Unidata Users Committee
Virtual Meeting Agenda

(All times are Mountain Daylight Time)

Thursday, September 10th

10:00 - 10:20: Welcome, logistics, and prior action items (Kevin and Josh)

10:20 - 11:00: Director's Report (Mohan)

11:00 - 12:10 Break

12:10 - 13:00: Around-the-table Reports (Committee)

13:00 - 13:30: NOAA Update (Anne Myckow)

13:30 - 13:40 Break

13:40 - 14:25: IDV Demonstration (Yuan)

14:25: End Day 1 Sessions

Friday, September 11th

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

10:05 - 11:00: Status Reports (Kevin)

11:00 - 11:10: Break

11:10 - 12:00: Blue Skies

12:00 - 13:00: Break

13:00 - 13:30: MSI Report

13:30 - 14:15: AWIPS Demonstration (Shay and Tiffany)

14:15 - 14:30: Close Meeting (Review Actions and Identify Next Meeting Date)

Friday, November 13th

10:00 - 10:05: Opening

10:05 - 10:55: What two things should the community be using? Possibly with demonstration. (From Unidata staff)

- (Dennis Heimbigner) General Cloud storage/computing and how to transition (25 minutes)

- (Julien Chastang) Direct Access to Jetstream resources (25 minutes)

10:55 - 11:05: Break

11:05 - 11:55: What two things are on the horizon worth watching and playing with to determine usefulness? Possibly with a short demonstration. (From Unidata staff)

- (Ryan May) Jupyter Real-Time Collaboration: (https://github.com/jupyterlab/rtc) (25 minutes)
- Ward Fisher Virtual-Reality (VR) and efforts currently underway across UCAR (25 minutes)

11:55 - 13:00: Break

13:00 - 13:55: Student Panel (Lena Huescher and Josh Young)

13:55 - 14:00: Break

13:55 - 14:30: University needs for IT-level training on installing/maintaining/running Unidata tools. Consider what format is needed for training materials. (Kevin Goebbert)

14:30 - 14:55: Student Workshop at AMS (Sean Arms - To be confirmed)

14:55 - 15:00: Closing and review of Actions

Status Report: Users Committee Actions

April 2020 - September 2020

Unidata Program Center Staff

Actions from the Previous Meeting

Action 1

Incorporate Kevin's 4 questions more fully into the in-person agenda (Josh)

Result

Given the reduced footprint of the September meeting, the current suggestion is to meet more frequently. If this is acceptable to the committee than the 4 question format would be the focus of an abbreviated virtual meeting in November, at a date and time to be selected by the committee.

Action 2

Explore developing a CONDUIT users survey (Staff)

Result

This action has been extended as other priorities superseded it during the intervening months.

Action 3

Committee should by next Thursday (4/9/20) provide any further feedback to Unidata on the A/V Recommendations (Committee)Committee should by next Thursday (4/9/20) provide any further feedback to Unidata on the A/V Recommendations (Committee)

Result

The committee received notice of the invitation to provide further feedback with a request to do so in one week. Therefore, this item is now complete.

Action 4

Schedule conversation between to explore a 45 minute jupyter notebook demonstration. (Julien, Kim Wood, Casey Davenport)

Result

In conversation with Julien, both Kim and Casey indicated that they would like to reschedule

this discussion/demonstration for a later time, since neither is teaching a Python-related course in the Fall 2020 semester.

Action 5

Please help contribute to an appropriate citation for Mike Zuranski by sharing your suggestions with Doug Dirks (Committee)

Result

A citation for the DeSouza award incorporating committee suggestions was created and circulated.

Action 6

Share with Doug Dirks experiences moving to online format for blog posts (Committee)

Result

This is an ongoing request, to be revisited as the semester progresses.

Prepared *August 2020*

Status Report: AWIPS and GEMPAK

April 2020 - September 2020

Shay Carter, Tiffany Meyer

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. How can we help you and your students? What new functionality or products would you like to see incorporated into Unidata's AWIPS?
2. What has frustrated you most about Unidata's AWIPS?

Activities Since the Last Status Report

Staffing Changes

Tiffany Meyer joined Unidata as the Lead AWIPS Developer in May 2020 bringing extensive AWIPS knowledge and valuable AWIPS contacts from previous positions.

AWIPS

Unidata's Jetstream production EDEX server continues to serve real-time AWIPS data to CAVE clients and through 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.

Since the last update, several EDEX changes have been tested on our backup server and applied to our production server. One of the most significant changes dramatically reduced some of the data decoding latencies. By concatenating grib messages periodically, we were able to reduce the number of files being ingested, without losing any of the data. This led to around a 50% reduction in grib latencies. This same approach was applied to bufr files for a model sounding data set and the latencies of the regular ingest decreased by 80-90%. Reducing EDEX latencies is important because real-time data is the most valuable to our users and if that data takes an hour to process, it is no longer real-time.

Recently, GFS released an update for Vertical Velocity and Total Precipitation. Visualization and parameter changes were made to our EDEX, which allowed users to view the updated

menus and data in CAVE without requiring a new CAVE install. More information about this can be found on our website at https://www.unidata.ucar.edu/blogs/news/entry/awips-updates-new-gfs-fv3.

Additional radar data was recently enabled to our EDEX. Products with the header starting with "SDUS6" had previously been omitted from the NEXRAD3 LDM data feed. These products were found on the HDS feed, and small alterations to the ldmd.conf, pqact.conf, and radar distribution files were made to ingest this data into AWIPS. These products include Spectrum Width, Hail Index, and Tornado Vortex Signature.

Our entire AWIPS documentation website (http://unidata.github.io/awips2/) has been reviewed and significant portions have been rewritten to reflect the current state of our AWIPS distribution. Outdated content was removed, and updated screenshots, descriptions, and a new WarnGen video tutorial were added. Moreover, a new page was created to address commonly known problems and detailed instructions for their solutions.

A follow up survey was sent out in these past few months that focused specifically on what datasets/features AWIPS users currently use and would like to see implemented in the future. The survey has received 40 responses so far, which is about 33% of our awips2-users mailing list. We have gathered good feedback and will try and use the survey results as guidance for expanding the data we offer in our cloud EDEX.

Progress has been made with our Elasticsearch/Kibana web service. We have successfully built a dashboard that monitors the disk usage of the /awips2 mount on all four EDEX machines. This is a crucial first step in being able to proactively detect an upcoming EDEX failure and correct it before it ever reaches our users. We are aiming to expand our Elasticsearch and Kibana functionality and possibly packaging it for our users who have their own EDEX server to use. It has the potential to be a very valuable EDEX monitoring tool, and since it's freely available, it might be a nice addition to offer our community.

Software Releases

No AWIPS software releases have been made since the last status update. We've begun investigating the development, build and release process that Michael used and are hoping to put out at least one updated release of version 18 before the end of this year. That release would include several EDEX configuration changes, small CAVE updates, and possibly new datasets as well. In the more extended future, we hope to be able to merge our Unidata AWIPS back with one of Raytheon's more recent builds (most likely version 20).

GEMPAK/NAWIPS

GEMPAK 7.5.1, released in February 2019, is the most recent Unidata release: https://github.com/Unidata/gempak/releases/tag/7.5.1

At this time, GEMPAK development has ceased, and support through Unidata is very limited. We are in contact with Unidata community GEMPAK experts who have volunteered to help with questions/problems as their time allows and hope that others in the community will be able to provide additional guidance to issues that arise.

Activities Ongoing/In-Progress

AWIPS development activities are currently in a state of flux. Currently the following activities are in progress:

- The AWIPS team is expanding the Elasticsearch and Kibana capabilities to build a thorough reliable metrics and monitoring tool.
- 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 is familiarizing themself with our AWIPS Github repository and starting to understand the differences of our distribution and how the deployment process works.

Future Activities

Future plans are currently in a state of flux. The AWIPS team is working on understanding the details and differences of our AWIPS distribution and hopes to have the ability to make release updates soon. Backend changes to our cloud EDEX have been in progress and we plan to continue to optimize and expand its capabilities.

Metrics

Downloads March 2020 - July 2020

AWIPS downloads: 2,584
GEMPAK downloads: 5,221

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**
   Both AWIPS and GEMPAK are freely available, and both incorporate LDM/IDD technology for accessing geoscience data.

3. **Supporting People**
   At this juncture, we are providing a limited amount of technical support for the community, and encouraging community members to assist each other through Unidata-managed forums.

Prepared *August 2020*

Status Report: Cloud Computing Activities

April 2020 - September 2020

Sean Arms, 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 use?
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

Jetstream 2 Proposal Funded

The NSF-funded Jetstream cloud facility has been integral to Unidata's efforts to help university educators and researchers adapt their workflows to better use cloud computing techniques.

In early June 2020, NSF announced that a "Jetstream 2" project, a follow-on to the current Jetstream facility, had been awarded funding, with the Pervasive Technology Institute (PTI) at Indiana University as the project lead and the Texas Advanced Computing Center (TACC) as a primary partner. Both PTI and TACC were heavily involved in the original Jetstream project. Other Co-PIs are from the University of Hawai'i and Arizona State University.

Unidata uses Jetstream resources as part of the Unidata Science Gateway project, providing geoscientists with an ecosystem that includes large volumes of real-time data, data-proximate analysis and visualization tools, and Jupyter Notebook workflows for conducting teaching and research. Jetstream2 will enable Unidata to continue and expand its Science Gateway efforts in the Jetstream 2 cloud facility going forward. Unidata is also a funded collaborator in the Jetstream 2 project, providing tools, outreach, and advice over the life of the project.

