Users Committee Meeting

4-5 April 2018

Agenda

Unidata Program Center, Boulder CO

Wednesday, 4 April 2018

8:30 - 9:00 Continental Breakfast
9:00 - 9:15 Welcome and Administrative Items

- Date for Fall meeting
- Review of Action Items from September 2016 meeting

9:15 - 9:30 Strategic Advisory Committee Report - Mohan Ramamurthy and Russ Schumacher
9:30 - 10:45 Director's Report & Questions - Mohan
10:45 - 11:00 Break
11:00 - 11:30 ECMWF Data Status - TBD
11:30 - 12:45 Lunch
12:45 - 1:30 Around-the-table reports from members
1:30 - 2:15 Status Reports - (Users Committee should review status reports prior to the meeting) - staff will be available for questions
2:15 - 2:30 Break
2:30 - 3:00 JupyterHub on the cloud - Rich, Ryan, Julien
3:00 - 3:30 Users Committee Scope - Josh
3:30 - 3:45 TDS 5.0 - Sean
3:45 - 4:30 NOAA - Carissa Klemmer
4:30 - 5:00 AWIPS Development Plans - Michael James
5:00 Adjourn
6:30 Collaborative discussion on the day's proceedings over dinner

- Dinner at Gondolier Italian Eatery 4800 Baseline Rd. Boulder, CO (map)

Thursday, 5 April 2018

8:30 - 9:00 Continental Breakfast
9:00 - 10:00 Triennial Workshop - John, Kevin, Victor, Josh
10:00 - 10:30 DeSouza Nominees - Josh
10:30 - 10:45 Break
10:45 - 11:45 Blue Sky Session
11:45 - 12:30 Transfer of Chair/AOB

Status Report: AWIPS and GEMPAK

October 2017 - March 2018

Michael James

Questions for Committee Members

- (from the GOES Status Report) Which GOES-16/17 bands and coverages should be converted from netCDF-4 to McIDAS AREA format for GEMPAK/NMAP2?
- What additional data and map resources could be added to AWIPS?

AWIPS

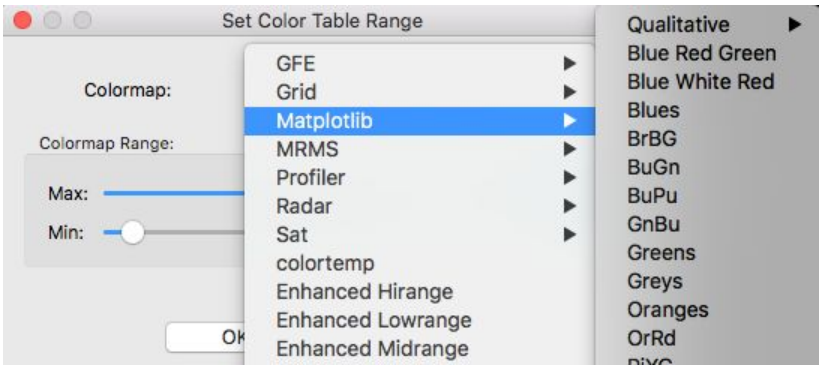
Unidata continues to maintain an EDEX data server on the Jetstream cloud, serving real-time AWIPS data to CAVE clients and through the python-awips data access framework (API). The distributed architectural concepts of AWIPS allow us to easily scale EDEX in the cloud to account for the size of incoming data feeds. We continue work using Jetstream to develop cloud-deployable AWIPS instances, as imaged virtual machines (VMIs) available to users of *Atmosphere* and *OpenStack*, and as docker containers available on Docker Hub and deployable (soon) with the *xsede-jetstream* toolset.

Jetstream AWIPS EDEX Standalone VMI
This EDEX image can serve either as a full standalone server or as a database/request server.
Jetstream AWIPS EDEX Ingest Node VMI
This image contains all AWIPS Python and EDEX software, but nothing database or http-related. This VMI makes it easy to deploy datatype-specific ingest nodes (Grid, Radar, Satellite, etc.) with simple edits to *ldmd.conf* and *modes-ingest.xml*

AWIPS 17.1.1-5 released January 2018

- **GOES East (16)** updates for operational mode, geostationary position adjustment, ingest of stitched CONUS and Full Disk netCDF files, menu updates for Geostationary Lightning Mapper (**GLM**), and Derived Motion Winds (**DMW**)
- EDEX no longer installs **awips2-python-setuptools**, which was not required and in some cases would prevent virtual machines from booting due to Python version conflicts.
- Removed IFPS server send dialogs from GFE perspective.
- Support for **NationalBlend**, **HIRLAM**, more complete **HRRRX** parameters
- Distributed computing support for ingest nodes writing data to a remote database/request instance - read more at http://unidata.github.io/awips2/edex/distributed-computing/
- Fix for long request times to fill colormap editor right-click menu, which now displays

only your own USER-level localization colormap edits
- Removed NCEP EDEX data decoders (rpm awips2-edex-ncep) from build and install (this does not affect the common/viz plugins or the NCP)
- Menu and bundle updates for WPC QPF/QPE grids
- Native handling of PNG-compressed GINI and AREA files with **ar.com.hjg.pngj**
- Pre-built el6 and el7 binaries of **edexBridge** now used for bundling during build in container
- Converted and added Matplotlib colormaps to AWIPS .cmap format, available in the right-click menu **Change Colormap** dialog:



- Removed bundled **netCDF** RPM packages to avoid conflict with **epel** repositories
- Removal of more unused AWIPS1, Hydro, RadarServer components
- Software updates for *qpid, httpd-pypies, postgresql, matplotlib, numpy, cython, shapely, python-qpid, dateutil, python-scientific, python-tables.*
- Removed unused python modules *nose, scipy, pupynere, pwm.*
- The script **qpidNotify.py** has been added to **/awips2/edex/bin** and is available in **$PATH** for non-root users on an AWIPS server.

