Unidata Users Committee
Virtual Meeting Agenda

(All times are Mountain Daylight Time)

Thursday, April 2

9:00 - 9:10: Welcome, logistics, and prior action items (Kevin and Josh)

9:10 - 9:25: Questions for Consideration (Kevin)

9:25 - 9:55: Director's Report (Mohan)

9:55 - 10:55: Program Recommendations on Unidata Analysis and Visualization Services (Mohan and Committee)
(25 minute presentation and 35 minute discussion)

10:55 - 11:10: Discussion of DeSouza Nominees (Committee)

11:10 - 11:35: Discussion and response to questions on Status Reports (Staff and Committee)

11:35 - 11:50 AM MDT: Identify potential fall meeting dates and close virtual meeting (Josh)

Friday, April 3 (Demonstrations)

9:00 - 9:40: Demonstration of Science Gateway (Julien)

9:40 - 10:10: Demonstration of new netCDF-Java and TDS features (Sean)

10:10 - 10:40: Demonstration of new Metpy features (Ryan)

Status Report: Users Committee Actions

November 2019 - April 2020

Unidata Program Center Staff

Actions from the Previous Meeting

Action 1

Review Equipment Award announcement language and communication strategy (All Committee members)

Result

Updated language was distributed for committee member comment on November 13th and finalized after incorporating feedback on November 18th: Completed

Action 2

Consider Equipment Award topics for special consideration and submit to UPC (All Committee members)

Result

The request for new topics for special consideration was announced during the fall joint session with the window for suggestions closing prior to AGU/AMS: Completed

Action 3

Provide any comments or revisions to the DeSouza Award text within the next 2 weeks (All Committee members)

Result

Revisions received and new language used in this year's solicitation: Completed

Action 4

Send out a note to the Users Community for any folks that are extensive AWIPS users (Kevin G.)

Result

Message to AWIPS user community was delivered: Completed

Action 5

Full list of available products on Conduit by spring and any help with reliability (Carissa)

Result

This is an ongoing action that will continue until the next committee meeting: Pending

Action 6 (from Joint Committee meeting)

Email committee members Tuesday, October 8th about MSI document review (Josh Young)

Result

Completed: Action Completed

Action 7 (from Joint Committee meeting)

Review and comment on draft MSI Engagement Strategy (Committees)

Result

Comments received and accepted: Action Completed

Status Report: AWIPS and GEMPAK

September 2019 - February 2020

Shay Carter

Since the previous report Evan Polster was hired as a contractor to help work with Shay in an effort to continue providing the community with high-level support for the AWIPS package. There has been an ongoing search to try and hire someone to help replace the position left vacant by Michael James. As of early March Tiffany Meyer has accepted the position of Software Engineer IV at Unidata to work as the lead of the AWIPS team. She will be working remotely for a year or so and will eventually be moving to Colorado. She will join Unidata sometime in the May - June time frame.

Activities Since the Last Status Report

AWIPS

Unidata's Jetstream 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 two datasets were re-established in our EDEX. Both NUCAPS and PIREP were not available through our EDEX when I arrived here. NUCAPS data had a feed change which required a modification in the EDEX distribution file. PIREP support was not included in the base install of EDEX and had to be manually added and have processing turned on afterwards.

An update to our installation page for CAVE was made that allows Windows users to download a Virtual Machine (VM) and run CAVE through a Linux environment. This was added because several of our users have reported problems rendering certain datasets on their Windows version of CAVE. This is believed to be because of a Python dependency within some of CAVES rendering methods, and trouble with getting it to install correctly when CAVE is installed. The VM offers a quick workaround for this problem, but the Python issue is still being looked into.

Additionally, a brief user survey specifically for AWIPS use was sent out to the awips-users email list at Unidata. This survey received a 33% response rate and helped describe a little

more of where our user base is coming from and how they use AWIPS. This was done in an attempt to better understand our user community to provide the best support and development of our AWIPS package possible.

Software Releases

No AWIPS software releases have been made since the last status update. We're hoping that it might be possible to have a software release by the end of 2020, to update to AWIPS version 20 and skip 19. The timing of this release will depend on the ability to understand and persist the changes Michael had made to Unidata's version of AWIPS compared to the baseline version developed by Raytheon.

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

- Shay is familiarizing herself with all Unidata publically available documentation and training materials related to AWIPS.
- Shay has established a back up EDEX system (with both a main EDEX machine and radar-specific machine) on the Jetstream cloud and is in the process of implementing measurable metric monitors.
    - This will be used to help quantify improvement attempts for data latency issues.
    - It can also be used to help prioritize support based on data use from our community.
- Shay is responding to all AWIPS support questions from the community and striving to provide realistic solutions in a timely manner.
- Shay is familiarizing herself with our AWIPS Github repository and starting to understand the differences of our distribution.

Future Activities

Future plans are currently in a state of flux. We expect the situation to begin to resolve as Shay comes up to speed and Tiffany Meyer joins the Unidata AWIPS team.

Metrics

Downloads September 2019 - February 2020

AWIPS downloads: 2,125
GEMPAK downloads: 4,928

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 *March 2020*

Status Report: Cloud Computing Activities

September 2019 - March 2020

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Do you need a Unidata hosted JupyterHub for your classroom or workshop use?
2. What new cloud technologies are our community members using and investigating on their own initiatives?
3. 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

JupyterHub Activities On Unidata Science Gateway

Unidata JupyterHub activities have been increasing since the last status reports. 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

We worked with Kevin Goebbert at Valparaiso University and Shawn Riley at OU to set up JupyterHubs for online instruction during the COVID-19 crisis. Separately, in collaboration with Keith Maull (UCP), we deployed a JupyterHub for a data science class at Southern Arkansas University for the fall 2019 semester.

JupyterHub Servers for OU and UNCC Regional Workshops

We worked with our partners at OU and UNCC to set up JupyterHub for use during the OU and UNCC regional workshops. The objective here was to provide pre-built environments to have the instructors and students focus on the instructional material rather than installing software on their laptops.

JupyterHub Server for AMS 2020 Student Conference

Unidata hosted a Python workshop at the Annual Student Conference for the American Meteorological Society 2020 annual meeting. The goal of this workshop was to deliver an introduction to Python for the atmospheric sciences to students in 90 minutes. While Unidata took the lead in organizing the workshop, students taught the material -- a workshop for

students by students. 140 students attended. We provided pre-installed and pre-configured JupyterHubs for this workshop. In collaboration with Doug Dirks, and those who organized and presented at this workshop, we are in the process of submitting a workshop summary for publication in BAMS.

JupyterHub Demonstration Server

Unidata continues to enhance the Unidata JupyterHub demonstration server.

We have been working with Ben Schenkel (Research Scientist, University of Oklahoma, Cooperative Institute for Mesoscale Meteorological Studies) who has been providing us feedback for this JupyterHub server. He is directing his NSF REU students to use this solution because it requires no installation of local software.

We assisted Alex Davies organize a Python instructional group at the US Naval Academy. The group employed the JupyterHub demonstration server as part of their instruction. This effort was ultimately described in Unidata blog entry: Unidata Science Gateway JupyterHubs are Helping U.S. Naval Academy Faculty Learn Python.

At this point, this demonstration server requires an update. In order to have this happen, we will ask all users to save any critical material they have on the JupyterHub and we will rebuild it with more up-to-date software. We especially need to incorporate the recently revamped Unidata python-training project.

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.

Ongoing Activities

NOAA Big Data Project

- 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 IDV, 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.
- Received a community contribution for an enhanced start up script for the LDM docker container.
- Updated THREDDS Docker container in order to address a zlib version compatibility issue between netCDF and Godiva. Also updated the version of netCDF contained within this project.
- Finally added an open source license to all Docker projects.

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 xsede-jetstream 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 can test solutions and modifications on the backup system and assess their viability before migrating the changes to the production system.

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.

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.

New Activities

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

Forthcoming Conference Attendance

- PEARC 2020 Conference | July 26-30, Portland, OR
    - Submitted a paper for presentation in collaboration with Andrea Zonca (XSEDE/ECSS/SDSC)
- 2020 ESIP Summer Meeting | July 14-20, Burlington, VT
- Gateways 2020 | October 19-21, Bethesda, MD

Jetstream Grant Renewal

We must renew our Jetstream allocation with XSEDE. We are making good use of the present 2019-2020 allocation and we are on target to make complete use of our Jetstream allocation for this time period. We will ask for at least the same amount of resources and perhaps more to accommodate the growing number of JupyterHub servers. We will be putting forward our grant proposal to XSEDE by April 15.