NOAA Big Data Program Reforecasts from GEFSv12

Reforecasts for the upcoming Global Ensemble Forecast System v12 are now available in NOAA's Big Data program at Amazon Web Services. More documentation will be forthcoming, but I expect you can puzzle your way through it now, using it on AWS or downloading (e.g., curl, wget) to your own system.

s3://noaa-gefs-retrospective

URL: https://lnkd.in/gAZyMgf

Browser: https://lnkd.in/gZVDsXP

Webinars that discuss GEFSv12 performance and the reanalysis/reforecast are at https://ufscommunity.org/news/ufs-webinar-archive/
Happy diagnosticizing and postprocessing.
This announcement is from Tom Hamill, on behalf of the larger GEFSv12 team

JupyterHub Activities On Unidata Science Gateway

Unidata JupyterHub activities continue to advance since the last status report. We give details of our progress in this area below. These JupyterHubs are deployed In collaboration with the eXtreme Science and Engineering Discovery Environment (XSEDE) Extended Collaborative Support Services (ECSS) team and the Jetstream team at Indiana University (IU).

JupyterHub Servers for Online Instruction During COVID-19 Crisis and Summer Internship Programs

We have supported a number of semester-long classes, workshops and summer internship programs 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.

Since the beginning of 2020, the JupyterHub resources have been accessed by roughly 350 users many of whom are undergraduate students in the atmospheric sciences. This table summarizes Unidata JupyterHub servers in various settings.

	Nº of Users	Point of Contact
https://jupyterhub.unidata.ucar.edu/	10	Julien Chastang

Spring 2020 Workshop		
AMS 2020 Student Workshop	140	Sean Arms, Unidata
OU Spring Regional Workshop	20	Shawn Riley, OU School Meteorology
UNCC Spring Regional Workshop	10	Prof. Casey Davenport, Department of Geography & Earth Sciences
University North Dakota (workshop)	19	Prof. David Delene, Atmospheric Sciences
Spring Semester 2020		
Valparaiso University	59	Prof. Kevin Goebbert, Department of Geography and Meteorology
Valparaiso University 2	3	Prof. Kevin Goebbert, Department of Geography and Meteorology
Embry Riddle Aeronautical University	1	Prof. Curtis James, Applied Aviation Sciences
University of Hawaii, Manoa	17	Prof. James Potemra, School of Ocean and Earth Science
University of Northern Colorado	17	Prof. Wendilyn Flynn, Department of Earth and Atmospheric Sciences
Summer 2020		
UCAR SOARS Internship	13	Keith Maull, UCAR/UCP
UC Santa Barbara	4	Torin White, UCSB Library
UCSB Workshop(s)	26	Torin White, UCSB Library
OU (NSF REU students)	5	Shawn Riley, Ben Shenkel OU School Meteorology
CUNY Internship	8	Hannah Aizenman (former SIParCS), CS Grad Student

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.

Publications

Bulletin of the American Meteorological Society Publication

S. Arms, J. Chastang, M. Grover, J. Thielen, M. Wilson, and D. Dirks. Introducing students to

scientific python for atmospheric science. *Bulletin of the American Meteorological Society*, 2020.

Gateways 2020

A. Zonca, R. P. Signell, J. Chastang, J. Fischer, J. M. Lowe, and R. S. Sinkovits. Deploy kubernetes and jupyterhub on xsede jetstream. In *Gateways 2020 Proceedings*, 2020. In press.

THREDDS Activities on Unidata Science Gateway

We recently deployed two special purpose THREDDS servers on the Science Gateway. The first THREDDS server was deployed on behalf of Angelie Nieves-Jimenez from the Universidad de Puerto Rico, a UCAR SOARS Protégé. For her SOARS internship, Angelie required access to 1TB of HWRF data for analysis from a JupyterHub environment (which the Science Gateway also provided). We moved the data from Cheyenne to a THREDDS server so that she could access her data in a data-proximate manner from her computational notebooks. This effort allowed her to complete her summer internship project.

The second THREDDS server was deployed for the University of Tennessee, Knoxville. In this case, users discovered a bug in the NetCDF Subset Service in the beta release of the TDS 5. This bug does not impact version 4.6, so we deployed a one-off server for them to access CORDEX data to compare with their previous results. Thus far, this has involved the transfer of 2.5 TB of NA-CORDEX output to JetStream.

The rapid deployment of these THREDDS servers builds upon the earlier containerization work. They demonstrate another aspect of the Science Gateway where we can quickly stage resources for specific aims related to research projects of community members.

Jetstream Grant Renewal

Unidata successfully requested and obtained a renewal to our Jetstream allocation worth $80,000:

- Awarded Resources: IU/TACC (Jetstream): 4,000,000.0 SUs
- U/TACC Storage (Jetstream Storage): 40,000.0 GB
- The Science Gateways Community Institute (SGCI) support
- XSEDE Extended Collaborative Support Services (ECSS)

Unidata Science Gateway Digital Object Identifier (DOI)

Because of increased use of Unidata Science Gateway resources, we have obtained a DOI: doi:10.5065/688s-2w73 so that users of the gateway can cite it in their publications.

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 and THREDDS Data Server (TDS) THREDDS Docker software are being employed to deliver these data.
- 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 .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.
- Unidata continues to get requests from other UCAR/NCAR groups, to partner and lend assistance in cloud computing, especially in the AWS cloud.

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 sys admin group to ensure that our containers are configured as securely as possible.
- Collaborated with Sean Arms for the deployment of TDS production and beta releases as new versions of the TDS are released.

Product Generation for IDD

For the past four years, Unidata generated products for the IDD, FNEXRAD and UNIWISC data streams have been created by a VM hosted in the Amazon cloud. This production generation has been proceeding very 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.

Recently, we have added a full backup EDEX system, which includes a main EDEX machine and dedicated radar machine (designed in the distributed EDEX architecture). This allows us to have a backup to fall upon if anything goes wrong with our production system. It also provides a reliable testbed for enhancements and improvements without affecting our live system directly. We have successfully used the backup EDEX servers to test out bug fixes, optimizations, and dataset changes in the past few months. The backup has allowed us to vet our solutions for a significant period of time (days or weeks in some cases) before migrating the changes to the production server. This has also reduced the amount of downtime for our production server because we can iron out these processes on the backend server first.

Lastly, with the passing of Michael James, we have been working with the Indiana University Jetstream team to understand and recover the work that was left behind by Michael. This investigation involves examining and trying to gain access to the EDEX VMs that Michael had been working on. This continues to be an ongoing process because of all the machines (physical and virtual) and data that Michael left behind.

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 have been working with the Unidata system administrator group to ensure that our web-facing technologies on Jetstream adhere to the latest security standards. This work involves such tasks as ensuring we are employing HTTPS , keeping cipher lists up-to-date, 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.

New Activities

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

Fall 2020 JupyterHub Servers

Fall 2020		
University of Oklahoma		Shawn Riley, Ben Shenkel OU School Meteorology
Farallon Institute		Dr. Chelle Gentemann, Senior Scientist, Farallon Institute (& UW, Physics)
U of North Dakota		Dr. Aaron Kennedy Assoc Prof Dept of Atmos Sciences U of North Dakota
Valparaiso University		Prof. Kevin Goebbert, Department of Geography and Meteorology
Southern Arkansas University		Keith Maull (UCAR/UCP)

Forthcoming Virtual Conference Attendance

- Gateways 2020
- Possibly AMS Annual Meeting 2021

AWS Public Dataset For NEXRAD Level 3

We were approached by AWS's public dataset program about NEXRAD Level 3 data. We have started uploading all NEXRAD level 3 data to the unidata-nexrad-level3 S3 bucket; this bucket is intended to store all data indefinitely, providing a full archive of data going forward.

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). We are currently experimenting with deploying a Dask Gateway on Jetstream to achieve a Pangeo-like environment on the Unidata Science Gateway.

Relevant Metrics

Github Statistics

	Watches	Stars	Forks	Open Issues	Closed Issues	Open PRs	Closed PRs
sience-gateway	5	9	7	9	148	2	414
tomcat-docker	8	35	38	2	35	0	62
thredds-docker	10	16	17	3	104	0	138
ramadda-docker	3	0	1	1	10	0	18
ldm-docker	7	10	9	0	31	0	45
tdm-docker	4	1	5	1	9	0	13

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 *August 2020*

Status Report: Community Services

April 2020 - September 2020

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:

- Unidata Program Center Welcomes Drew Camron
- Bringing Back the NIU Weather Server
- Meet Unidata's 2020 Summer Interns
- Unidata Program Center Welcomes Tiffany Meyer
- 2020 Community Equipment Awards
- Summer 2020 Unidata Interns Wrap Up Their Projects
- Offer: Unidata Science Gateway JupyterHub Resources Available for Fall 2020 Courses
- Software release information
- Community job postings
- Community meetings and other announcements

Dependencies, challenges, problems, and risks include:

- Finding community members willing to contribute stories (or story ideas) for the blog is an ongoing challenge. 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 and SACNAS

- Engagement with CUAHSI to support the NFIE and WRF-Hydro at the NWC.
- Continue to serve on the CUAHSI HIS standing committee.
- We continue to update Unidata's social media channels (Facebook, Twitter, Google+).
- We continue to publish short videos/screencasts on the Unidata YouTube channel.
- 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.

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 allows
- Engage other UCAR/NCAR divisions regarding Unidata software use i.e. CESM/IDV
- Active participation in CUAHSI HIS (Hydrologic Information System)
- 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

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
- Seek additional opportunities to engage and listen to the community

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:

- 50,067 unique pageviews (44,873 in previous period)
- 7.6% of total unique pageviews (8.0% in previous period)

Top community pages

1. All blog pages
   38443 unique pageviews (34211 in previous period)
   77% of total community pageviews (76% in previous period)
2. www.unidata.ucar.edu/community
   4682 unique pageviews (5292 in previous period)
   9% of total community pageviews (12% in previous period)
3. www.unidata.ucar.edu/events
   3438 unique pageviews (2790 in previous period)
   7% of total community pageviews (6% in previous period)
4. www.unidata.ucar.edu/about
   2992 unique pageviews (2175 in previous period)
   6% of total community pageviews (5% in previous period)

Social media statistics, August 17, 2020

1. # of Twitter followers: 1386 (up from 1280)
2. # of Facebook followers: 846 (up from 817)

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 *August 2020*

Status Report: Unidata Community Equipment Awards

Sponsored by the National Science Foundation

April 2020 - September 2020

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Possible theme(s) for the 2021 Unidata Community Equipment Awards;
2. Please consider volunteering to serve on the 2021 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 November 19, 2019 with a March 20, 2020 submission deadline. The Review Panel met virtually on April 1 and recommended that the following proposals be funded:

- College of Charleston, Dr. Gabriel J. Williams, Jr., "Acquisition of AWIPS II CAVE Client Computing Infrastructure at the College of Charleston" proposal
- Mississippi State University, Kimberly M. Wood, ""Advancing Weather Visualization at Mississippi State University through a Dedicated AWIPS Server" proposal
- Texas A&M University, Dr. Christopher J. Nowotarski, "Building a Real-Time Running Archive EDEX Server for Meteorology Instruction" proposal
- University of Buenos Aires, Moira Doyle, "Upgrading LDM server at DAOC University of Buenos Aires" proposal
- University of Hawaii at Manoa, Jennifer D. Small Griswold and Steven Businger, "Enhancing Education in Atmospheric Sciences using AWIPS II and CAVE at the University of Hawaii at Manoa" proposal

Congratulations to all of the recipients and a special thank you to the Review Panel and the NSF for making the Unidata Equipment Awards program possible.