Python Data Access Framework (python-awips)

An early version of the AWIPS Data Portal has been published at http://js-157-49.jetstream-cloud.org:8080, including EDEX Server Inventories, Forecast Models, NEXRAD level 3, and instructions on how to build a GeoJSON API.

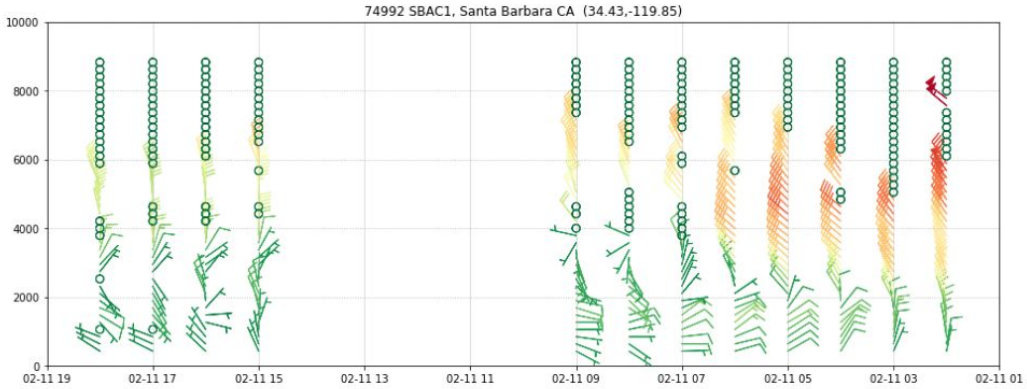
0.9.12 released March 9, 2018

- New NEXRAD dictionary in tables.py programmatically-generated from AWIPS WSR-88D radar definition file with class GenerateNexradTable (`from awips.tables import nexrad`)
- Added pandas and h5py to conda environment
- All EDEX-side Derived Parameters are now **callable** from python-awips and returned as 2-d Numpy arrays (as of EDEX 17.1.1-5)

0.9.11 release Feb 11, 2018

- Added profiler site metadata to tables.py profiler time series notebook (`from awips.tables import profiler`)

- Created and published a [Profiler Wind Barb Time-Series Jupyter notebook] using newly-acquired ESRL/PSD profiler hourlies.



GEMPAK/NAWIPS

Docker Container

https://hub.docker.com/r/unidata/gempak/

- This image contains the latest GEMPAK RPM (7.4.1), **epel-release**, and **python-pip**
- Python packages installed via pip include **python-awips**, **numpy**, **shapely**, **six**
- To run:
  - `docker pull unidata/gempak`
  - `docker run --user gempak -it --rm=true unidata/gempak:latest`
  - `./rungempak.sh` from within the container to run GDINFO using remotely-retrieved AWIPS grids.

GEMPAK 7.4.1 (September 2017)

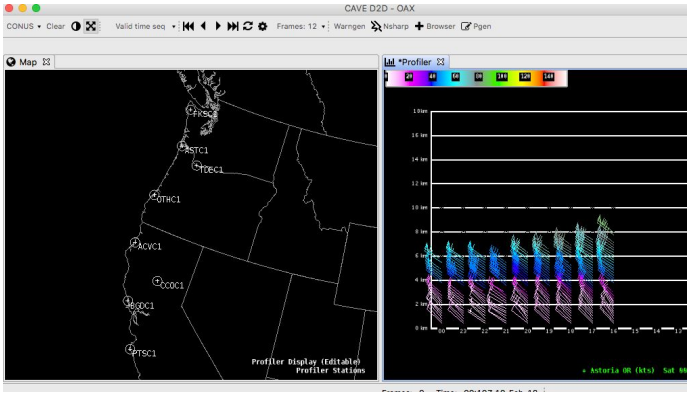
https://github.com/Unidata/gempak/releases/tag/7.4.1

- Grib2 table updates for moisture, momentum, mass, radiation, hydrology, cloud/aerosol, more.
- Removed repetitive dcmetr log messages.
- Pattern actions in ldm/etc/templates/pqact.gempak_decoders_grid for various NamDNG grid and the National Blend of Models.
- Alias/mod_res updates, increased max grids in file for gfs, nam, National Blend, SST, URMA, others.
- Bug fix for AWIPS grid retrieval in scripts/python/GridDataRetriever.py

Activities Ongoing/In-Progress

AWIPS

- Continue work on a deployable **edex-ingest** Docker container.
- Continue work on a deployable edex-ingest-wrf instance
- Collaborate with AWIPS developers at GSD on distributed containerization of AWIPS/EDEX.
- Continue to create and publish datatype workflow **Jupyter notebooks for python-awips**
- Continue expanding the capabilities of the **AWIPS Data Portal** as new Jupyter notebooks are created.
- Next AWIPS release will be **17.1.1-6**
- **GOES East (16)** updates for operational mode, geostationary position adjustment, ingest of stitched CONUS and Full Disk netCDF files, menu updates for Geostationary Lightning Mapper (**GLM**), and Derived Motion Winds (**DMW**)
- Moved **/awips2/edex/bin/edex** to **/usr/bin/edex** to find in sudo user path
- **Profiler map resource** added to CAVE, selectable sites to load a time-series



