the change?

Activities Since the Last Status Report

Staffing Changes

As of 17 September 2021, Sean Arms has moved on from Unidata to pursue new opportunities. Unidata has received NSF funding for a two-year THREDDS developer position, and we will be beginning that search shortly.

The THREDDS Project

The THREDDS Project encompases four projects: **netCDF-Java, the THREDDS Data Server (TDS), Rosetta, and Siphon** (the Unidata Python client to interact with remote data services, such as those provided by the TDS). For specific information on Siphon, please see the Python Status Report. This status report contains updates on cloud compatibility within netCDF-Java and the TDS; for updates on further cloud efforts, including the popular Docker container effort, please see the Cloud Computing Activities Status Report.

Released netCDF-Java 5.4.2

- NetCDF-Java version 5.4.2 was released on 27 August 2021.
- The new release includes updates to the following GRIB tables:
    - MRMS GRIB2 Tables to v12.0.
    - NCEP GRIB 1 Tables.
    - Experimental HRRR GRIB2 tables to v4.

    - NDFD GRIB2 tables to reflect degrib v2.25.
- Support for GRIB2 PNG encoding (24- and 32-bit depth) is new in 5.4.2

Released netCDF-Java 6 Beta1

- Version 6 of the library was released in beta on 24 May 2021. So far, feedback has been minimal.
- The most notable change in version 6 is introduction of fully immutable data structures.
- More information on netCDF-Java versions can be found on the netCDF-Java github wiki.

Released TDS version 4.6.17 (Maintenance)

- TDS version 4.6.17 was released on 13 September 2021. This release contains a variety of bug fixes, as well as updates to third-party libraries, including security updates. Version 4.6.17 of the TDS is considered a maintenance release; we will continue to provide security updates through May 2022. We recommend upgrading to TDS v5 at or prior to that time.

Released TDS version 5.0 (Stable)

- TDS version 5.0 was released on 20 September 2021. This is the first feature-complete release of TDS v5.
- TDS v5 requires the use of Java 11 at a minimum (TDS 4.6.x only requires Java 8).
- Starting with version 5, the TDS has been decoupled from netCDF-Java, allowing the two products to be released separately. We anticipate that the TDS will not move to netCDF-Java 6+, but will continue to use the latest release of netCDF-Java 5.x.
- We anticipate releasing a stable version of TDS 5.0.0 this summer.
- TDS v5 includes general object store support (AWS S3 compatible API). In partnership with NASA JPL, initial TDS benchmarks show that the move from local disk based data holdings to Object Store based data holdings is mostly a horizontal move in terms of performance from an end users perspective.

Rosetta

Rosetta continues to progress following a very successful NASA ACCESS grant (the Oceanographic In-situ data Interoperability Project, or **OIIP**), in which Unidata partnered with the PO.DAAC at JPL and UMASS-Boston.

During this period of flux and low-staffing for the THREDDS projects, Rosetta will temporarily enter a maintenance mode; no new development is planned for the short-term future.

Ongoing Activities

We plan to continue the following activities:

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.

Development

- netCDF-Java
    - Wrap-up initial work on version 7 of netCDF-Java, using immutable arrays and requiring Java 11 or higher.
    - Work on version 8 of netCDF-Java will begin soon. This will be the first version of netCDF-Java which uses only the new public API. __Many thanks to John Caron, who continues to make the vast majority of this work happen__ .
    - Continued work to implement read support for Zarr and NCZarr. Zarr support is expected to be included in the next minor release of netCDF-Java (5.4.3).
    - Continued work to curate existing documentation into four documentation sets: tutorial, developer/reference (nitty-gritty details, for those interested in learning more or hacking on the netCDF-java codebase), THREDDS Client Catalogs (language agnostic), and NetCDF Markup Language (NcML) (language agnostic).
- gCDM (gRPC for the Common Data Model)
    - gCDM is an ongoing effort which currently exists for netCDF-Java versions 6 and 7 (branches __6.x__ and __develop__) and is a new way to allow netCDF-Java to communicate remotely. See April 2021 THREDDS status report for an in depth description of the gCDM project.
    - gCDM stands for "gRPC Cdm", where gRPC is a recursive acronym that stands for "gRPC Remote Procedure Calls". For more information on gRPC, checkout the gRPC FAQ.
    - Work is beginning soon to port gCDM bask to netCDF-Java version 5.
    - Porting gCDM to netCDF-Java 5.x will allow the TDS to include a gCDM endpoint. This will be part of a collaborative effort with the AI/ML group and NASA JPL to create a machine learning data pipeline for oceanic anomalies. The THREDDS component of the project involves exposing TDS-hosted data to an ML client via gRPC.