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 *August 2020*

Status Report: GOES-R Series Satellites

April 2020 - September 2020

Mike Schmidt, Tom Yoksas

Questions for Committee Members

- A reorganization of the content of the LDM/IDD NIMAGE was finalized and went live in June 2019. A reorganized **UNIWISC** datastream went live on March 2; details are included below

  Given the datastream reorganizations, 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 in the process of updating ADDE remote data serving of GOES-16/17 NOAAPort ABI imagery by the deployment of the latest McIDAS-X release, v2020

  By any measure, the 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 the 4.5m satellite dish at the NCAR Mesa Lab. An outcome of the discussions 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.

GOES-16 ingest on the Foothills Lab 2 (FL-2) satellite dish has been working well except during periods where maintenance is being done in the tree-lined plaza between FL-1 and FL-2. The worst interference is experienced when cleaning crews use gas-powered leaf blowers in the plaza. A secondary source of "interference" (signal degradation is a better description) is the trees that fill the plaza. As expected, signal levels and quality improve when the leaves fall off of the trees during the fall, and worsen when the leaves return in the spring. In the long term, the trees in the plaza will be removed (by UCAR) to facilitate construction to resolve drainage issues. Afterwards, some trees may be restored to the area, but we expect they will be sized and placed to avoid future problems. In the interim, we have been preparing to install a 3.8m mesh dish (from NCAR/EOL) on the western satellite pad at the Mesa Lab with the expectation that this location will be shielded from the TI problems that plagued our GOES-16 ingest.

- 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

- Investigate additions to the IDD CONDUIT data stream that would be useful for creation of new GOES-16/17 based Level 2 products

Future Activities

NOAAPort SBN

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.

The volume of data available in the **NIMAGE** and **UNIWISC** 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

VALUE-ADDE Products

Texas Tech University (Eric Bruning) is 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 added these Level 2 products to the NIMAGE IDD datastream at the same time that we added the NOAAPort-delivered Level 2 products. The GLM Level 2 products are directly usable by all of the analysis and display packages that we make available with the exception of GEMPAK.

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

Establish a test bed for the creation of Level 2 (L2) products from GOES-16/17 imagery, model output and observational data.

The objective is 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 (12 GB/hour ave and 16 GB/hour max; this is the most voluminous IDD feed!) can be seen in the real-time statistics plot from our GOES-R ingest machine:

  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 *August 21, 2020*

Status Report: Internet Data Distribution

April 2020 - September 2020

Steve Emmerson, 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, 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.

Activities Since the Last Status Report

Internet Data Distribution (IDD)

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

```
20200821

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume  98013.960 M bytes/hour
Average hourly volume  61788.411 M bytes/hour

Average products per hour     450001 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)   number/hour
FSL2                    15025.704    [ 24.318%]    25292.602    21111.667
SATELLITE               11844.822    [ 19.170%]    19623.624     5189.333
CONDUIT                  8954.141    [ 14.492%]    31000.958    81462.262
NEXRAD2                  8280.869    [ 13.402%]    11405.511   102319.119
NGRID                    7498.523    [ 12.136%]    12970.609    57545.929
NIMAGE                   5116.299    [  8.280%]     8965.813     4839.929
NEXRAD3                  2439.584    [  3.948%]     3184.550   100656.714
HDS                      1305.791    [  2.113%]     1905.559    34226.667
UNIWISC                   734.078    [  1.188%]     1099.757      719.286
GEM                       218.244    [  0.353%]     2423.627     1288.405
NOTHER                    173.277    [  0.280%]      875.763       30.357
FNEXRAD                   105.698    [  0.171%]      156.818       86.619
```

```
IDS|DDPLUS                85.652     [  0.139%]      118.871     40188.452
LIGHTNING                  5.629     [  0.009%]       11.499       335.738
GPS                        0.100     [  0.000%]        1.233         0.714
```

New Data Distribution:

IDD CONDUIT feed:

The GEFS model is being upgraded and the 1 degree files will no longer be created. Half degree files are available. Examples of this data on para-nomads - https://para.nomads.ncep.noaa.gov/ Unidata has requested a test feed that contains the new GEFS model output so that the impact on data volume can be assessed both internally and at the community top level IDD relay sites, UW/AOS and Penn State.

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 facility in Cheyenne, WY.

The data volume seen in the **SATELLITE** (which is known as **DIFAX** in LDM distributions prior to v6.13.6) feed 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 0.8-1.1 GB/hr volume in 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.

- Unidata receives the NOAA/GSD experimental High Resolution Rapid Refresh (**HRRR**) grids (both 2D and 3D fields) in an LDM/IDD feed from NOAA/GSD and feed these products to a small number of university sites on **hrrr.unidata.ucar.edu** (which is also known as **lead.unidata.ucar.edu**).

- The NCEP operational HRRR is being served with other model output at:

  http://thredds-test.unidata.ucar.edu/thredds/catalog/idd/forecastModels.html
  (.xml for machine access)

- Other data sets we continue to explore with NOAA/GSD/ESRL are:
  - FIM
  - HIWPP
  - HRRRx

- 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 assists LSU/SRCC in the maintenance of their NOAAPort ingest capability. Activities have included providing a spare LNB to bringing their NOVRA S300N receiver to Boulder for testing, configuration, power supply replacement and routine monitoring of their data and distribution.

  During this reporting period 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.

- The NOAAPort-derived data streams (**HDS**, **IDS|DDPLUS**, **NGRID**, **NIMAGE**, **NEXRAD3** and **NOTHER**) are redundantly injected into the IDD at three geographically separate locations: UCAR/Unidata, UW/SSEC, and LSU/SRCC. 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 on March 2.

We continue to look for a fourth ingest site to increase robustness of the IDD distribution of NOAAPort derived data.

- Unidata's NOAAPort ingest package is bundled with current versions of the LDM. The current LDM release is v6.13.11. A new LDM distribution is being readied for release.

Relevant Metrics

- Approximately **577** machines at **222** 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 changes. We are not certain why this may be the case, but out best guess is that some sites do not keep their LDMs running all of the time.

  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.

  Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

- 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.

  Over the period from April through mid-July 2020 the average volume of LDM/IDD data flowing from the UCAR/NCAR network averaged around 5.2 Gbps (~56 TB/day), and peak rates reached 10.8 Gbps (which would be ~116 TB/day if the rate was sustained).

Our new cluster machines have 10 Gbps Ethernet capability.

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 *August 21, 2020*

Status Report: IDV with RAMADDA

April 2020 - September 2020

Yuan Ho, Julien Chastang

Areas for Committee Feedback

We have no questions at this time.

Activities Since the Last Status Report

IDV Release

IDV 5.7 has been released on February 5 of 2020.

IDV System Changes

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the latest stable IDV version (5.7) is the 4.6.15-SNAPSHOT . The prior version of netCDF-Java to be distributed with the IDV was 4.6.12. There have been many improvements and bug fixes in that range. The complete release notes for these versions can be found here. The nightly release version 5.7u1 has integrated with the netCDF-Java 5.4 distribution. There are many changes related to the new features and enhancements.

__IDV Certificates__

Java Webstart, Windows app and MacOS certificates have been renewed and will be valid until at least May 30, 2020 (MacOS certificate is valid until 2024). Moreover, as properly signing the IDV under these different environments can be an involved process, this information has thoroughly documented here.

IDV Display Changes

__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.

__Color Table__

New enhanced color tables have been added in this version, including AWIPS color tables and Satellite products customized color tables.

__ NWS Hazard Warning__

NWS Tornado, Severe Thunderstorm, Flash Flood and Special Marine Warning KML files were supported in this release. A new menu item has been added to access the official NWS Warnings site directly under Data > Special data list.

__ABI and AHI Formulas__

Added a series of formulas specific to ABI (Advanced Baseline Imager), AHI (Advanced Himawari Imager) and MSG (Meteosat-8 through 11)to create displays of RGB and band subtraction products. These formulas can be found in the Field Selector under Formulas > Satellite. Also, derived fields for these products will be automatically added to the Field Selector when the multi bands ABI data source is loaded that includes bands required to create the product.

__Grid Diagnostics Formulas__

Added a variety of formulas specific to gridded datasource to calculate Saturation Equivalent Potential Temperature, Heat Index, Helicity, Ellrod Index, LP Index, and Richardson number.

__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 transitioning 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.

IDV EarthCube Drilsdown Project

In collaboration with University of Miami Professor Brian Mapes, the project facilitates the use of data visualizations within curated, executable notebook documents, in the service of improving statistical algorithms for datarich geosciences. The project integrates Jupyter Notebooks as the medium for literate computing and documentation; Unidata's IDV for its powerful data integration, visualization, and algorithm design capabilities; and RAMADDA as a metadata-rich repository for the packages that are the objects of this workflow.

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

__2019 AGU Fall Meeting__

- Hydrology Model Water Resource Estimation with Advanced 3D Visualization and Analysis
- An update in Hydrology Model Development and Application

__2019/2020 IDV lectures at WRF tutorial workshop__

- Yuan delivered two IDV introduction lectures on two semi-annual WRF tutorial workshops.