- Added qpid-stat as "**edex qpid**" command to monitor datatype message queues
- Fix for missing colormap to default to **Grid/Gridded Data** (rainbow) and alert/WARN the user that the colormap in their bundle could not be found. This solves the problem created by simply renaming a **.cmap** file which used to throw an error and fail to load the product/bundle. Because colormap definitions are controlled on the EDEX side of the server-client exchange, this catch allows us to rename and update colormaps without users having to update or reinstall CAVE.
- CreateOpenStack-created VMIs

GEMPAK

- Next release will be **7.5.1**
- GOES **fixed grid projection** McIDAS navigation added to development GEMPAK 7.5.1
- GOES16/17 sectors and products to be made available in AREA format for GEMPAK is still TBD
- **gdpvsf** added back to release after fix for max number of grids to avoid memory

relocation error

New/Future Activities

- Dockerize gempak on top of LDM.
- Test AWS API for public data set retrieval via EDEX.
- Expand AWIPS geospatial city table to include major world cities.
- Deploy a standalone AWIPS Archive EDEX server in the cloud to contain notable case study events.
- Investigate how the Weather Event Simulator 2 Bridge (WES-2 Bridge) can be (and if it should be) merged into the Unidata AWIPS baseline.
- Migrate NCEP Viz plugins from the NCP to D2D.
- Investigate creating an AppImage executable for CAVE for wider Linux distro support.
- Add networking to the CAVE Archive Case Study creation tool.
- Incorporate USER-level bundles and modifications in the D2D data selection menus

Metrics

AWIPS Downloads

- 11246 total
- See core institutions
- See all EDUs
- See per month
- See affiliation types
- See countries
- See US vs non-US
- See Core vs non-Core
- See list of download sites

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Both AWIPS and GEMPAK are freely available, and both incorporate LDM/IDD technology for accessing 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. **Develop and provide open-source tools for effective use of geoscience data**
   Both AWIPS and GEMPAK are open-source, and while GEMPAK is now in maintenance mode, AWIPS is continuously being developed.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**

Unidata is the only known entity to provide a freely-available and non-operational version of the AWIPS software package. Unidata continues to find and make available new datasets through the AWIPS project.

4. **Build, support, and advocate for the diverse geoscience community**
   Using LDM/IDD technology to provide access to real-time meteorological data; providing visualization tools for data analysis.

Prepared *March 2018*

Status Report: *Cloud Computing Activities*

October 2017 - March 2018

Julien Chastang, Ward Fisher, Michael James, Ryan May, Jen Oxelson, Mohan Ramamurthy, Christian Ward-Garrison, Jeff Weber, Tom Yoksas

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. What clouds are our community using, either commercial (e.g., Amazon) or non-for-profit (e.g., NSF XSEDE Jetstream)?
2. What new cloud technologies are our community using and investigating on their own initiative?
3. Who would like to volunteer to beta test CloudIDV?
4. Who would like to volunteer to beta test: https://jupyter-jetstream.unidata.ucar.edu