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

XSEDE ECSS Jetstream JupyterHub Collaboration

We plan to continue our collaboration with Andrea Zonca (XSEDE ECSS, San Diego Supercomputing Center) to migrate from a Kube Spray to Magnum deployment for our JupyterHubs. This transition will allow for simpler workflows as well as giving us access to clusters that can automatically scale to add more cluster nodes as more users come online and remove nodes when they are no longer needed.

Relevant Metrics

Github Statistics

	Watches	Stars	Forks	Open Issues	Closed Issues	Open PRs	Closed PRs
xsede-jetstream	5	8	6	4	146	2	354
tomcat-docker	8	32	28	2	32	0	52
thredds-docker	10	14	16	1	104	0	136
ramadda-docker	3	0	1	1	10	0	18
ldm-docker	7	7	8	0	31	0	45
tdm-docker	4	2	5	1	9	0	12

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.

Prepared *March 2020*

Status Report: Community Services

November 2019 - April 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:

- Pete Pokrandt Receives 2019 DeSouza Award
- A Visit to the SACNAS Conference and the University of Hawai'i
- MetPy Project Brings GEMPAK Features to Python
- Unidata Staff at AMS 2020 Meeting
- AMS 2020 Conference Highlights from the Unidata Staff
- What's New in IDV 5.7
- Unidata Science Gateway JupyterHubs are Helping U.S. Naval Academy Faculty Learn Python
- 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

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:

- 44,873 unique pageviews (51,066 in previous period)
- 8.0% of total unique pageviews (10.5% in previous period)

Top community pages

1. All blog pages
   34211 unique pageviews (42256 in previous period)
   76% of total community pageviews (83% in previous period)

2. www.unidata.ucar.edu/community
   5292 unique pageviews (3112 in previous period)
   12% of total community pageviews (6% in previous period)

3. www.unidata.ucar.edu/events
   2790 unique pageviews (2856 in previous period)
   6% of total community pageviews (6% in previous period)

4. www.unidata.ucar.edu/about
   2175 unique pageviews (2395 in previous period)
   5% of total community pageviews (5% in previous period)

Social media statistics, March 13, 2020

1. # of Twitter followers: 1280 (up from 1132)

2. # of Facebook followers: 817 (up from 786)

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 *March 2020*

Status Report: Unidata Community Equipment Awards

Sponsored by the National Science Foundation

November 2019 - March 2020

Admin Group

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 will meet on April 1 before the Users Committee Meeting at the Unidata Program Center to finalize their review and to make a recommendation on funding.

Relevant Metrics

Since taking over the management and administration of the Equipment Awards program in 2003 on behalf of the NSF, Unidata has made 99 awards totaling over $1,250,000.

Prepared March 4, 2020

Status Report: GOES-R Series Satellites

November 2019 - April 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. 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

- Upgraded the IDD NIMAGE datastream by including the full content of NOAAPort-delivered GOES-16/17 ABI and other Level 2 imagery and products.

- Upgraded the IDD **UNIWISC** datastream by adding select wavelength channels (0.64um VIS, 6.2um WV and 10.3um IR) products for all coverages (CONUS/PACUS, FullDisk, Mesoscale-1 and Mesoscale-2) for GOES-16/17 imagery. The FullDisk images have reduced spatial resolution (nominal 4km) from the images that are available in the **SATELLITE** feed in which all GOES ReBroadcast (GRB) products are distributed. All products in the **UNIWISC** feed are distributed in PNG compressed AREA format that is fully supported by GEMPAK, the IDV, McIDAS-V and McIDAS-X and partially supported by AWIPS.

- Updating remote data serving of GOES-16/17 NOAAPort ABI imagery community data servers that we operate in UCAR.

  By any measure, the GOES-16/17 imagery is **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

- Canvas the community to learn more about GOES-16/17 data needs
- 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 have been made aware of a possible expansion of the NOAAPort SBN that would, if implemented, increase available bandwidth twofold. We are trying to keep on top of this possibility in order to minimize any effects it might have on the Unidata community.

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 anticipate that TTU will not be the only site creating value added products.

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.

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 *March 2020*

Status Report: Internet Data Distribution

November 2019 - April 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:

```
20200313

Data Volume Summary for lead.unidata.ucar.edu

Maximum hourly volume 125230.316 M bytes/hour
Average hourly volume  75252.268 M bytes/hour

Average products per hour     496635 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)            (M byte/hour)   number/hour
FSL2                    22931.217    [ 30.472%]    54443.397    34122.651
SATELLITE               14374.532    [ 19.102%]    20420.168     6451.651
CONDUIT                 11786.783    [ 15.663%]    33389.863   107130.651
NGRID                    9106.208    [ 12.101%]    13822.809    60820.977
NEXRAD2                  6161.061    [  8.187%]     8328.219    83258.442
NIMAGE                   5787.311    [  7.691%]     9001.989     5791.884
NEXRAD3                  2070.583    [  2.752%]     2806.330   104862.093
HDS                      1311.325    [  1.743%]     1748.879    39198.930
UNIWISC                   864.808    [  1.149%]     1102.187      886.814
GEM                       410.518    [  0.546%]     3344.875     2619.279
NOTHER                    210.973    [  0.280%]      733.793       40.860
IDS|DDPLUS                117.647    [  0.156%]      133.027    50835.302
```

```
FNEXRAD                 111.076    [  0.148%]      130.390      104.767
EXP                       6.622    [  0.009%]       13.719       99.953
LIGHTNING                 1.474    [  0.002%]        2.544      409.744
GPS                       0.129    [  0.000%]        1.315        1.070
```

New Data Distribution:

Very recently, a feed named FSL2 was added to the mix of feeds provided by the toplevel IDD relay clusters idd.unidata.ucar.edu and iddb.unidata.ucar.edu. This feed currently contains only Alaskan wind profiler data in BUFR format. We will add other profiler data we are currently receiving from NOAA/GSD in a test mode to the feed in the very near future.

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**.

- 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.

- The NOAAPort-derived data streams (**HDS**, **IDS|DDPLUS**, **NGRID**, **NIMAGE**, **NEXRAD3** and **NOTHER**) are being 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 **601** machines at **231** sites are running LDM-6 **and** reporting real-time statistics to the UPC.

  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.