Ongoing Activities

We plan to continue the following activities:

__NetCDF Java 5.0.0 integration__

The new version of netCDF-Java will be integrated with the IDV. There have been many improvements and bug fixes in this new netCDF java library.

__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 August 2020

Status Report: Information Technology

April 2020 - September 2020

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 *August 2020*

Status Report: LDM

April 2020 - September 2020

Steve Emmerson, Tom Yoksas, Mike Schmidt, 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:

- Product-Queue:
  - Corrected bug in function's return code
- pqact(1):
  - Corrected truncation of output file by the FILE action when both "-overwrite" and "-metadata" are specified
  - Demoted "`Deleting ... entry`" to DEBUG if it's not due to an error in order to save log file space
- Logging:
  - Documented internally how `log_refresh()` will cause the log file to be closed/opened on the next log message and not before
- ldmd(1):
  - Top-Level Server:
    - Eliminated superfluous error messages at process termination about the upstream LDM database not existing if the "`-n`" (do nothing) option is specified.
  - Upstream LDM:
    - Replaced use of `select()` with `poll()` for better error-control
    - Increased service timeout from 60 s to 120 s
- noaaportIngester(1):
  - Corrected discussion in the documentation of UDP packet reassembly and the setting of system parameter `net.ipv4.ipfrag_max_dist` to zero
  - Added `Vertical Velocity (Geometric)` to the NCEP GEMPAK table
  - Changed logging-level of messages for first fragment of GOES imagery from NOTE to INFO
- Misc:
  - Improved logic in ghostname() for determining the fully-qualified name of the local host, which is used by rtstats(1)
  - Removed lint found by 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

Dependencies, challenges, problems, and risks include:

The LDM is sometimes held responsible for decisions made by the NWS 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.

Mathew Lazarra (University of Wisconsin–Madison) is spearheading an effort to formally integrate the LDM into the Antarctic meteorological data-distribution network.

Multicast LDM (aka LDM-7)

The multicast LDM project is separately funded by CISE in NSF. The goal is 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 the new, semi-reliable multicast protocol developed jointly with the University of Virginia with the time-tested unicast capability of the current LDM.

This project ended April 1st, 2019, but a no-cost extension was approved because funding was delayed.

Progress has been made on the following:

- Fixed bugs, particularly in the mapping from multicast product-index to LDM product-signature and in the backstop mechanism for products missed by the multicast module
- More sites added to the test deployment
- The switch handling LDM7 traffic from UCAR was replaced. This switch was mistakenly limiting the outgoing rate of multicast packets to 47 Mbps for an unknown reason.

Dependencies, challenges, problems, and risks include:

- The amount of manual intervention required to maintain the multipoint VLAN.

Ongoing Activities

We plan to continue the following activities:

- Support and maintain the LDM

New Activities

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

- Add message-authentication to the data-packets multicast by LDM7 (Unidata time to be paid by a separate grant from NSF)
- Work with NSF's Polar Programs regarding integration of the LDM into the Antarctic meteorological data-distribution network

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 *August 2020*

Status Report: McIDAS

April 2020 - September 2020

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 UW/SSEC McIDAS-X/XCD v2020 updates into the Unidata MciDAS-X v2020 release.

Current Activities

- Unidata McIDAS version 2020 will be made available for release in early to mid-September.

  v2020 will include all SSEC versions up to and including the current McIDAS-X and -XCD releases, both of which are v2020.1. A planned v2020a release will contain updates to the GOES-R/S ADDE servers.

- 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 features 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.

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 over 43 TB/month since the since the late spring of 2019

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 will remain 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 will continue through Thursday, September 3, 2020 1600 UTC. GOES-15 was returned to service to augment surveillance during the Pacific hurricane season.

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 *August 21, 2020*

Status Report: netCDF

April 2020 - September 2020

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What use cases
2. 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?
3. How can we encourage more user testing of the release candidates we provide?

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 **130** open issues for netCDF-C, **56** open issues for netCDF-Fortran, and **34** 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:

- We have provided releases of the core C library, as well as the Fortran and C++ interfaces.
- The initial version of Zarr-based data storage into the netCDF-C master branch.
- Further enhancements to the netCDF-C documentation, modernization of the netCDF-Fortran documentation.
- Progress moving the netCDF Users Guide (NUG) into its own repository; the NUG will be released as a stand-alone project when the next netCDF-C release is made.
- An architecture roadmap is available describing how the netcdf-c library will support thread-safe operation in *nix* and Windows environments. The draft proposal is available as netcdf-c github issue #382.
- 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.
- We have presented on the netCDF-Zarr roadmap and plans to working groups at both NASA and NOAA.

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.

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 begun the technical work of adopting Zarr functionality in the core C library. This will allow for object-based storage (Amazon S3, etc). We expect to have the initial release in the next 1-2 months.

Status

- Metadata is being properly written and read
- The content data of variables is not yet supported.
- ncgen and ncdump -h work with some limitations (as shown)
- Much testing:
    - move and modify tests from the ncdump/ and the nc_test4/ directories to zarr_test/.

Support for pure zarr

- Requires producing simulated data when _ncz... is missing:

    - simulated _nczarr: get the zarr version from the root group, use library built-in value for nczarr_version.
- simulated _nczcontent: Assume we have a group whose key is e.g. /y/z/g.
- dims: All variables whose shape is, say, (m,n), create dimensions in the root group of form dim_m=m and dim_n=n.
- vars: collect values of X for all keys of form "/y/z/g/X/.zarray".
- grps: collect values of X for all keys of form "/y/z/g/X/.zgroup".
- simulated _nczvar - contains netcdf-4 specific information for a
- Variable
    - dimrefs: Using the shape of the variable to figure out the dim names
    - contiguous is always false
- simulated _nczattr: infer attribute type from the values of the
- Attribute
    - (process is somewhat complex and is similar to algorithm in ncgen)

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

- Release the first test version of netCDF with Zarr support (ncZarr).
- Release subsequent versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Continue modernizing 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:

- **910,000** for netCDF-3
- **860,000** for netCDF-4
- **1,070,000** for HDF5
- **85,100** for GRIB2

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

- **416** for netCDF-3
- **906** for netCDF-4
- **15,900** for HDF5
- **1,320** 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 *August 2020*

Status Report: Outreach to Underserved Communities

April 2020 - September 2020

Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We are requesting 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?

Activities Since the Last Status Report

Attended the NSBE Virtual Conference

- Primarily a job fair
- Much learned about virtual conferencing AND engaging this community
- This is not a good venue for this activity and recommend we do not repeat attendance

Engaged with Rising Voices

- Planned attendance to annual conference in Washington State (postponed COVID-19)
- Coordinating with other Rising Voices UCAR staff for active engagement
- Joined and attended multiple sessions of the Indigenous Peoples Climate Change Working Group **(IPCCWG)**
- Joined and attended multiple sessions of the Indigenous-**FEWSS**. Indigenous Food, Energy & Water Security and Sovereignty

Attended AGU MSI sessions

- Met with students and faculty from MSI's around the country with a focus on Colorado for local engagement. Created some very good connections with CSU
- Continue to meet and work with UCAR's broader MSI initiative

Wizard Outreach

- Created content for education modules
- Performed live for K-12 during normal hours via Zoom conferencing

- Preparing for doing Super Science Saturday on Nov 7th from our homes and UCAR/NCAR (if using chemicals)

Internships

- Active engagement in the SOARS program
    - Selection Committee and mentoring - completed for 2020

- SCIParCS
    - Mentoring of Obsmara Ulloa completed and a new IDV kiosk version was created

- Unidata Internship
    - Mentoring completed

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

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

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 *August 2020*

Status Report: Python

April 2020 - September 2020

Ryan May, Drew Camron, Sean Arms, Julien Chastang

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. When do you update your Python environment for your classrooms (e.g. 1 month before classes start)? How often do you update your personal environment?
2. What headaches or challenges have you encountered while trying to incorporate MetPy/Siphon in your classes or research?
3. Would it be helpful if the Python Training material were easier to incorporate into classroom curricula? Is there some way we can better organize, package, or grow the Python Training to facilitate this?

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. One challenge is to balance time dedicated to creation of training materials, workshop preparation, and logistics against time devoted to support and Python software development. We're looking at ways to restructure our Python training material to work better in a virtual setting; this is extremely relevant in a world with COVID but also helps us reach a broader audience than through in-person events alone.

In addition to the need to reorganize for a virtual setting, we are currently working on additional re-organization of the Python training site as we work to complete the merging of sites we began last fall. Current plans are to rework the site using Sphinx along with a new Unidata Sphinx theme. This would unify the look and feel of the site with that of the documentation for the Python projects. This will also pave the way for a better organization of the material in general.

Progress has been made on the following:

- A collaborative Earthcube project with CISL, CGD, and the University of Albany, focusing on building out a larger Python training portal with a cloud-deployable platform to execute the training materials, was awarded by NSF in August 2020.
- Ryan May and Drew Camron, together with Kevin Goebbert, submitted a proposal for a short course, an intermediate course on using MetPy for calculations and analysis, at

the 2021 AMS Annual Meeting. This proposal was accepted.
- 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). The latest developments have focused on updating and improving MetPy’s infrastructure:

- Switching to manage CI dependencies using dependabot. This has greatly helped us keep updates in other packages from causing failures and blocking us from merging other PRs.
- Moving to GitHub actions, which has shortened time on Windows builds, as well as other benefits.
- Switching to a new theme for documentation, which has an improved, modern visual style, as well as better similarity to other newly styled documentation for Unidata projects.
- Testing MetPy without optional dependencies, ensuring that it is possible to use MetPy without these packages.

For MetPy 1.0, the full release is planned for this fall. First, 1.0rc2 will be released along with a 1.0 upgrade guide. The goal is to give users some more warning as to what changes will be necessary when upgrading to 1.0. About one month later, the final 1.0 release will follow.