Activities Since the Last Status Report

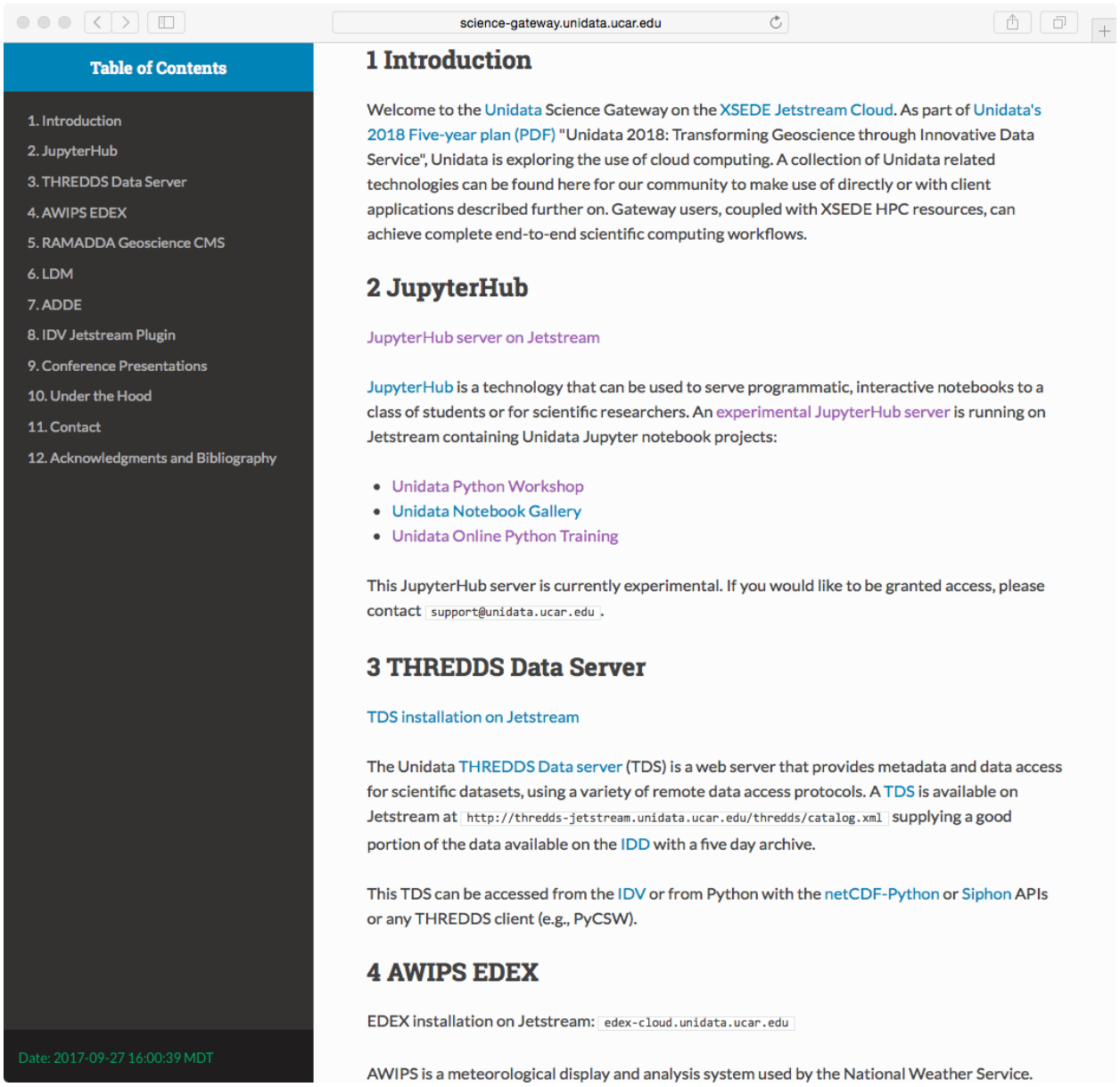
Updates to NEXRAD in AWS

Two separate requests were received from the community (NCAR, Oklahoma) to enhance the AWS Simple Notification Service topic for the realtime (i.e., individual chunk) S3 bucket. The goal was to allow users to utilize message filters in order to only subscribe for a subset of the data; for example to only request data for a particular site or sites. Ryan May worked with AWS and NCAR to prototype a new topic, and the new AWS API calls have been added and deployed to the AWS NEXRAD processing software.

Serverless Experiment

Ryan May has begun experimenting with so-called “serverless” cloud technology. The idea with “serverless” is that, instead of managing a virtual machine running in the cloud, various other cloud infrastructure services are leveraged to create an application that runs in an event-driven fashion, without the need for a VM that is idle much of the time; for simple applications, this can represent significant cost savings over running a compute instance. As a concrete example, AWS provides API Gateway as a service for routing web requests and Lambda as a service for running code (e.g. Python, Java). Using these services, Ryan has put together a web application (for syncing GitHub issues to the Asana project management tool) that requires no continuously running server. Ryan plans to continue to explore this space as a way for creating simple, scalable web services running affordably in the commercial cloud.

Unidata Science Gateway on Jetstream



Building upon our previous containerization efforts, we are continuing to enhance the Unidata Science Gateway on NSF-funded XSEDE Jetstream Cloud: http://science-gateway.unidata.ucar.edu/. A collection of Unidata related technologies can be found here for our community to make use of directly or with client applications such as the IDV. The following resources are available on this gateway:

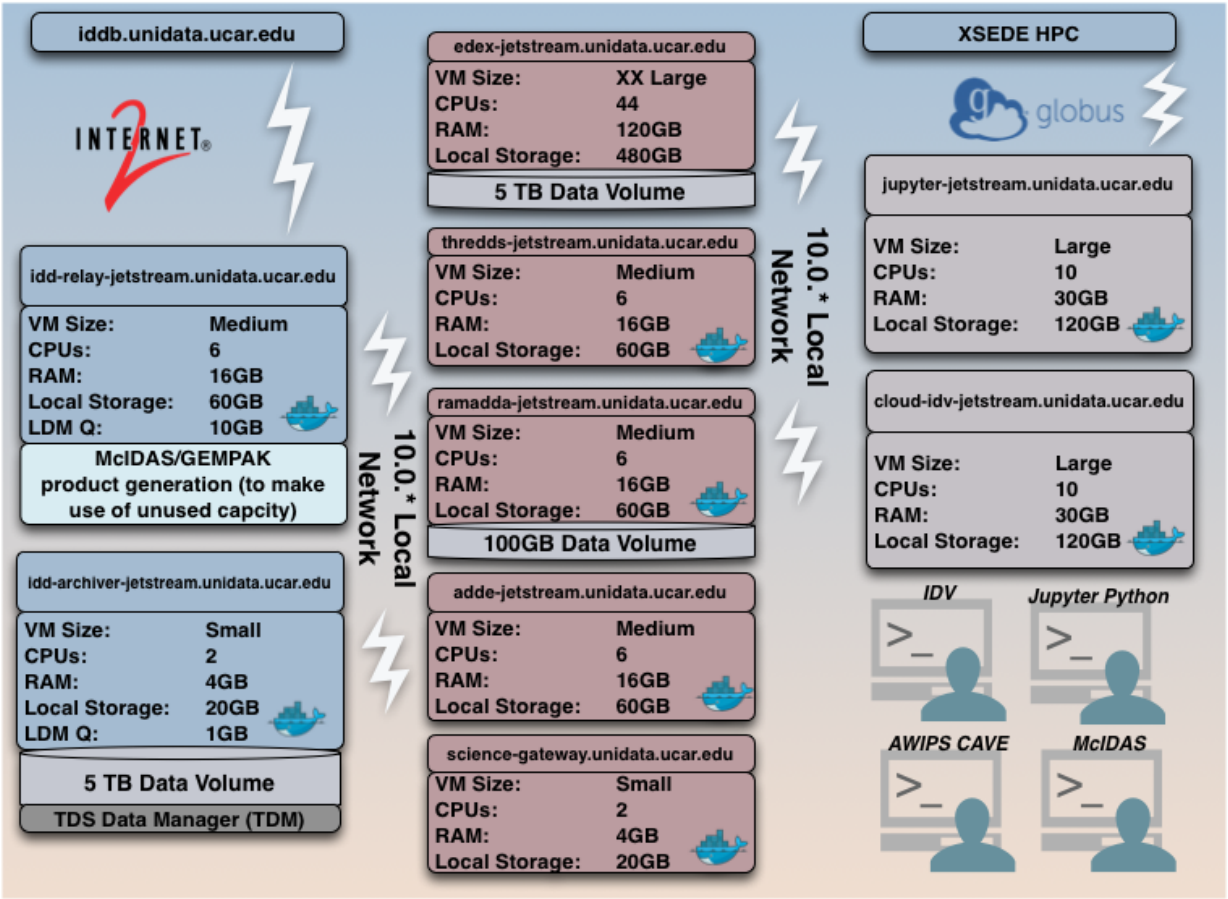
- An experimental JupyterHub server containing Unidata Jupyter notebook projects (see Python status report for more information).
- A TDS supplying a good portion of the data available on the IDD with a five day archive.