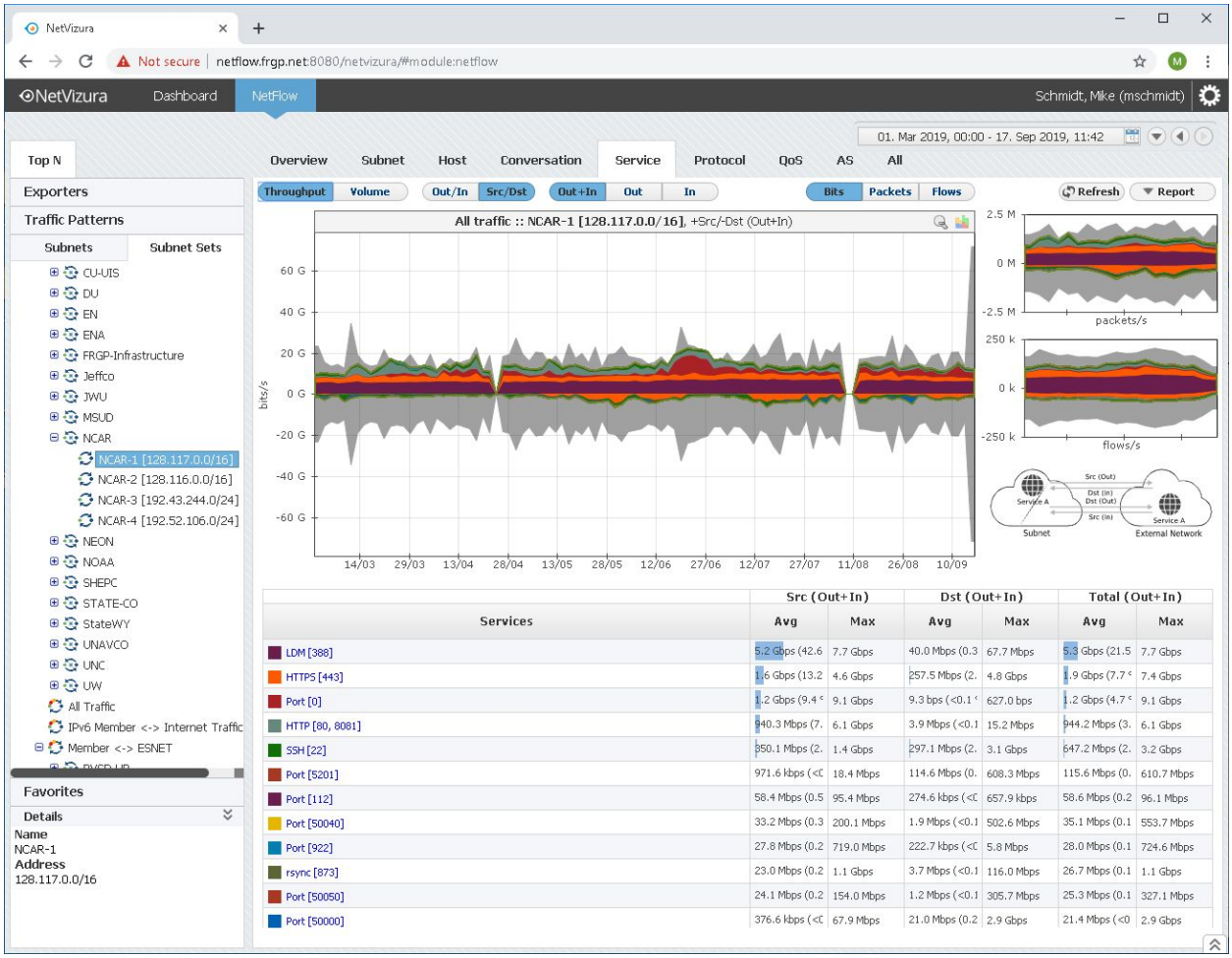
  Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

- 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 cluster, described in the June 2005 CommunitE-letter article Unidata's IDD Cluster, routinely relays data to more than 1076 downstream connections. It was moved to the NCAR/Wyoming Super Computing facility in Cheyenne, WY in late August.

  Over the period from March 11, 2019 through September 17, 2019 the average volume of LDM/IDD data flowing from the UCAR/NCAR network averaged around 5.2 Gbps (~54 TB/day), and peak rates reached 7.7 Gbps (which would be ~126 TB/day if the rate was sustained). The volume of data being distributed via the IDD from UCAR increased significantly with the augmentation of the **NIMAGE** feed in June. Updates to the **UNIWISC** feed and the distribution of the Alaskan wind profiler data in the **FSL2** feed in March further increased feed volumes.



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 *March 2020*

Status Report: IDV with RAMADDA

October 2019 - March 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 (5.6) 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.

__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

__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.

__CDM Profile data type__

Common Data Model profile feature data is a set of data points along a vertical line, including a collection of a time series of profile features at named locations. CDM Profile data can be loaded in the Data Choosers: RAOB > Soundings Local.

__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 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 the 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 workflows.

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 ~20 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 March 2020

Status Report: Information Technology

October 2019 - March 2020

Mike Schmidt, Matt Perna, & Jennifer Oxelson

Major Activities

****Cybersecurity Analysis**** -- UCAR has received the final report from the completed cybersecurity evaluation and assessment commissioned by the UCAR Board of Trustees. While the report had a few noteworthy findings, by and large, the findings show that UCAR exceeded the industry standard for phishing preparedness and only a small handful of systems had issues considered serious.

****IDD cluster**** -- We've switched the operational IDD cluster from UCAR's Foothills Lab campus to the NCAR Wyoming SuperComputing Center (NWSC) in late August. Notable changes with this cluster include 10Gb/s network service and the ability for LDM queue sizes to hold several hours worth of data .

****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.

****UCAR FISMA**** -- UCAR currently has contractual FISMA requirements for a small parts of the organization (none for Unidata currently), but the plan is to start down the path of making the entire organization FISMA compliant at some level yet to be determined (probably low, possibly medium). We continue to attend meetings to stay connected to the process

****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.

****Removal of website authentication/database**** -- Given that only Unidata staff, governance committee members, and users of McIDAS-X need to access restricted resources on the Unidata website, we will be looking to remove general website authentication for everyone else. This will result in the removal of the Login & Register links from the website pages. Links for the aforementioned groups to authenticate to access restricted materials will be available on the Unidata site in context-specific locations. Messages will be sent to the relevant mailing lists announcing this information when the change is made.

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 *March 2020*

Status Report: LDM

November 2019 - April 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:

- pqact(1):
  - Added one-day timeout for inactive entries that, nevertheless, had open file descriptors to eliminate the possibility of pqact(1) holding open thousands of files that have been scoured and, consequently, filling up the disk with invisible files
  - Ensure valid time-of-last-insertion before writing state file at exit
  - Correct parsing/formatting of state file with negative seconds
  - Modified "Starting Up" message to log the command-line
  - Corrected truncation of output file by the FILE action when both "-overwrite" and "-metadata" are specified. Before, the file would be truncated to the size of the data-product even though it now has a metadata header
- Logging:
  - refresh_logging(1): replaced "command kill" with "/bin/kill" to ensure correct behavior
- ldmd(8):
  - Downstream/Receiving LDM:
    - Log a warning that the (PRIMARY/ALTERNATE) transfer-mode won't work correctly if there are more than two REQUESTs for the same feed
  - Upstream/Sending LDM:
    - Don't terminate previously-executing upstream LDM with a different protocol version number
- ldmadmin(1):
  - Changed URL for GEMPAK tables updates from "git@github..." to "https://..." because the former required a key
- NOAAPort:noaaportIngester(1):
  - Changed pre-installation instruction to recommend disabling UDP packet sequence checking
  - Created keep_running(1) script to ensure that NOAAPort ingesters are restarted if they crash. Added examples to NOAAPort documentation
  - Updated main page of Doxygen-based documentation
  - Added new parameters to the NCEP GEMPAK tables:

        - Flat Ice Accumulation (FRAM)
        - Freezing Spray
        - Tropical Wind Direction
        - Tropical Wind Speed
        - Precipitation Duration
    - updateGempakTables(1): changed the "make install..." to "mv -f *.tbl..." to make it independent of the build mechanism
- Misc:
    - Improved documentation
    - 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.

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 at 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:

- Bug fixes
- More sites added to the test deployment
- Identifying network problems:
    - Reconfiguration of the Internet2 multipoint VLAN (due to a site joining the multicast or leaving it) can take about 100 seconds -- leading to missed data as the VLAN is first destroyed and then recreated. Internet2 has been made aware of this problem.
    - Two test sites (NCAR and UCSD) mistakenly limit the outgoing rate of multicast packets to 47 Mbps. The reason is unknown -- even by the network administrators -- and is being actively investigated.

Dependencies, challenges, problems, and risks include:

- Creating a workaround in the FMTP control-plane for the time it takes to reconfigure the multipoint VLAN
- Identifying and eliminating the cause of the 47 Mbps limit on multicast at NCAR and UCSD

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:

- Advise graduate students at the University of Virginia on modifications to the FMTP control-plane (Unidata time to be paid for by a separate grant from NSF)
- Working with Internet2

Relevant Metrics

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

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

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

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

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

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

Prepared *March 2020*

Status Report: McIDAS

November 2019 - April 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 was the creation of ADDE servers for NOAAPort-delivered GOES-16/17 satellite imagery and merging the resulting code into the McIDAS core at UW/SSEC.

Current Activities

- Unidata McIDAS version 2019 was made available for release, and an updated version, v2019a is in its final stage of preparation.

  V2019 includes all SSEC versions up to and including the current McIDAS-X and -XCD releases, both of which are v2019.1. V2019a will contain updates to the GOES-R/S ADDE servers along with a handful of bugfixes and additions from the SSEC provided XRD package.