The following active proposals directly involve THREDDS work:

- USGS has awarded Unidata a project with the goal of improving the TDS's ability to leverage cloud computing technologies (like object stores) to serve very large datasets. The award will support just over 0.5 FTE over the next year. The three focus areas are 1) Zarr support in netCDF-Java and TDS; 2) add a TDS Service API to simplify adding new data services to the TDS; and 3) add support for the OGC EDR (Environmental Data Retrieval) API using the new TDS Service API.
- We are in our final six months of the NOAA IOOS award titled "A Unified Framework for IOOS Model Data Access", in which we are partnered with Rich Signell and Axiom

Data Science. The goal is to enable support of the UGRID specification within the THREDDS stack, as well as create a GRID featureType to allow for serving large collections of gridded datasets (including UGRID). This work **strategically aligns with the Unidata 2024 focus area “Managing Geoscience Data, Making Geoscience Data Accessible** by improving the reliability and scalability of the TDS to handle very large collections of gridded datasets, as well as **“Managing Geoscience Data, Enhancing Community Access to Data”** through the addition of UGRID support (example: MPAS output is on a mesh, a.k.a. “unstructured”, grid).

- Step 1 proposal submitted for NASA Advanced Information Systems Technology (AIST) grant. This proposal includes continued work on gCDM (see above). More information on the proposal can be found in the AI/ML status report.

New Activities

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

- netCDF-Java
    - Release Initial support for reading Zarr and ncZarr.
    - Begin beta testing of netCDF-Java 7.
    - Enable gCDM in netCDF-Java 5.

- TDS
    - Better curate existing documentation into four documentation sets: server administrator (with quick start guide), end user (browser), developer (web access via api), reference (nitty-gritty details, for those interested in learning more or hacking on the TDS codebase).

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

- netCDF-Java
    - Obtain 90%+ test coverage of the public API in version 6+.
    - Initial support for the majority of NumCodecs supported filters for reading Zarr and NCZarr.
    - Complete unstructured grid support.
    - Release netCDF-Java version 8.
- TDS
    - Re-evaluate new development of the TDS in light of gCDM.
    - Re-evaluate the TDS dependency on Java and consider development options to optimize maintainability.

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

- netCDF-Java
    - Fully support Java 11 and the Java Platform Module System (end of Java 8 support)
- TDS
    - Begin development of a new product based on microservices.

Relevant Metrics

NetCDF-Java

End of 12 month period	Number of Downloads (cdm/cdm-core.jar, netcdfall.jar, toolsui.jar)
2020-02-29	36,897
2020-07-31	49,181
2021-03-31	104,516
2021-10-08	

THREDDS Data Server

Downloads

End of 12 month period	Number of Downloads (thredds.war, thredds-classes.jar)
2020-03-31	2,218
2021-10-08	

Startup Metrics

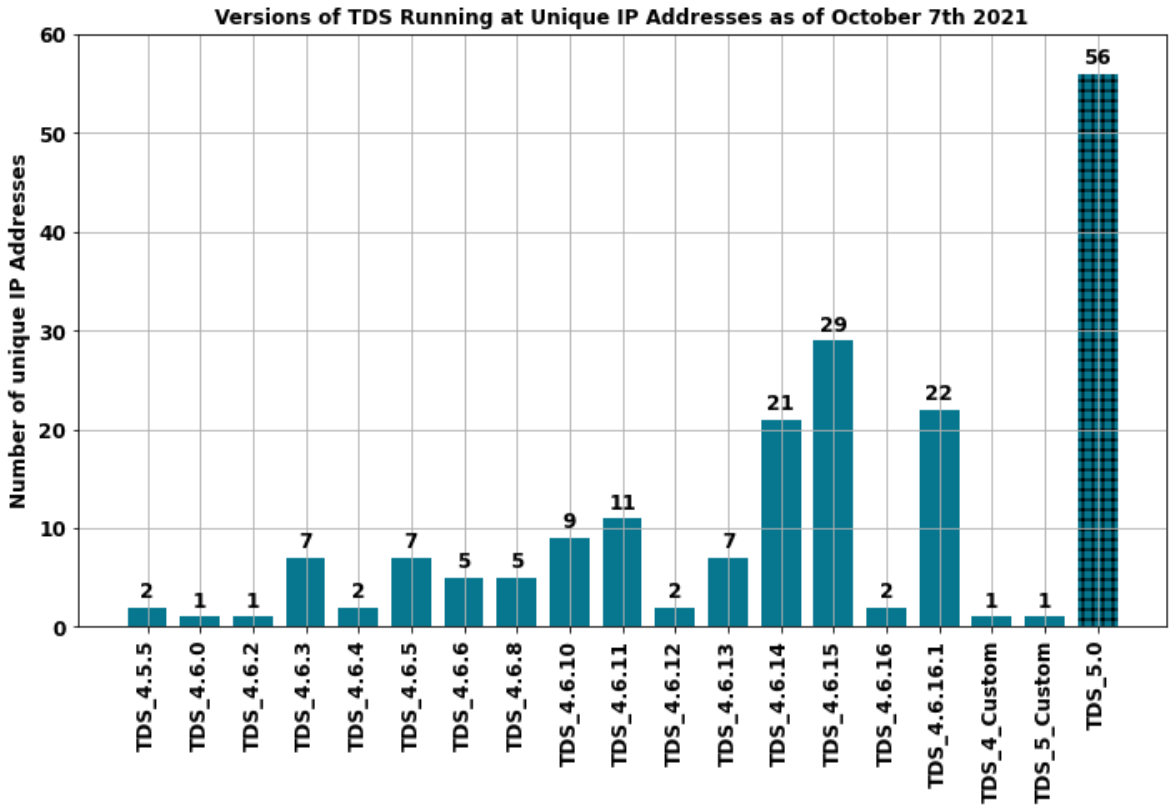
	2021-04 — 2021-10	2014-08 — 2021-04
TDS Startup (unique IP address count)	3,522	36,974