- EDEX server
- RAMADDA containing IDV bundles that retrieve data from Jetstream data servers.
- Two LDM nodes
- ADDE Server
- IDV Jetstream plugin that allows easy access to Jetstream installations of the TDS, RAMADDA and ADDE from the IDV.

Gateway users, coupled with XSEDE HPC resources, can achieve complete end-to-end scientific computing workflows. In the past six months, we have done three presentations on this work:

- Gateways 2017 | October 23–25, 2017 – Ann Arbor, MI USA
- 2017 AGU Fall Meeting | December 11–15, 2017 – New Orleans, LA USA
- 2018 AMS Annual Meeting | January 7–11, 2018 – Austin, TX USA

A complete bibliography of this effort is available here.



Dependencies, challenges, problems, and risks include:

- We would like to transition from an experimental and research and development approach, to an operational mode. To achieve this objective, we are working with the Jetstream team to improve VM uptime availability. To that end, Unidata has ameliorated its monitoring of Jetstream VMs so that we have accurate metrics concerning VM availability that we can communicate to the Jetstream team.
- Unidata software staff continues to work closely with Unidata system administrators to ensure cloud VMs (especially on Jetstream) are adhering to Unidata security standards.

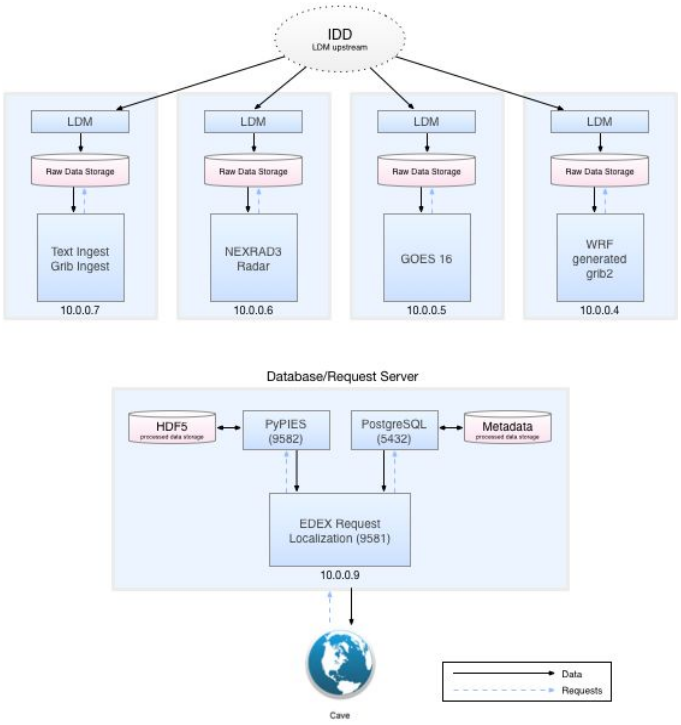
AWIPS EDEX in the Cloud

Unidata continues to maintain an EDEX data server on the Jetstream cloud, serving real-time AWIPS data to CAVE clients and through the python-awips data access framework (API). The distributed architectural concepts of AWIPS allow us to easily scale EDEX in the cloud to account for the size of incoming data feeds. By isolating the database/request processes to a single machine, we avoid *data serving* competing with *data decoding* on the same machine, minimizing the chance of reaching system memory limits which can result in EDEX shutdown.

We continue work using Jetstream to develop cloud-deployable AWIPS instances, both as imaged virtual machines (VMIs) available to users of *Atmosphere* and *OpenStack*, and as docker containers available on Docker Hub and deployable (soon) with the *xsede-jetstream* toolset.

Jetstream AWIPS EDEX Standalone VMI
This EDEX image can serve either as a full standalone server or as a database/request server.