- 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 latest releases feature the following:

- Updated GOES-R Series ABI servers to store GOES-16 and -17 Temperature Data Quality Flag (TDQF) information and report the values in **IMGLIST** command's FORM=AUX output, support ABI L2 image products that derive from NOAAPORT distributed image tiles, remove restrictions on names of directories containing the ABI files, and to skip past corrupted files so that IMGLIST and other commands will proceed to subsequent "good" files.
- Updated VIIR servers with improvements and enhancements. The changes include increasing the number of files allowed in a dataset from 5,000 to 20,000 and removing previous dataset descriptor naming requirements. The servers were also updated to use the SATANNOT file's new sensor source numbers 320 (SNPP-SDR) and 321 (NOAA20-SDR).

- Added DAY= and TIME= keywords to **IMGPROBE** command to allow use with archive or other servers that require the DAY= keyword to be specified. The **D** command (Alt-D) was also updated to include the day value from the frame directory in its request so that it also works with those servers.
- Updated **FRMLABEL** and **ZA** commands to include new *(DAYJ)* placeholder that plots the date in format *dd mon ccyyddd*, e.g., 12 OCT 2019285. Also updated FRMLABEL to identify RGB frames (displayed using the **RGBDISP** command) and thus apply an appropriate image label with no overlapping characters.

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 is 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 44 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 will be reactivated in late August 2020 for surveillance of eastern Pacific hurricanes.

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 very active use by those interested in satellite meteorology worldwide.

Prepared *March 2020*

Status Report: netCDF

October 2019 - April 2020

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Are there other cloud-based block storage formats/locations (TileDB, Azure, etc) that are actively in use? What is the next venue for investigation once we have our Zarr support in place?
2. Are there any emergent avenues (stack overflow, etc) for user support which the netCDF team should investigate?
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 **101** open issues for netCDF-C, **48** open issues for netCDF-Fortran, and **31** 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:

- Multiple releases of the core C library, as well as the Fortran and C++ interfaces.
- Refinement of user-defined compression filters.
- Continuing Work towards enhanced parallel I/O.
- Further enhancements to the netCDF-C documentation, modernization of the netCDF-Fortran documentation.
- Progress moving the netCDF Users Guide (NUG) into its own repository.
- Extended continuous integration platforms have been adopted.
- 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 have seen a high volume of contributions to the netCDF code base(s) from our community. While these contributions require careful review and consideration, it is encouraging to see this model of development (enabled by our move to GitHub) being more fully embraced by our community.

Dependencies, challenges, problems and risks include:

- Small group of developers for supporting large project.
- Dependency on HDF5, controlled by external group.
- 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 organization of Doxygen-generated 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

- Meta data 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:

- Work on reducing the defects reported by static analysis.
- Release the first version of netCDF with Zarr support.
- Release the next 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.
- Enhance thread-safety for 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:

- **900,000** for netCDF-3
- **861,000** for netCDF-4
- **924,000** for HDF5
- **106,000** for GRIB2

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

- **407** for netCDF-3
- **853** for netCDF-4
- **14,900** for HDF5
- **1,240** 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 *March 2020*

Status Report: Outreach to Underserved Communities

October 2019 - April 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

Participated in the AISES Annual Conference

- Identified Tribal Liaison within NOAA and initiated relationship
- Promoted Unidata internship opportunities to participating students

Engaged with SACNAS

- Further developed relationship with NOAA Tribal Liason
- Attended annual conference
- Made valuable connections from a wide range of MSI's
    - Enhanced, during the visit, and potentially (via equipment grants) upgrade U of Hawai'i - Manoa's VizLab

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

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

Internships

- Active engagement in the SOARS program
    - Selection Committee and mentoring

- SCIParCS
    - Mentoring

- Unidata Internship
    - Mentoring

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 *March 2020*

Status Report: Python

October 2019 - April 2020

Ryan May, Drew Cameron, Sean Arms, Julien Chastang

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What can we do to improve the utility of our training resources for you and your students? How can Unidata do more as a core Python training resource for the community? How do we grow our reach and serve the community better in this area?
2. Are there additional opportunities (e.g. conferences) we should explore as convenient venues for teaching workshops/courses?
3. Are we missing something? Anything you notice as lacking in MetPy or Siphon?

Activities Since the Last Status Report

Staffing Changes

Mid-February 2020, Drew Cameron was hired to fill the position vacated by Zach Bruick in November 2019. Drew has been working hard coming up to speed on the various projects and responsibilities of the Python team.

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 also continue to explore other options for conferences for teaching Python workshops. One thought is to look at traditionally science-focused conferences as a potential venue to teach more short courses.

We have completed a first cut at the new Unidata Python Training site, which debuted at the 2020 AMS Annual Meeting. This site serves as a single resource for discovering our Python training resources, combining the suite of materials previously available in the Python Gallery, Unidata Python Workshop, and Unidata Online Python Training. We will continue to iterate on both the user experience of the site as well as its technical underpinnings.

Progress has been made on the following:

- A collaborative Earthcube proposal with CISL, CGD, and the University of Albany was submitted to NSF, focusing on building out a larger Python training portal with a

cloud-deployable platform to execute the training materials.
- Ryan May and John Leeman, together with Kevin Goebbert, taught a short course using MetPy at the 2020 AMS Annual Meeting, focused on MetPy's simplified plotting interface. The workshop was sold out, with 23 attending, mostly from universities.
- Ryan May, John Leeman, and Zach Bruick taught an intermediate Python workshop at Texas Tech University in October 2019. This workshop focused on covering some more advanced topics, and left plenty of open time to work with attendees in small groups. We are still looking at ways to incorporate this format in future workshops.
- Ryan May and John Leeman taught a Python workshop at the University of Manitoba in Winnipeg, MB, Canada in February 2020. This workshop was well attended by about 30 people, including a sizable contingent from the University of North Dakota.
- Ryan May and Julien Chastang helped teach triennial follow-on workshops at the University of Oklahoma and University of North Carolina Charlotte in March 2020. Both workshops were well attended.
- 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 primary efforts are focused around improved units support and integration with xarray, along with working on the framework to automate computation of some derived quantities. MetPy 1.0rc1 was released at AMS, and 1.0 final is planned for release in the spring. Staffing issues, as well as technical challenges around xarray integration, slowed progress versus the original plan. This release was showcased in a talk on MetPy 1.0 at the 2020 AMS Annual Meeting.

Progress has been made on the following:

- Community awareness continues to grow, with the volume of engagement and mentions on social media growing; the MetPy twitter account has reached 1168 followers (25% growth in 6 months).
- MetPy 0.12 was released with some bug fixes and new features, including METAR parsing, as well as some warnings for future changes to the library. This is also the first release that only supports Python 3.
- MetPy 1.0rc1 was released, with mainly backward-incompatible changes and fixes on top of 0.12.
- 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. Aodhan Sweeney, one of Unidata's summer interns, presented his work in Siphon in a talk at the 2020 AMS Annual Meeting.

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

External Participation

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

Progress has been made on the following:

- Ryan May will attend a NumFocus Project Summit for the matplotlib project
- 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 Cameron are planning to attend SciPy 2020, provided it is still held.
- We also continue to actively engage with the xarray, numpy, 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:

- Teach additional Python regional workshops, including one at a minority-serving institution (pending being able to travel again)
- Release MetPy 1.0 final
- Continue developing Unidata Python Training site

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

- Teach another short course on MetPy at AMS 2021, possibly two days
- Teach another short python overview course at AMS 2021 student conference
- 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: 54
- Downloads for the releases made in the last year (Conda + PyPI):
    - 0.10.2: 6721
    - 0.11.0: 1259
    - 0.11.1: 6821
    - 0.12.0: 16621
    - 1.0.0rc1: 177
- Since 1 October 2019
    - Active Issues: 116 (84 created, 59 closed)
    - Active PRs: 92 (72 created, 72 closed)
    - External Issue Activity: 43 opened, 159 comments
    - External PR Activity: 49 opened, 120 comments
    - Unique external contributors: 43
    - Stars: 75 (549 total)
    - Forks: 4 (219 total)
    - Commits: 300
- Since 1 March 2019
    - Active Issues: 212 (148 created, 134 closed)
    - Active PRs: 191 (179 created, 167 closed)
    - External Issue Activity: 78 opened, 258 comments
    - External PR Activity: 77 opened, 183 comments
    - Unique external contributors: 74
    - Stars: 174 (549 total)
    - Forks: 14 (219 total)