	Total Servers	Information page updated

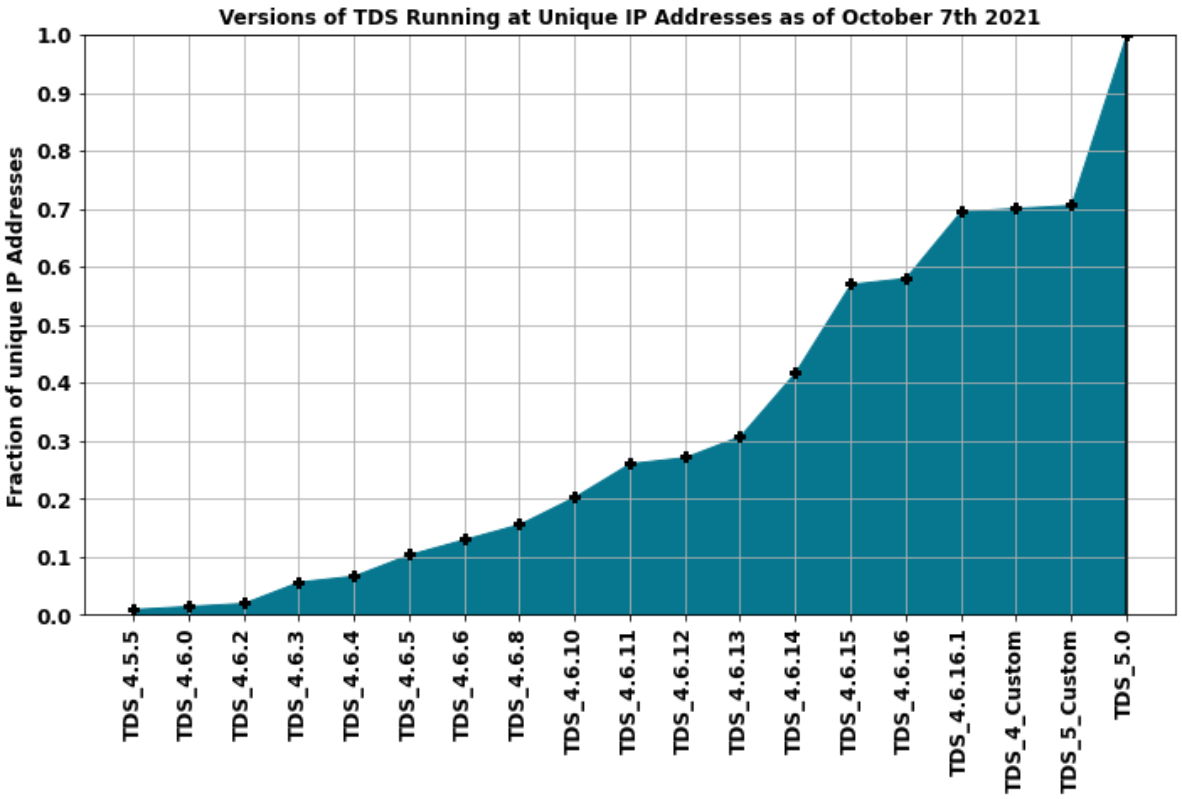
Publicly Accessible[1] TDS count	191	90

Over the past six months, **3,522** unique IPs started up the TDS (April 2021 through September 2021). Since we've started tracking these metrics (v4.5.3, August 26th, 2014), we've seen the TDS startup from **36,974** unique IP addresses. There are currently **191** publically accessible TDSs running "in the wild" (one less than our last report) . Furthermore, of the **191** publically accessible servers, **90** 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 (v5.0 was the most current release during this period, and was released on 20 September 2021). TDS v5 is the dominant specific version running in the wild.



[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.



In the next year, we will be working towards enabling TDSs, on an opt-in basis, to officially advertise their availability to the community through a centralized resource.

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 (such as adding Zarr support) and services.
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, 2021*