Jetstream AWIPS EDEX Ingest Node VMI
This image contains all AWIPS Python and EDEX software, but nothing database or http-related. This VMI makes it easy to deploy datatype-specific ingest nodes (Grid, Radar, Satellite, etc.) with simple edits to *ldmd.conf* and *modes-ingest.xml*



Nexus Server on Jetstream

Unidata continues to run a Nexus Server on Jetstream for the distribution of netCDF-Java artifacts (e.g., `netcdfAll.jar`, `toolsUI.jar`, `ncIdv.jar`): https://artifacts.unidata.ucar.edu. netCDF-Java documentation is also hosted at that location.

On February 26, 2018, the VM that hosts this server experienced a 12 hour downtime due to a Jetstream network issue. We are working with Jetstream staff to avoid these lengthy downtimes in the future.

Preparing for Renewal of Research Allocation on Jetstream

Unidata is preparing to renew its Research allocation starting on March 15, 2018. Our current $425,000 allocation in cloud computing resources is on schedule to be totally consumed by June 30, 2018. We plan on requesting an allocation of equal or greater size to the original allocation.

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, AWIPS, and Python with Unidata Technologies. 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.

Unidata / NOAA Docker Collaborations

We have been collaborating with NOAA in two different containerization efforts:

- We have been advising Evan Polster on Docker technology and the containerization of AWIPS and the LDM.
- We recently worked with Marcus England on use of the LDM Docker container. Marcus found several problems with this container that we were able to quickly resolve. We thank Marcus for reporting these problems so that we can improve these containers.

Progress has been made on the following:

Unix permission issues have been a long-standing problem with Docker technology with permission "mismatches" involving the user running the container, and the non-root user **inside** the container. These issues resulted in frequent and frustrating "permission denied" errors in situations where a directory on the Docker host was mounted in the container. We recently made significant progress in this area by allowing the user running the container to supply a Unix user ID and group ID to the container thereby allowing the user inside the container to effectively be the same as the user running the container. The end result is permission problems of this type are now resolved with this new technique. This amelioration is implemented in most Unidata Docker containers.

Dependencies, challenges, problems, and risks include:

It is unlikely that most of our community will use these containers directly. Rather they will be leveraged by experts on behalf of the community, or they will be abstracted from users by being integrated into a user-friendly workflow. For example, on Jetstream we have a JupyterHub server currently in development: https://jupyter-jetstream.unidata.ucar.edu. This server was deployed with the aid of cloud computing technologies including Docker. These details, however, are hidden from the user.

In addition, there are overlapping (perhaps, competing or complementary) technologies such as Ansible that are emerging alongside Docker that need to be investigated.

Ongoing Activities

Amazon Web Service Activities and NOAA Big Data Project

NOAA Big Data Project

- In collaboration with Unidata, NOAA is delivering 20+ years of NEXRAD Level II data via Amazon Web Services. LDM and THREDDS Data Server (TDS) software are being employed to deliver these data.
- Started transferring GOES-16 data to the Amazon cloud S3 bucket.
- TDS on AWS for level II NEXRAD (For .edu access only): 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.

Product Generation for IDD

For the past three 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.

CloudIDV, CloudStream, Cloud Control

- We have released the technology enabling CloudIDV in a form that can be easily leveraged by other projects looking to bring legacy software to the Cloud. We are currently trying to build the CloudStream community via conference presentations and outreach.
- In addition, we continue to experiment with CloudIDV on the Jetstream Cloud. We are investigating CloudIDV for data-proximate visualization of the WRF-hydro modeling system.
- We presented at both AGU 2017 and AMS 2018 on CloudIDV and CloudStream.

Open Commons Consortium Award

The Open Science Data Cloud, a resource of the Open Commons Consortium (OCC), provides the scientific community with resources for storing, sharing, and analyzing terabyte and petabyte-scale scientific datasets. The OSDC is a data science ecosystem in which researchers can house and share their own scientific data, access complimentary public datasets, build and share customized virtual machines with whatever tools necessary to analyze their data, and

perform the analysis to answer their research questions. Unidata is a beta user of resources in the Open Science Data Cloud ecosystem and we have been provided cloud-computing resources on the Griffin cloud platform. Our allocations are renewed on a quarterly basis and Unidata is partnering with OCC on the NOAA Big Data Project. Given the limited staff resources and many ongoing cloud activities on AWS, Azure, and XSEDE environments, Unidata's activities on the OSDC have been in a temporary hiatus. We are hoping to ramp up our OSDC efforts in the upcoming months.

New Activities

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

Unidata Science Gateway

Forthcoming Presentations

- EGU General Assembly 2018 | April 8–13, 2018 - Vienna, Austria
- PEARC 18 | July 22-27, 2018 - Pittsburgh, PA USA

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

Unidata Science Gateway

We would like to promote and advertise the science gateway (http://science-gateway.unidata.ucar.edu/) to our community.

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

Unidata Transitioning to the Cloud

In the long-term, we would like to explore the possibility of migrating some core Unidata services onto the cloud.