    - Commits: 593

Siphon

- 97% test coverage
- Watchers: 17
- Downloads for the last year (Conda + PyPI):
    - 0.8.0: 33973
- Since 1 March 2019
    - Active Issues: 13 (10 created, 4 closed)
    - Active PRs: 13 (10 created, 8 closed)
    - External Issue Activity: 5 opened, 2 comments
    - External PR Activity: 9 opened, 10 comments
    - Unique external contributors: 8
    - Stars: 18 (117 total)
    - Forks: 1 (46 total)
    - Commits: 15
- Since 1 September 2018
    - Active Issues: 23 (18 created, 7 closed)
    - Active PRs: 22 (21 created, 17 closed)
    - External Issue Activity: 5 opened, 2 comments
    - External PR Activity: 9 opened, 18 comments
    - Unique external contributors: 12
    - Stars: 27 (117 total)
    - Forks: 2 (46 total)
    - Commits: 51

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 *September 2019*

Status Report: Support

November 2019 - April 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

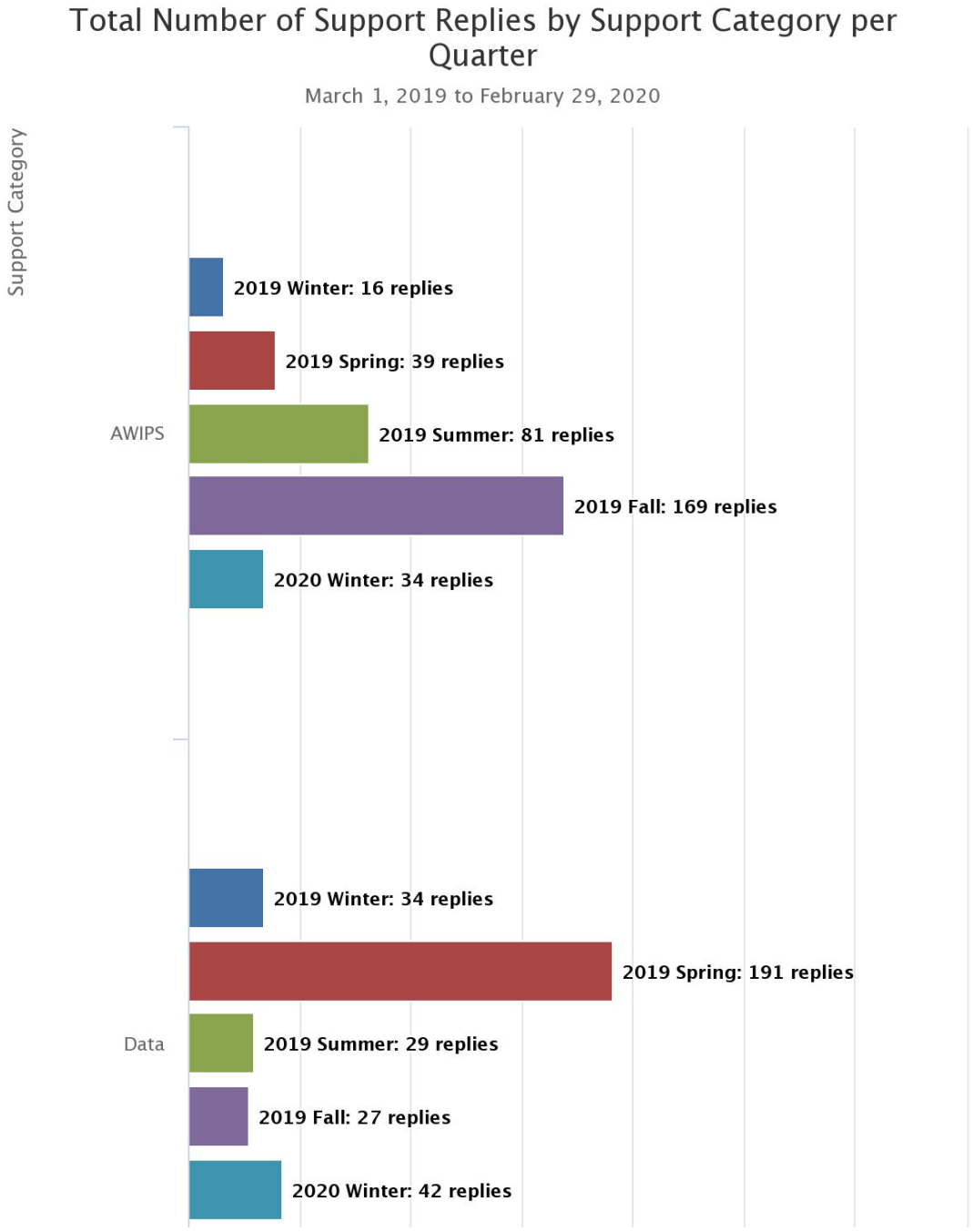
Since January 26, 2006 over 61999 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

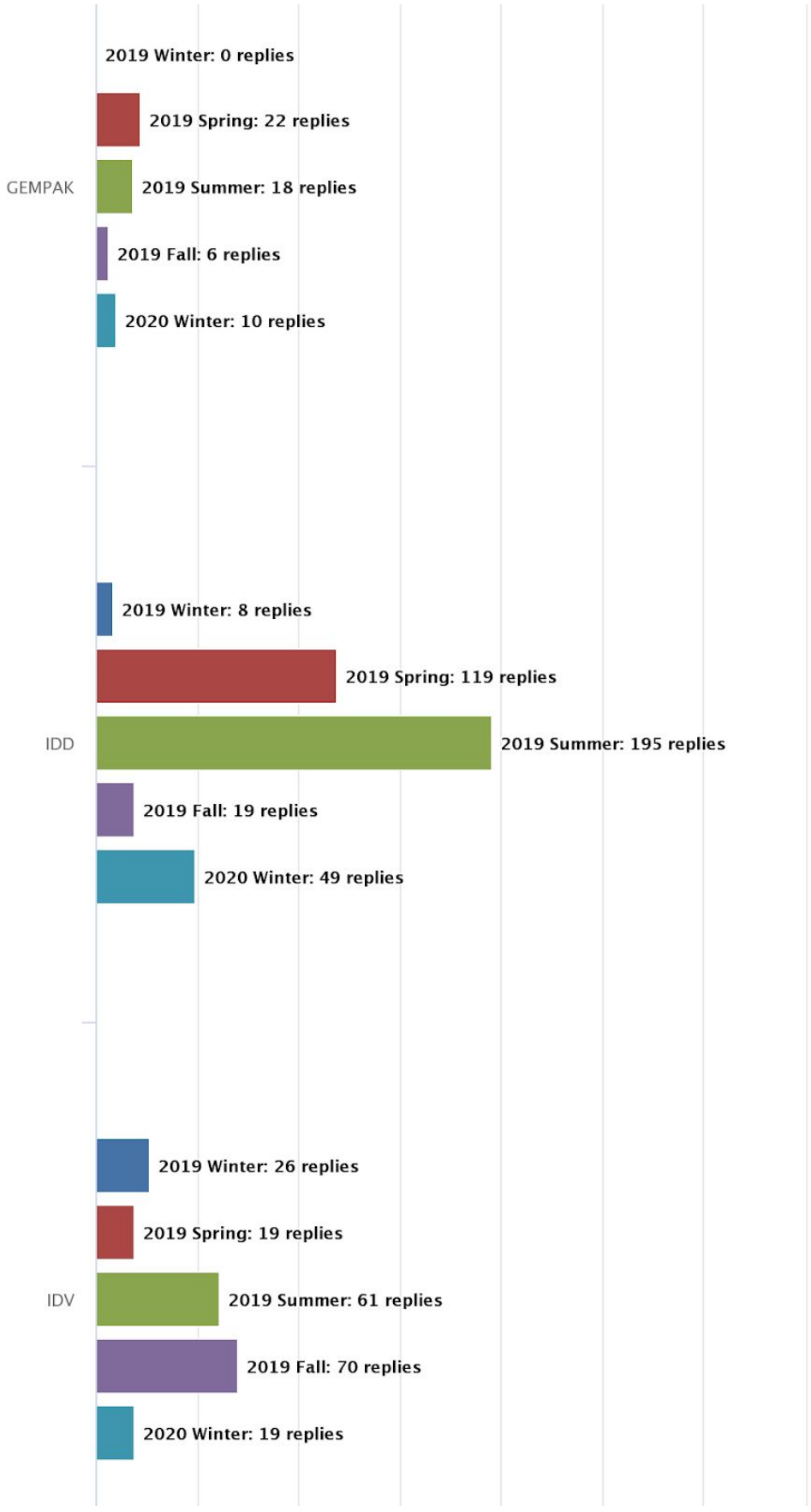
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