MetPy’s usage metrics continue to grow. During the last year, MetPy averaged almost 24k page views per month, a 45% increase over the previous year. Also, in the last year, 22 new scholarly publications have cited MetPy, a large increase over the total count of 19 previously. On social media, the MetPy twitter account has reached 1369 followers (17% growth in 6 months). We also completed the 2020 MetPy User Survey, which generated 55 responses. When asked “In your experience how easy is it to use MetPy for your various activities?”, 60% of respondents rated MetPy 4 or higher on a 5 point scale (with 1 being “not easy” and 5 being “very easy”). Other interesting results from the survey is that almost 40% of respondents have not used both MetPy and GEMPAK, and that the 3 most used features of MetPy are: SkewT, calculations, and interpolation.

Progress has been made on the following:

- MetPy 0.12.1 and 0.12.2 were released with bug fixes on top of 0.12.0, including fixes to work with CartoPy 0.18.
- MetPy 1.0rc2 will be released imminently, with 1.0 following a month after that.
- Work towards requirements of MetPy-related NSF awards

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 continues to see community contributions trickle

in. Our next priority will be to transfer over infrastructure improvements made to MetPy. Beyond that, development in Siphon will explore ways to facilitate access to cloud-based public datasets.

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. In the last six months, this has included 10 issue reports and 28 pull requests. 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.

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.
- Ryan May and Drew Camron attended SciPy 2020, held in a fully virtual setting.
- We also continue to actively engage with the xarray, conda-forge, and pint projects.

Ongoing Activities

We plan to continue the following activities:

- Unidata Python training workshop
- 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:

- Continue updating and reorganizing the Unidata Python Training site
- Offer a virtual Python training workshop
- Release MetPy 1.0 final

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

- Teach another short course on MetPy at AMS 2021
- Present annual update on Python libraries at AMS 2021

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
- Evaluate ways to improve MetPy performance and scalability using tools like Dask and Numba

Relevant Metrics

MetPy

- 96% test coverage
- Watchers: 58
- Downloads for the releases made in the last year (Conda + PyPI):
    - 0.11.0: 4351
    - 0.11.1: 12465
    - 0.12.0: 39491
    - 0.12.1: 33205
    - 1.0.0rc1: 469
- Since 1 March 2020
    - Active Issues: 114 (69 created, 29 closed)
    - Active PRs: 118 (81 created, 75 closed)
    - External Issue Activity: 40 opened, 111 comments
    - External PR Activity: 33 opened, 95 comments
    - Unique external contributors: 53
    - Stars: 85 (617 total)
    - Forks: 7 (247 total)
    - Commits: 218
- Since 1 September 2019
    - Active Issues: 239 (160 created, 114 closed)
    - Active PRs: 217 (178 created, 169 closed)
    - External Issue Activity: 83 opened, 259 comments
    - External PR Activity: 78 opened, 194 comments

    - Unique external contributors: 87
    - Stars: 158 (617 total)
    - Forks: 16 (247 total)
    - Commits: 623

Siphon

- 97% test coverage
- Watchers: 17
- Downloads for the last year (Conda + PyPI):
    - 0.8.0: 36212
- Since 1 March 2020
    - Active Issues: 4 (3 created, 1 closed)
    - Active PRs: 14 (3 created, 1 closed)
    - External Issue Activity: 3 opened, 5 comments
    - External PR Activity: 3 opened, 3 comments
    - Unique external contributors: 9
    - Stars: 12 (125 total)
    - Forks: 1 (48 total)
    - Commits: 1
- Since 1 September 2019
    - Active Issues: 23 (15 created, 6 closed)
    - Active PRs: 26 (14 created, 10 closed)
    - External Issue Activity: 6 opened, 7 comments
    - External PR Activity: 9 opened, 10 comments
    - Unique external contributors: 16
    - Stars: 28 (125 total)
    - Forks: 3 (48 total)
    - Commits: 21

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 *August 2020*

Status Report: Support

April 2020 - September 2020

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?

Activities Since the Last Status Report

Training

- Since 2018, the UPC has been focusing its in-person training efforts on regional workshops and short courses.
- Additional resources will be directed towards developing online 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

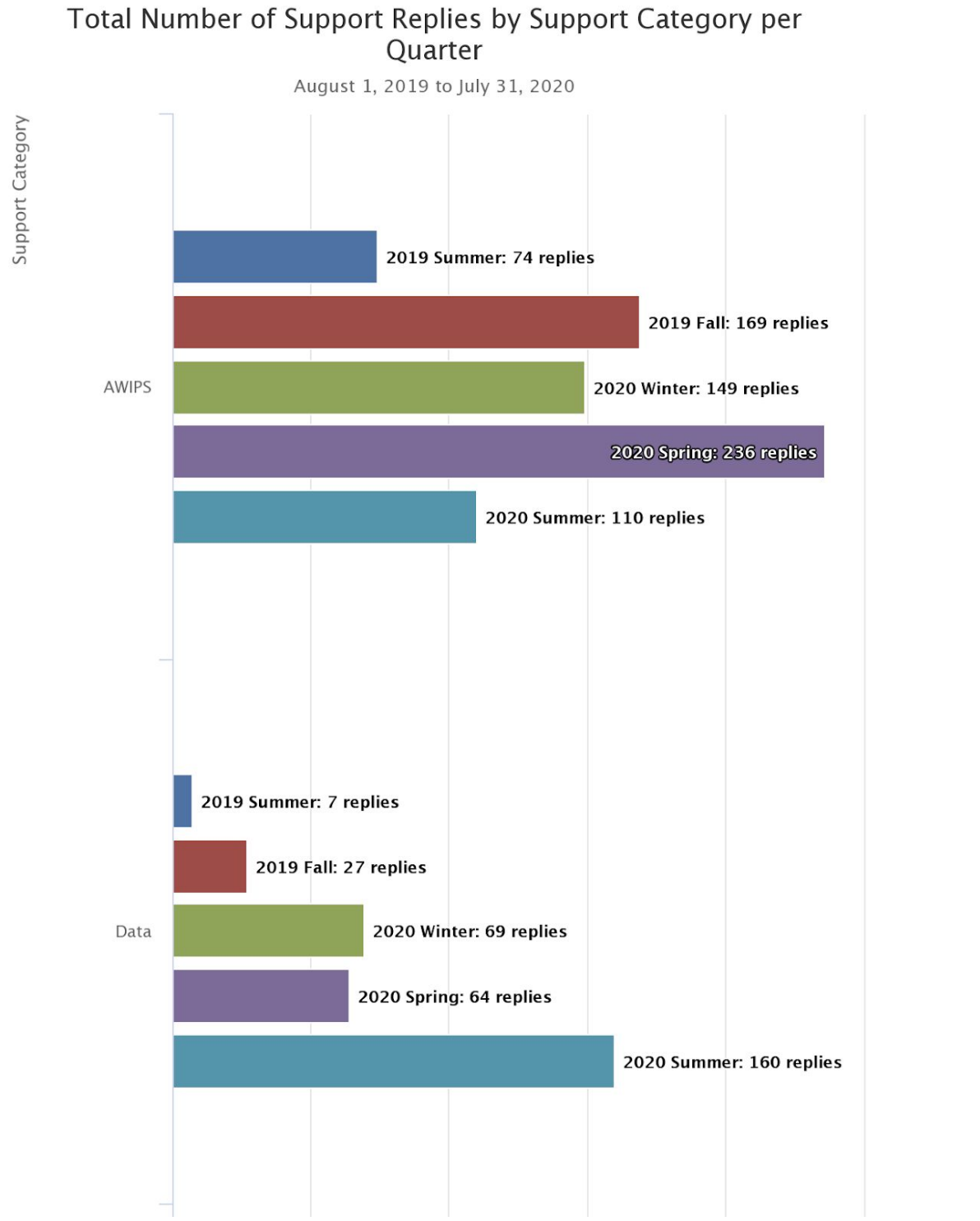
Since January 26, 2006 over 62278 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

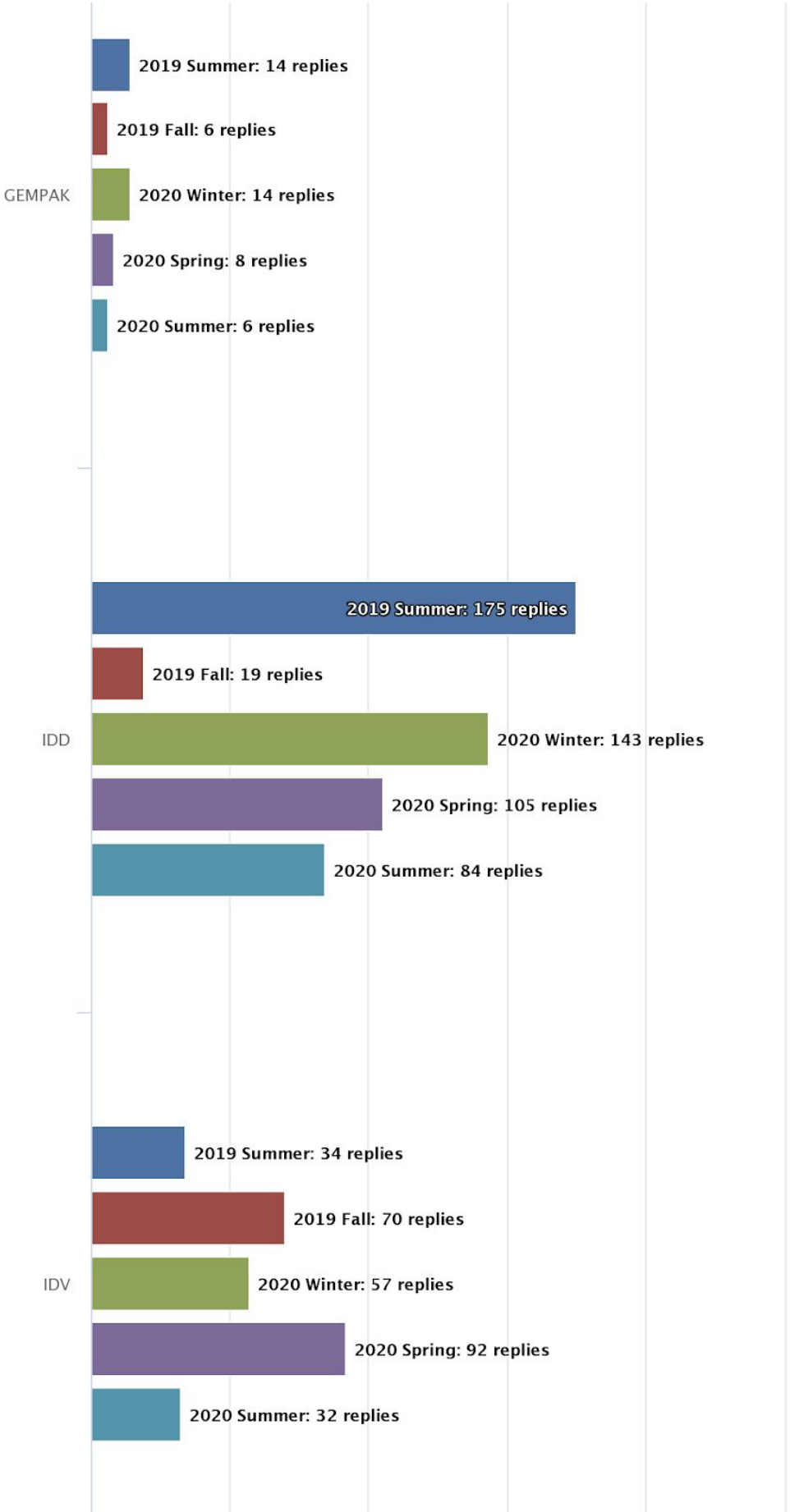
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