Relevant Metrics

Docker image downloads are available from Unidata's Dockerhub repository.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Making Unidata data streams available via various commercial (e.g., Amazon) and

not-for-profit (e.g., NSF XSEDE Jetstream) cloud services will allow our community to access data quickly and at low or even no cost. Moreover, our users can benefit from high data bandwidth capability provided by various cloud computing platforms, and in some cases, Internet2 capability. Lastly, cloud computing offers the possibility of accessing geoscience data in a "data-proximate" manner where users can perform analysis and visualization on, at times, unwieldy data sets next to where the data reside.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Containerization technology complements and enhances Unidata technology offerings in an open source manner. 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. Unidata staff develop Docker containers in an open-source manner by employing software carpentry best-practices and distributed version control technology such as git.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   Unidata is uniquely positioned in our community to experiment with cloud computing technology in the areas of data discovery, access, and use. Our efforts to determine the most efficient ways to make use of cloud resources will allow community members to forego at least some of the early, exploratory steps toward full use of cloud environments.

4. **Build, support, and advocate for the diverse geoscience community**
   Transitioning Unidata technology to a cloud computing environment will increase data availability to new audiences thereby creating new and diverse geoscience communities.

Prepared *March 2018*

Status Report: Community Services

October 2017 - March 2018

Doug Dirks, Jeff Weber, Joshua Young

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Please suggest potential Triennial speakers by visiting this spreadsheet.

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:

- Kevin Tyle Receives 2017 DeSouza Award
- 2018 Unidata Users Workshop: Save the Date
- Additional DOIs Available for Unidata Technologies
- Models in the Cloud: A Cost Exploration of Cloud Computing for the Atmospheric Sciences
- MetPy Short Course at 2018 AMS Annual Meeting
- Call for Proposals: Unidata 2018 Community Equipment Awards
- Unidata Program Center Welcomes Cece Hedrick
- AMS 2018 Conference Highlights from the Unidata Staff
- Coming Soon: Public Beta of TDS 5.0
- Registration Open for 2018 Unidata Users Workshop
- 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
- Maintaining community involvement - no DeSouza nominees this cycle

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:

- Finalize text of the Unidata Strategic Plan and produce printed version.
- Planning and support for the 2018 Users Committee Triennial Workshop has accelerated. Registration is now open and detailed event planning is in progress.
- Background work on the upcoming core funding proposal is under way.
- Planning for participation in upcoming professional society conferences (2018 AGU Fall Meeting and 2019 AMS Annual Meeting).
- Engagement with CUAHSI to support the NFIE and WRF-Hydro at the NWC.
- Assisting the EarthCube Science Support Office staff with back-office tasks.
- 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:

- Event planning activities for the Users Workshop in June will continue (and increase)
- Work on the Unidata core funding proposal to NSF
- Engagement with EarthCube, ESIP, and 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

New Activities

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

- Expanded emphasis on cloud-related activities
- Support the Users Committee's work towards planning the next Triennial

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

- 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:

- Provide additional data management and 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:

- 53,919 unique pageviews (56,470 in previous period)
- 10.1% of total unique pageviews 10.1% in previous period)

Top community pages

1. All blog pages
   42422 unique pageviews (41969 in previous period)
   79% of total community pageviews (74% in previous period)
2. www.unidata.ucar.edu/community
   4678 unique pageviews (2782 in previous period)
   9% of total community pageviews (5% in previous period)
3. www.unidata.ucar.edu/events
   3888 unique pageviews (8476 in previous period)
   7% of total community pageviews (15% in previous period)
4. www.unidata.ucar.edu/about
   2393 unique pageviews (2660 in previous period)

4% of total community pageviews (5% in previous period)

Social media statistics, March 15, 2018

1. # of Twitter followers: 824 (up from 738)
2. # of Facebook followers: 717 (up from 707)

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to 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. We provide user workshops, tutorials, and community workshops to help build supportive relationships between community members.

2. **Develop and provide open-source tools for effective use of geoscience data**
   We promote Unidata tools and software for multi-disciplinary use, with an eye toward finding additional research and educational communities that can benefit from our work.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   We work with government and industry data providers to secure access to data for Unidata community members.

4. **Build, support, and advocate for the diverse geoscience community**
   We coordinate with our governing committees to find ways to expand Unidata's community participation. We use our web site, electronic newsletters, and social media to keep community members informed about enhanced data services, software tools, and cyberinfrastructure.

   We participate in UCAR/NCAR and NSF projects for underrepresented populations and minority communities (SOARS, AIHEC, outreach to HBCUs). We provide services and tools to facilitate education and research in diverse communities. We work to broaden the Unidata community by participating in student and professional conferences.

Status Report: Community Equipment Awards

Sponsored by the National Science Foundation

October 2017 - March 2018

Admin Group

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Possible theme(s) for the 2019 Unidata Community Equipment Awards;
2. Please consider volunteering to serve on the 2019 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 2018 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.

This year, in conjunction with the Triennial Users Workshop, special consideration will be given to proposals that included one or more of the following:

- Projects that advance data-proximate analysis of large remote datasets (e.g. GOES-16 satellite data).
- Projects that facilitate the use of machine learning techniques and data analytics.
- Installation of equipment that provides student access to and use of GOES-R series satellite data.
- Installation of equipment for the operational use of ensemble models and the development of those models' predictions to share with the broader geoscience community.

A Request for Proposals was sent out on December 19, 2017 with a March 16, 2018 submission deadline. A Review Panel has been formed and will meet after the Users Committee meeting on April 5 at the Unidata Program Center to recommended which proposals should be funded.

Relevant Metrics

Since taking over the management and administration of the Equipment Awards program in 2003 on behalf of the NSF, Unidata has made 92 awards totaling close to $1,200,000.