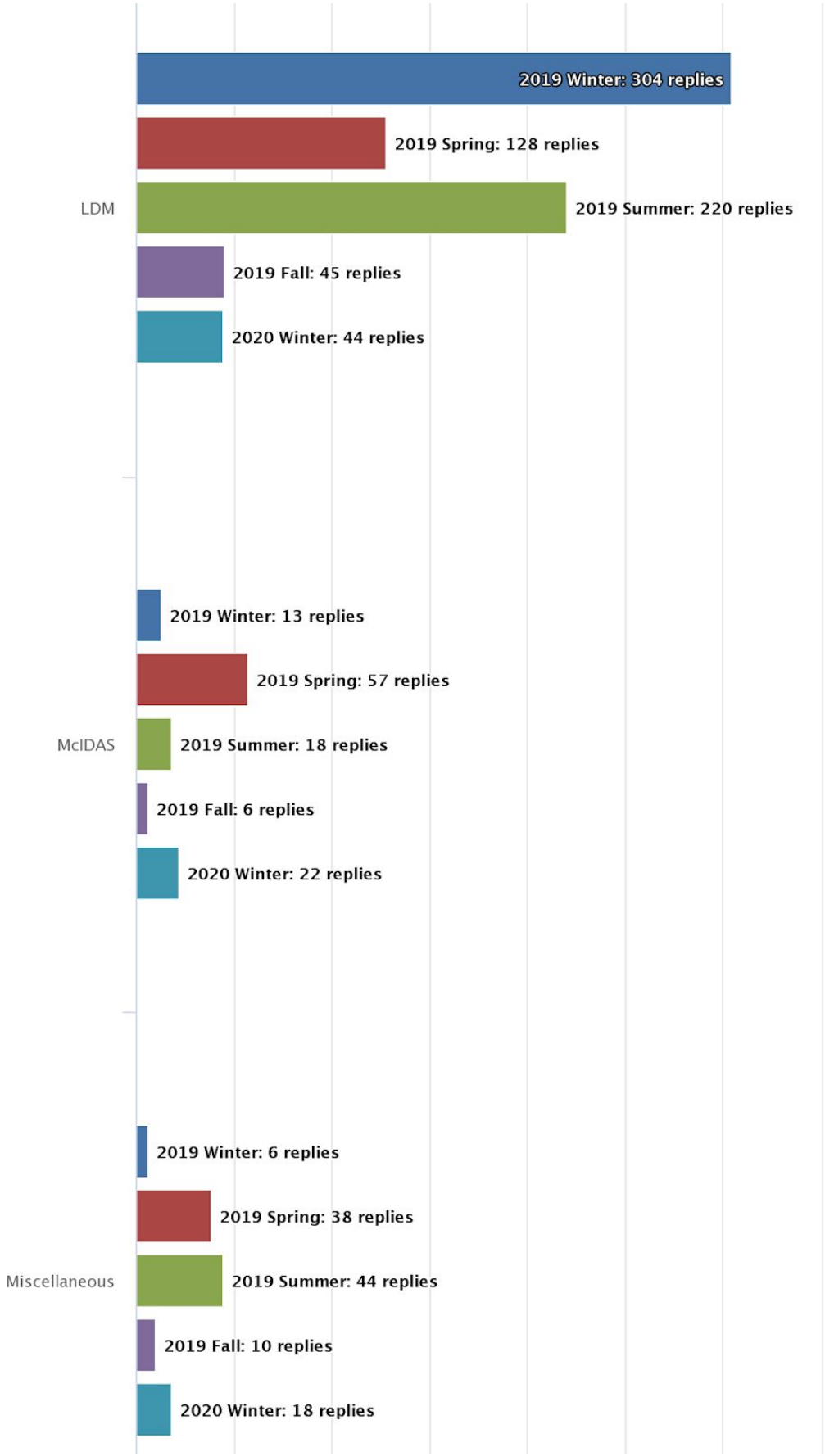
Fig. 1: Below are histograms that portray the number of Unidata email responses for categories of support logged in the Unidata Inquiry Tracking System for the 12 month period from September 1, 2018 until August 31, 2019.

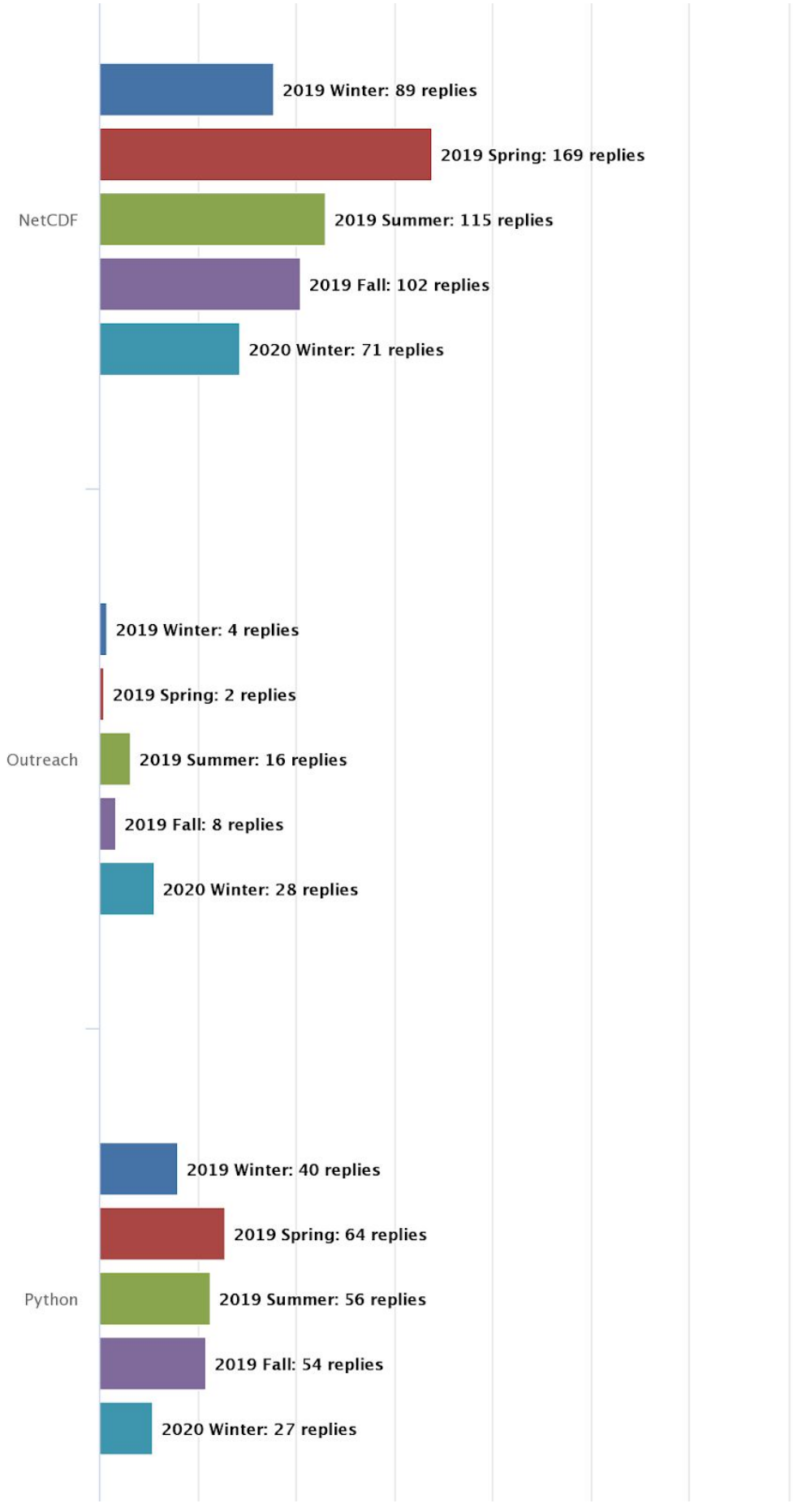
The quarters shown are defined as:

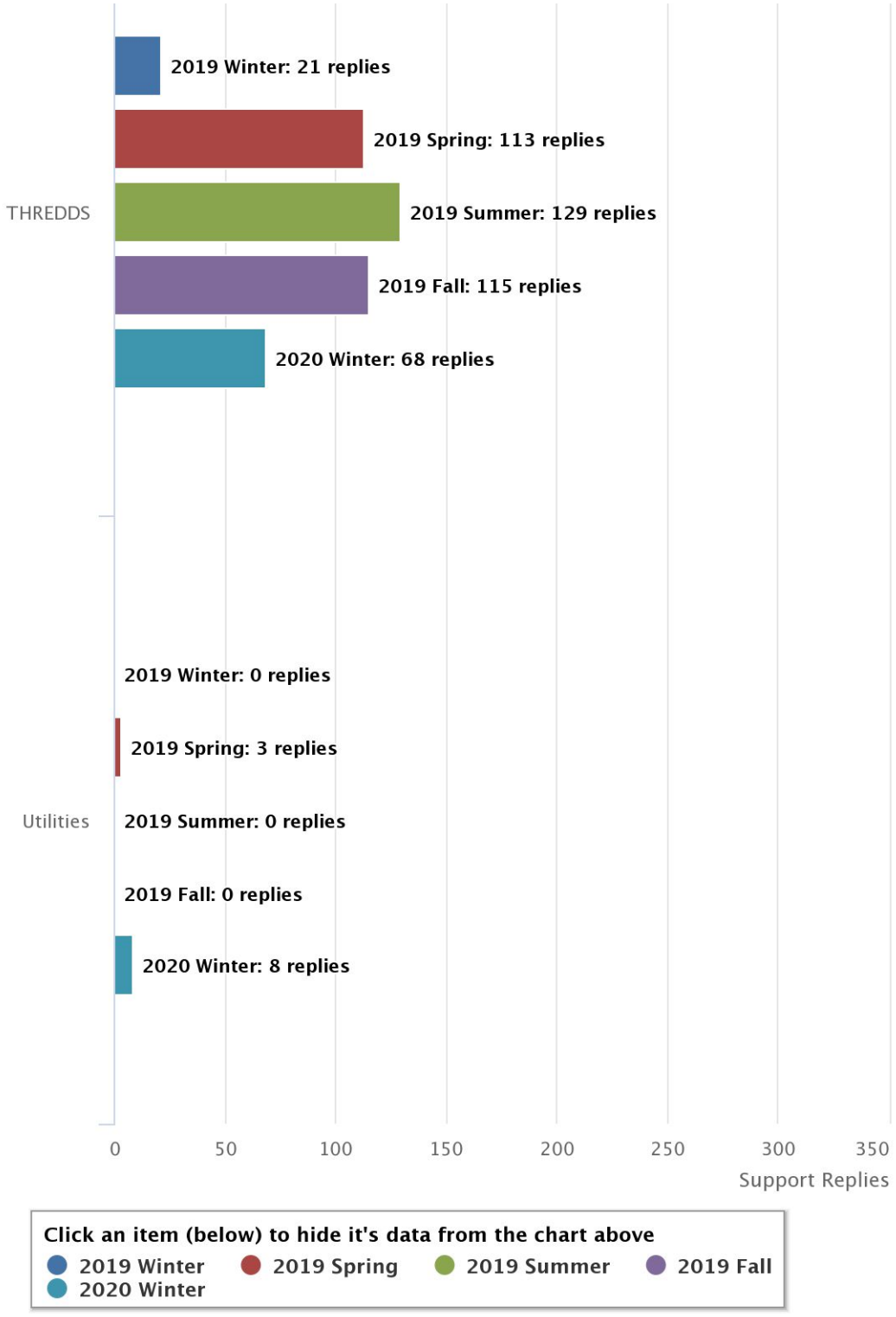
Winter:	**Spring:**	**Summer:**	**Fall:**
January, February, March	April, May, June	July, August, September	October, November, December




2019 Winter: 0 replies
2019 Spring: 22 replies
GEMPAK
2019 Summer: 18 replies
2019 Fall: 6 replies
2020 Winter: 10 replies
2019 Winter: 8 replies
2019 Spring: 119 replies
IDD
2019 Summer: 195 replies
2019 Fall: 19 replies
2020 Winter: 49 replies
2019 Winter: 26 replies
2019 Spring: 19 replies
IDV
2019 Summer: 61 replies
2019 Fall: 70 replies
2020 Winter: 19 replies


LDM
2019 Winter: 304 replies
2019 Spring: 128 replies
2019 Summer: 220 replies
2019 Fall: 45 replies
2020 Winter: 44 replies
McIDAS
2019 Winter: 13 replies
2019 Spring: 57 replies
2019 Summer: 18 replies
2019 Fall: 6 replies
2020 Winter: 22 replies
Miscellaneous
2019 Winter: 6 replies
2019 Spring: 38 replies
2019 Summer: 44 replies
2019 Fall: 10 replies
2020 Winter: 18 replies


NetCDF
2019 Winter: 89 replies
2019 Spring: 169 replies
2019 Summer: 115 replies
2019 Fall: 102 replies
2020 Winter: 71 replies
Outreach
2019 Winter: 4 replies
2019 Spring: 2 replies
2019 Summer: 16 replies
2019 Fall: 8 replies
2020 Winter: 28 replies
Python
2019 Winter: 40 replies
2019 Spring: 64 replies
2019 Summer: 56 replies
2019 Fall: 54 replies
2020 Winter: 27 replies


THREDDS
2019 Winter: 21 replies
2019 Spring: 113 replies
2019 Summer: 129 replies
2019 Fall: 115 replies
2020 Winter: 68 replies
Utilities
2019 Winter: 0 replies
2019 Spring: 3 replies
2019 Summer: 0 replies
2019 Fall: 0 replies
2020 Winter: 8 replies
0
50
100
150
200
250
300
350
Support Replies
Click an item (below) to hide it's data from the chart above
2019 Winter
2019 Spring
2019 Summer
2019 Fall
2020 Winter

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 *March 2020*

Status Report: THREDDS

October 2019 - March 2020

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

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Do you use netCDF-Java as a library? We are currently working on defining the public API and would love your feedback!
2. Do you know any students in your program who actively use Java?
3. We are currently in the process of hiring. Do you know of anyone skilled in Java?

Activities Since the Last Status Report

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 5.3.1

- NetCDF-Java version 5.0.0 was released on 29 July 2019. NetCDF-Java version 5.3.1 was released on 12 March 2020.
- 5.3.1 includes AWS S3 support, so any file netCDF-Java is able to read can now be read directly off of an S3 object store.
- 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 4.6.14 was released on 20 March 2020. This is mostly a maintenance release and recommended for anyone running the TDS.
- As part of this release, netCDF-Java version 4.6.15 artifacts (jar file) were made available. However, the 4.6.x line of development for netCDF-Java is now in maintenance mode, and will only include security related fixes. All users of netCDF-Java are strongly encouraged to move to the latest netCDF-Java (as of this

report, version 5.3.1).

TDS version 5.0.0

- We anticipate releasing a stable version of TDS 5.0.0 before the beginning of summer 2020.
- 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/5.1/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:

- 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.

General dependencies, challenges, problems, and risks include:

- While nearly all java based components in the THREDDS project run under Java 11, all except Rosetta currently build with Java 8. Portions of our build infrastructure need to be reworked to use Java 11. The end of 2020 is generally marked as end-of-life for Java 8, and thus moving to Java 11 is a priority. netCDF-Java will be the first to build

with Java 11, and we anticipate that happening before Summer 2020. The TDS WMS server currently does not run on Java 11, and we are in contact with the University of Reading team to see what can be done.
- 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.!).
- While the list of names on this report seems large, the current staffing levels on all three components covered in this report is less than 1.5 FTE (including externally funded efforts). A similar resource landscape can be seen for nearly every other project run by the Unidata Program Center. External funds help, but rarely provide the ability to bring on and sustain new staff members, which results in taking resources away from other projects and efforts within the Unidata Program Center. Given your position as members of our governing committees, you play a critical role in helping us set priorities. Your feedback is greatly desired and very much appreciated.
- We are currently in the process of hiring. Do you know of anyone skilled in Java?

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.