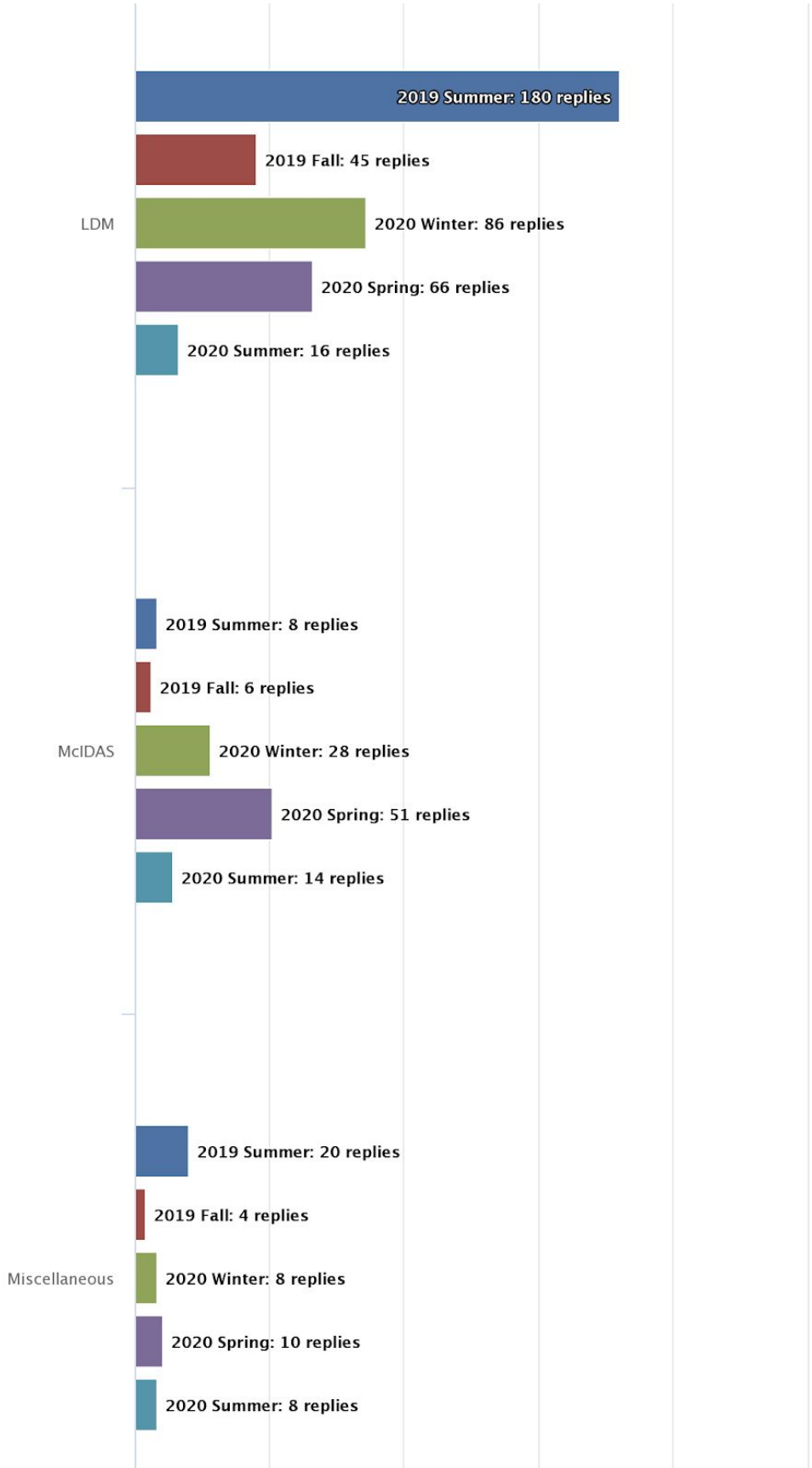
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 12 month period from August 1, 2019 until July 31, 2020.

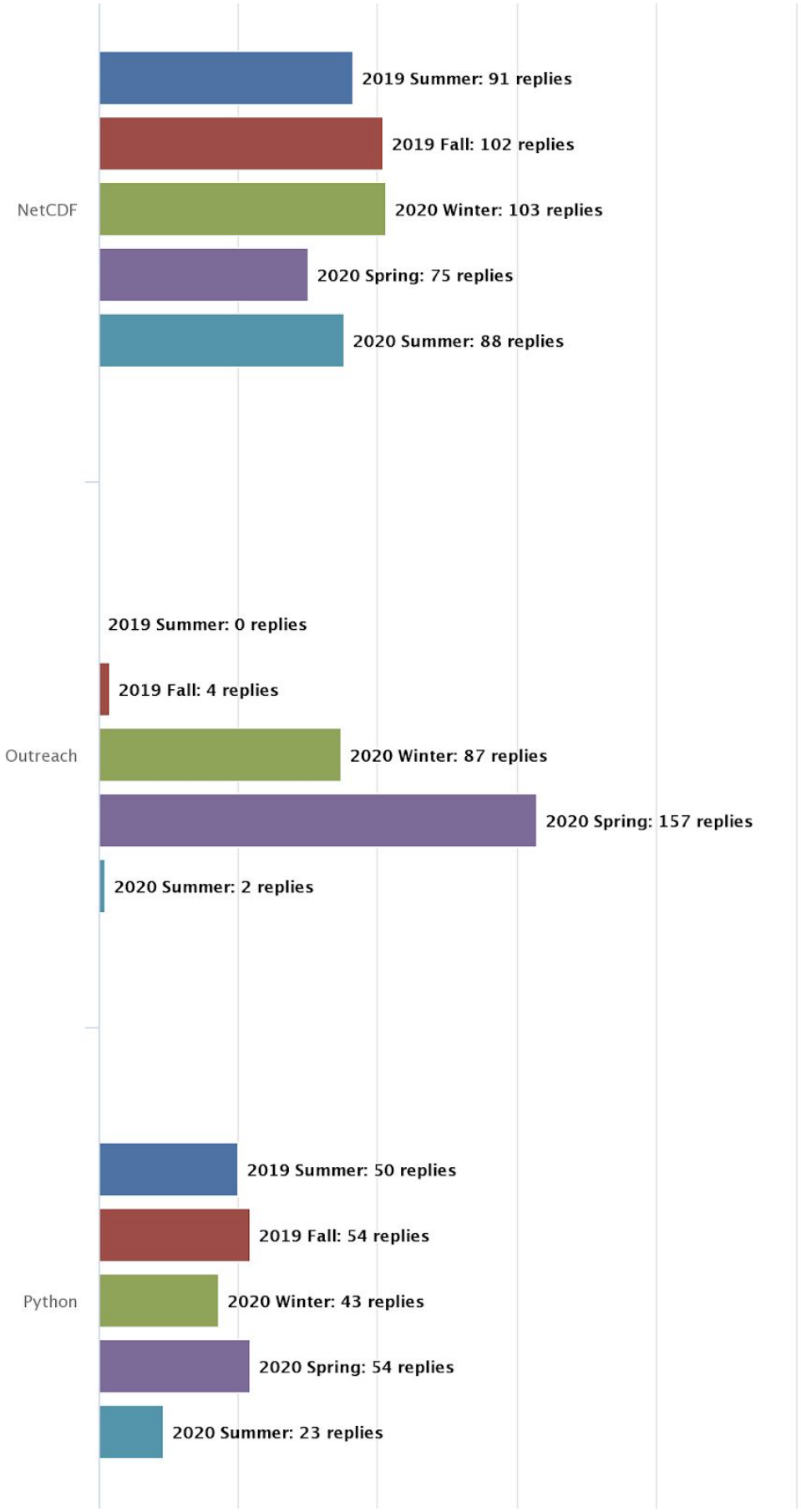
The quarters shown are defined as:

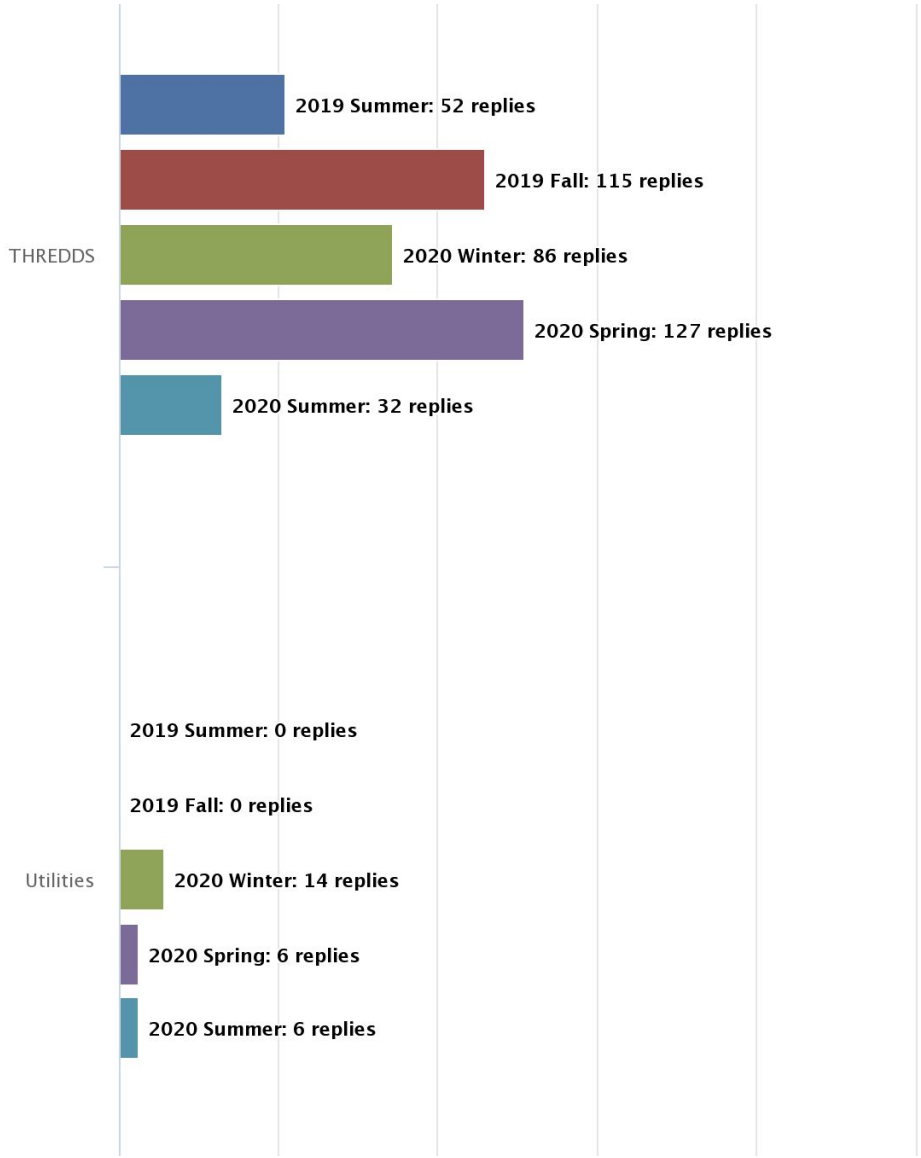
Winter:	**Spring:**	**Summer:**	**Fall:**
January, February, March	April, May, June	July, August, September	October, November, December




GEMPAK
2019 Summer: 14 replies
2019 Fall: 6 replies
2020 Winter: 14 replies
2020 Spring: 8 replies
2020 Summer: 6 replies
IDD
2019 Summer: 175 replies
2019 Fall: 19 replies
2020 Winter: 143 replies
2020 Spring: 105 replies
2020 Summer: 84 replies
IDV
2019 Summer: 34 replies
2019 Fall: 70 replies
2020 Winter: 57 replies
2020 Spring: 92 replies
2020 Summer: 32 replies


LDM
2019 Summer: 180 replies
2019 Fall: 45 replies
2020 Winter: 86 replies
2020 Spring: 66 replies
2020 Summer: 16 replies
McIDAS
2019 Summer: 8 replies
2019 Fall: 6 replies
2020 Winter: 28 replies
2020 Spring: 51 replies
2020 Summer: 14 replies
Miscellaneous
2019 Summer: 20 replies
2019 Fall: 4 replies
2020 Winter: 8 replies
2020 Spring: 10 replies
2020 Summer: 8 replies


NetCDF
2019 Summer: 91 replies
2019 Fall: 102 replies
2020 Winter: 103 replies
2020 Spring: 75 replies
2020 Summer: 88 replies
Outreach
2019 Summer: 0 replies
2019 Fall: 4 replies
2020 Winter: 87 replies
2020 Spring: 157 replies
2020 Summer: 2 replies
Python
2019 Summer: 50 replies
2019 Fall: 54 replies
2020 Winter: 43 replies
2020 Spring: 54 replies
2020 Summer: 23 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 *August 21, 2020*

Status Report: THREDDS

April 2020 - September 2020

Sean Arms, Jennifer Oxelson, Hailey Johnson, 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? Would you be interested in a classroom demonstration of how the THREDDS projects can be used to access data (note - Java programming **not** required - we love Python too!)?

Activities Since the Last Status Report

Staffing Changes

Please welcome the newest member of the THREDDS Development Team - Hailey Johnson! Hailey was a summer intern with Unidata in 2018, where she revamped the TDS html interface (bringing our html catalogs out of the 90s), as well as implemented the highly anticipated Jupyter Notebook service! Welcome Hailey!

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. An update regarding cloud efforts related to the TDS, including the popular Docker container effort, can be found in the Cloud Computing Activities Status Report.

Released netCDF-Java 5.3.3

- NetCDF-Java version 5.0.0 was released on 29 July 2019. NetCDF-Java version 5.3.3 was released on 18 June 2020.
- AWS S3 support was introduced in 5.3.1, so files readable by netCDF-Java can be read directly off of an S3 object store.
- The netCDF-Java build system was migrated from Gradle 3.5.4 to Gradle 6.5.1. With this move, we are now able to build and test netCDF-Java with Java 11+. While moving our build system three major versions ahead was not an easy task, the gain of Java 11 support is huge, as the end of 2020 is generally marked as end-of-life for Java 8 support by Oracle. Other JVM implementations, such as the JVM provided via the AdoptOpenJDK project, have different timelines set for Java 8 EOL.
- A critical caching bug was addressed in version 5.3.3, and as such, anyone using netCDF-Java should upgrade to 5.3.3 (or 4.6.15) asap. More infomration can be found in the release post on our blog at

https://www.unidata.ucar.edu/blogs/news/entry/tds-version-4-6-15.

- Prior to version 5.0.0, the netCDF-Java/CDM library and the THREDDS Data Server (TDS) have been built and released together. Starting with version 5.0.0, these two packages have been decoupled and now live in separate git repositories, allowing new features or bug fixes to be implemented in each package separately, and released independently. The codebase of netCDF-Java can be found at https://github.com/unidata/netcdf-java

Released TDS version 4.6.15 (Stable)

- TDS version 4.6.1 was released on 18 June 2020. A critical caching bug was fixed in this version, and as such anyone using the TDS to serve data should be using this version or else incorrect data could be returned by the Netcdf Subset Service (ncss). For more information, please see the release announcement at https://www.unidata.ucar.edu/blogs/news/entry/tds-version-4-6-15. If you notice a TDS server running a version prior to 4.6.15 (or 5.0.0-beta8), please let them know of this issue. We have reached out to several large sites, and while most upgraded immediately, some continue to run versions impacted by this bug.

TDS version 5.0.0

- We anticipate releasing a stable version of TDS 5.0.0 this fall.
- As with netCDF-Java, the Gradle build system for the TDS project was upgraded from version 3.5.4 to 6.5.1. There are a few places where the TDS codebase fails to build with Java 11, but these are pretty minor and will be addressed shortly, bringing full Java 11 compatibility to the THREDDS java projects.
- AWS S3 support added in netCDF-Java 5.3 has been extended to the TDS (current snapshot), and allows the TDS to serve individual files stored on AWS S3. CatalogScan, aggregation support is in the works.
- Starting with TDS v5.0.0-beta7, the TDS codebase can be found at https://github.com/unidata/tds

Documentation for netCDF-Java / TDS (Beta) v5

- Both netCDF-Java and the THREDDS Data Server documentation sets have changed significantly, and now use Jekyll and a markdown driven documentation theme for documentation generation. Previously, the documentation was a loose collection of Adobe Dreamweaver generated HTML. We hope the new documentation system provides for a consistent set of docs, and facilitates documentation contributions from users. The new netCDF-Java documentation can be found at https://docs.unidata.ucar.edu/netcdf-java/current/userguide/index.html, and the new THREDDS Data Server documentation can be found at https://docs.unidata.ucar.edu/tds/5.0/userguide/index.html.

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.

Progress has been made on the following:

- A lot of bedrock level work has been done on the codebase, including writing new tests, fixing bugs, and upgrading 3rd party libraries. While this type of work does not result in new features, it's the kind of ongoing work any web-based service needs to remain secure and usable into the future.
- Supporting user selection of an appropriate standard name from the CF Conventions has been particularly challenging. Work has been done to create a mind-map of the various standard names as a design prototype for building a standard name selection widget. While it is easy to stuff all of the available standard names into a dropdown list (>4000 names), it is not particularly user friendly, does not promote discoverability, and is certainly not performant. **Please let us know if you would be interested in testing this mind map** (which would essentially involve point-and-clicking your way through a tree to see how easy or difficult it is to find the CF standard names of variables you work with on a regular basis).

General dependencies, challenges, problems, and risks include:

- Calling all beta testers! The goal of beta testing TDS 5 is to ensure that the current capabilities of 4.6.x are working in the new version (and if some bugs get fixed in the process, even better!). Beta testing by our users is critical, and so far we have had several community members offer their help (special thanks to Rich Signell, Peter Pokrandt, Victor Gensini, the NCAR RDA, etc.!). We are currently on beta 8, and will be releasing beta 9 shortly. Beta 9 will be the final beta before the TDS 5 release candidate.

Ongoing Activities

We plan to continue the following activities:

- 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.
- Clearly define the public API of netCDF-Java.
    - Work on version 6 of netCDF-Java has started. Much of the previously deprecated code has been removed, and only the public API is used.
    - Many thanks to John Caron, who has been making the vast majority of this work happen. You likely remember John (or have heard his name) from his many dedicated years on staff at Unidata as the lead of the THREDDS projects since their inception in the late 90s. John left Unidata in 2015, and has remained a critical and active part of the community since then, even while holding a full time position at Google. John's talent, insights, and mentorship

have been greatly appreciated and cannot be overstated.