The following active proposals directly involve THREDDS work:

- 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).

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.
    - Facilitate migration from netCDF-Java 4.6.x to 5.x where we can.
    - 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 late 2019/ early 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.

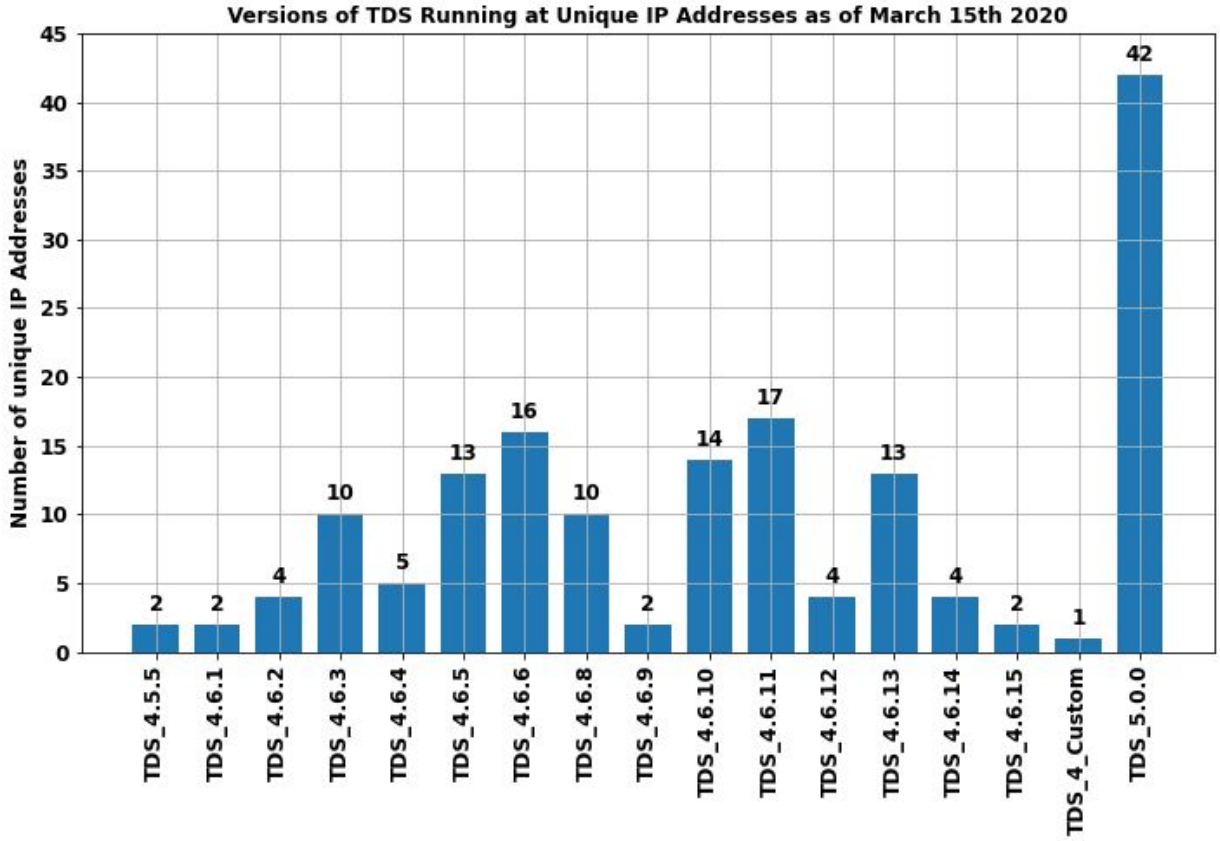
Relevant Metrics

NetCDF-Java

In our last report, we mentioned that the method by which we count download statistics for the netCDF-Java library has been excluding a significant source of downloads. Traditionally, users have downloaded a single jar (netcdfAll.jar) to use with their packages. However, the way Java developers have been consuming our libraries has changed significantly over the past several years, as many projects now rely on a build system, such as maven or gradle, to pull in just the components of the library they need, and we have not taken that into account.

Last report, we showed the yearly downloads for netCDF-Java to be 8,714 (September 1, 2018 - August 31, 2019). Using the updated method of counting for the past 12 months, the yearly downloads for netCDF-Java is **36,897** (March 1, 2019 - February 29, 2020). Note that this does not double count downloads - a user downloading netcdfAll.jar would not also download cdm-core.jar, and vice versa.

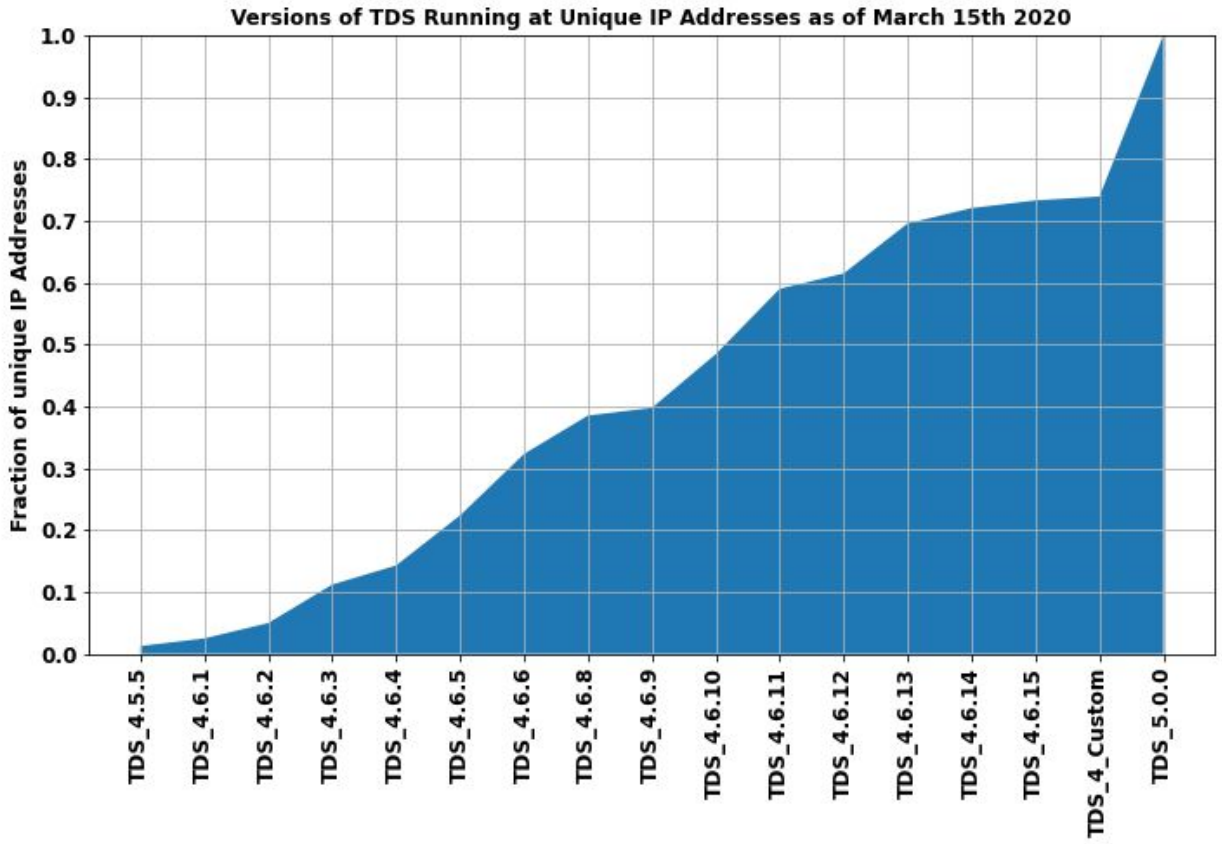
THREDDS Data Server



We've seen that over the past six month period of time, **3,469** unique IPs started up the TDS (August 2019 through March 2020), **62** 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 **28,675** unique IP addresses. Overall, we currently have a total of **161** publically accessible TDSs running "in the wild" (up from our last report, which had **106**) . "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 (v4.6.14 was the most current release during this period, and was released on 26 July 2019). We now see that TDS v5 has become the dominant specific version running in the wild.



Furthermore, of the **161** publically accessible servers, **64** 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…).

In the next six months, 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 *March 2020*