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.
- Thanks to Rich Signell, we, along with Axiom Data Science, submitted and were **awarded** a NOAA IOOS grant. The proposal was entitled "A Unified Framework for IOOS Model Data Access", and the goal 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 will fund a Undiata staff member at 0.5 FTE for two years, as well as two co-PIs at Axiom Data Science at a slightly lower level. This work **stragiticialy 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). The award was granted a 1 year no-cost extension.

New Activities

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

- netCDF-Java
    - Continue the deep look into dependencies and reduce as much as possible.
    - For the most part, this is a "baking" period for 5.x - focus on bug fixes, non-visibile library changes, few new features.
        - Exception - enhance support for S3 storage
    - Initial support for reading Zarr
    - Add aggregation support for files stored on AWS S3.

- TDS
    - Getting TDS v5.0 to a stable release (release candidate targeted for early fall 2020).

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

- netCDF-Java
    - Begin to modularize (Java Platform Module System) but maintain Java 8 compatibility.
    - Define public API and get 90%+ test coverage of it.
    - API breaks likely as we restructure our current artifacts

    - Initial support for writing Zarr and ncZarr
    - Complete command line tool creating a WRF intermediate file from a subsetted GRIB dataset.
    - Remove deprecated code
- TDS
    - Release TDS version 5 (Stable)
    - Create a TDS Registry
    - Implement option to create WRF intermediate files from GRIB datasets via TDS user interface. Support storage of pre-defined dataset variables for ease of WRF file recreation.

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

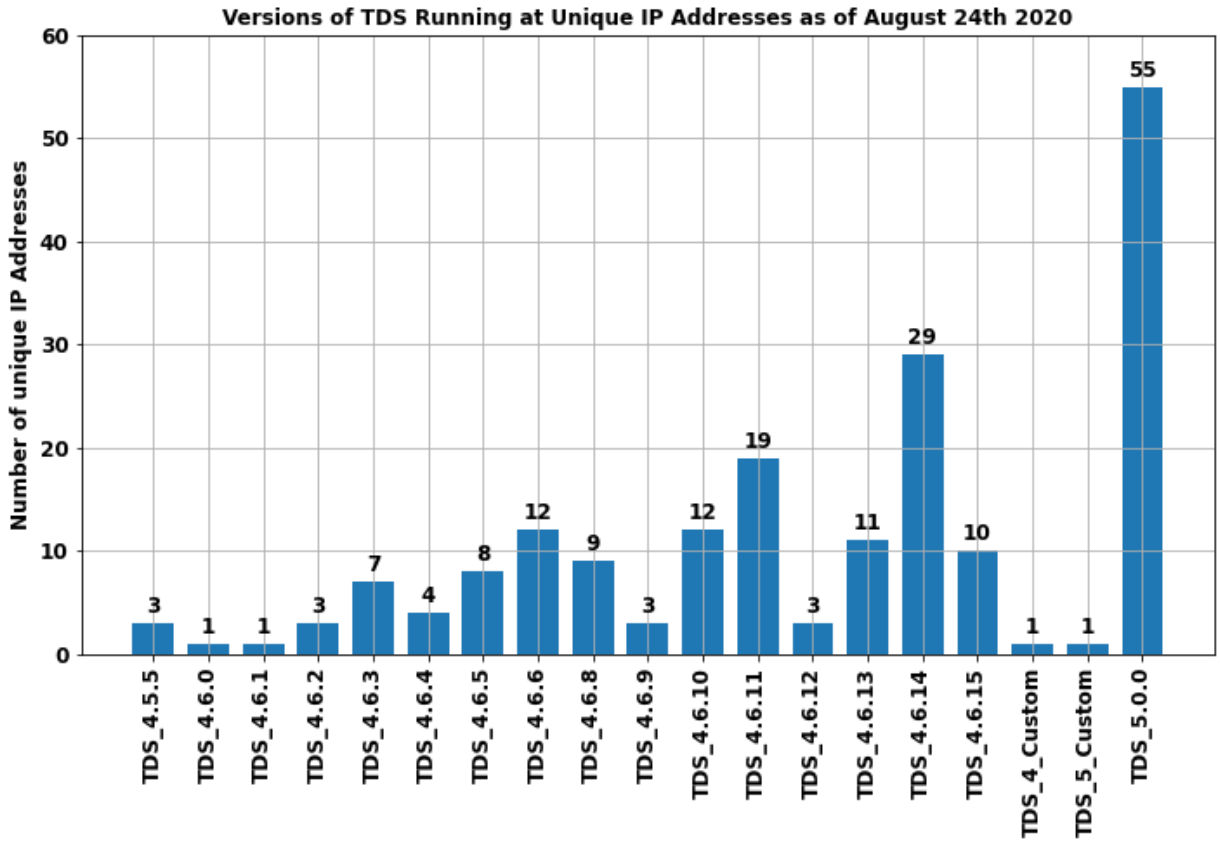
- netCDF-Java
    - Fully support Java 11 and the Java Platform Module System (end of Java 8 support)
    - Commit to semantic versioning
- TDS
    - Reduce dependency footprint
    - Modularize (Java Platform Module System)
    - Create a collection level update notification system
    - Create a collection level metadata search across TDSs
    - Look at re-architecting the TDS to “really” run on the cloud
        - TDS as a collection of autoscalable microservices vs monolithic web application. With our current resource levels, this is a stretch.

Relevant Metrics

NetCDF-Java

Last report, we showed the yearly downloads for netCDF-Java to be 36,897 (March 1, 2019 - February 29, 2020). The new 12 month download count for netCDF-Java is **49,181** (August 1, 2019 - July 31, 2020).

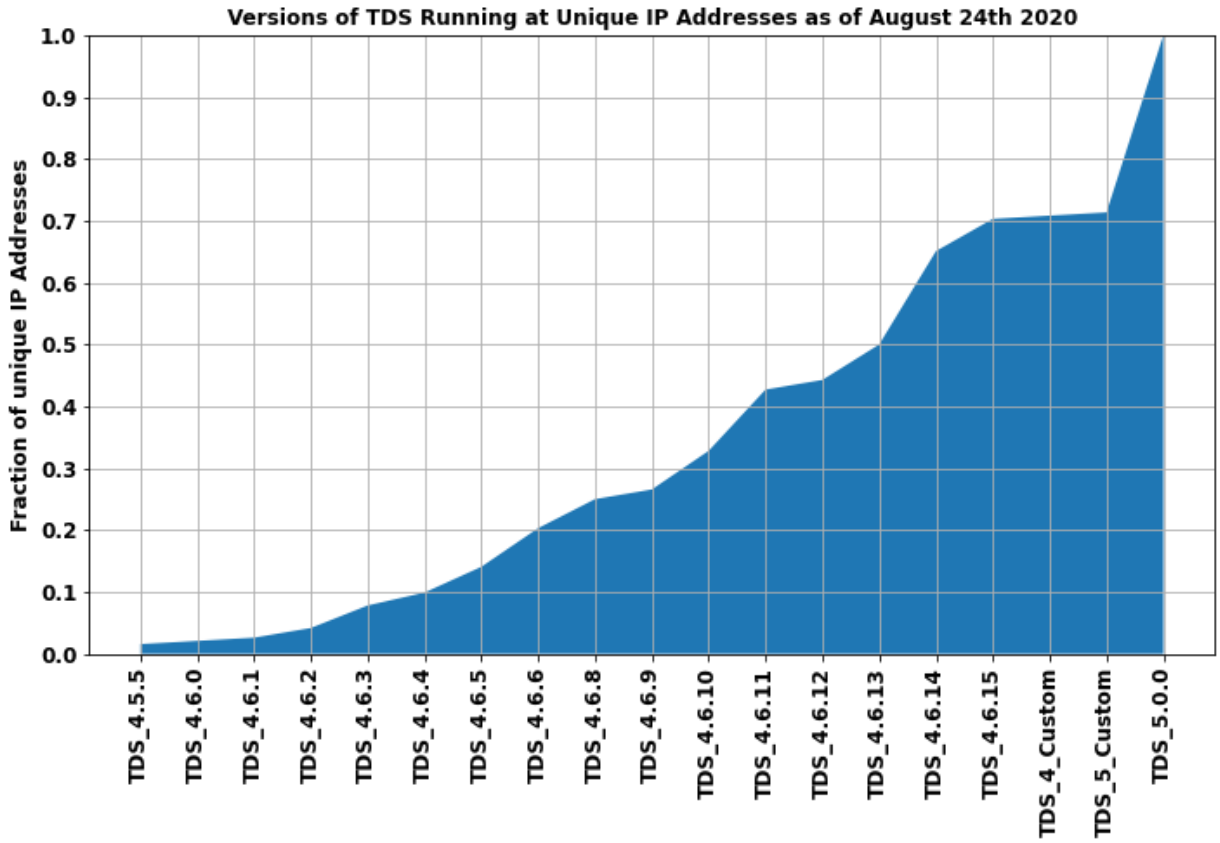
THREDDS Data Server



We've seen that over the past six month period of time, **3,041** unique IPs started up the TDS (March 2020 through August 2020), **96** of which are publicly accessible servers. Since we've started tracking these metrics (v4.5.3, August 26th, 2014), we've seen the TDS startup from **31,167** unique IP addresses. Overall, we currently have a total of **192** publically accessible TDSs running "in the wild" (up from our last report, which had **161**) . "Publicly accessible" means we could find a reachable URL using common url patterns (<server>/thredds/catalog.xml, <server>/thredds/catalog/catalog.xml).

For this plot above, the version includes betas and snapshots, not just the official release of that version, for presentation simplicity.

Note 1: the majority of the publicly accessible servers are running v4.6.10 or above (v4.6.15 was the most current release during this period, and was released on 18 June 2020). We now see that TDS v5 has become the dominant specific version running in the wild.



Furthermore, of the **192** publically accessible servers, **103** 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…). This is up from our last report, which showed **64** servers with updated contact information.

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 *August 2020*