Prepared March 2018

Status Report: GOES-R/S

October 2017 - March 2018

Mike Schmidt, Tom Yoksas

Questions for Committee Members

- What new coverages and increased spatial and temporal resolutions should be considered for UNIWISC IDD feed imagery?
- Which GOES-16 bands (wavelength channels) and coverages should be converted from netCDF-4 to McIDAS AREA format so that GEMPAK users will continue to be able to use currently satellite imagery?
- What kind(s) of data access methods are most desired/usable for the community?
- Other questions?

Activities Since the Last Status Report

- Implemented IDD distribution of GOES-16 ABI imagery, Space Weather and Geostationary Lightning Mapper (GLM) products

  NCAR/RAL was the first site to connect to this service to provide GOES-16 data for their science objectives. Other sites including two international sites are now receiving feeds of mostly ABI full disk imagery and GLM lightning data.

- Install robust data serving platforms to be used to provide programmatic access to GOES-16 ABI imagery, space weather products and the GLM
  - atm.ucar.edu was the first machine to be replaced by the new equipment.
  - Motherlode.unidata.ucar.edu will be the next machine to be upgraded.

- Implemented UW/SSEC's "fanout server" (redistribution of the GRB-200 UDP unicast stream over TCP)

  We are feeding from one of the SSEC fanout servers, and they are feeding from two that we are running. SSEC is currently experimenting with the creation of a composite GOES-16 datastream that will be composed of the best packets from all GOES-16 ingest systems that are participating in the experiment.

- Secured funding to install GOES-S ingest and data serving capabilities in UCAR

- Dealing with Terrestrial Interference (TI) on the GOES-16 satellite dish we installed at the NCAR Mesa Lab

  In the fall of 2017 we began seeing TI in the signal being received by the GOES-16 dish at the NCAR Mesa Lab. We contacted Quorum Communications to get advice on how to identify the source of the TI and what, if anything, can be done to mitigate it.

Quorum sent an engineer to Boulder (no cost to us) to investigate the cause of the TI, and his findings were that the noise was in-band, so it can not be filtered out. His educated guess about the cause of the TI was that the source of the noise was close to the dish, so it may be from power lines down the hill south of the Mesa Lab or from somewhere in the Mesa Lab itself.

As part of the discussion we had, we described the GVAR GOES and NOAAPort ingest capabilities at the UCAR Foothills Lab 2 (FL-2) location. The engineer's assessment was that we could probably ingest GOES-16 data on the 3.8 m solid dish located outside of the cafeteria at FL-2. He then offered to loan us one of their new GRB LNBs and GRB 200 receivers to use to test this hypothesis.

We installed the loaner LNB on the FL-2 cafeteria dish and installed the GRB 200 in the 2nd floor RAL computer room, and then realigned the dish to point at GOES-16. Ingest results (so far) have been very promising - we are receiving all products in the Left Hand Circularly Polarized (LHCP) channel of the GRB with very few errors. We are unable to ingest the Right Hand Circularly Polarized (RHCP) channel on the cafeteria dish since there is only one signal cable from it to the RAL computer room. We are now talking with UCAR facilities personnel about pulling an additional signal coax (LMR-400) from the cafeteria dish to the RAL computer so that we can also ingest the RHCP channel. In the interim, we have been feeding the RHCP channel from our Mesa Lab GOES-16 satellite dish (the RHCP channel has much fewer errors than the LHCP on the Mesa Lab dish) to the ingest machine for the FL-2 cafeteria dish. Combining the two GRB channels from two dishes has resulted in a very good GOES-16 ingest capability.

Ongoing Activities

We plan to continue the following activities:

- Ingest the GOES ReBroadcast (GRB) from GOES-16 in real-time using the 4.5 m satellite dish that we installed on the eastern pad at the NCAR Mesa Lab using the NOAA funded UW/SSEC/CIMSS Community Satellite Processing Package (CSPP) for Geostationary Data (GEO) package.

- If we can establish error free ingest for both GOES-16 channels on the FL-2 cafeteria dish, we will repoint the Mesa Lab dish to GVAR GOES-West to see if the TI problem disappears.

- Continue uploading all ABI imagery, Space Weather products and Geostationary Lightning Mapper (GLM) to an S3 bucket in Amazon Web Services (AWS)
    - Transfer of ABI imagery and space weather products has been active since the first day of GRB broadcast using storage resources provided free-of-charge by Amazon, but there is a multi-month gap in the data record in the S3 bucket that was caused by a configuration error when we switched from 'rsync'ing the data to moving products via an LDM pattern-acton file action.

- Continue working with NCAR/EOL to support their CSPP GEO installation at the NCAR Mesa Lab
    - This installation will provide an in-house redundancy for GOES-16 ingest
- Canvas the community to learn more about their GOES-16 data needs
- Investigate additions to the IDD CONDUIT data stream that would be useful for creation of new GOES-16 based Level 2 products

Future Activities

- Serve GOES-16 Advanced Baseline Imager (ABI) imagery via McIDAS ADDE and the TDS
    - Serving is already in progress
    - We previously commented that the expected data volume of GOES-16 data would be too large to relay in the IDD. Our opinion was changed when the data volume was drastically reduced by the turning on of netCDF4 compression of the data being received in the GRB
- Investigate approaches that would insulate GRB ingestion from long standing twice per year power downs in the NCAR Mesa Lab facility
- Unidata-Wisconsin (UNIWISC) IDD imagery has been updated to incorporate GOES-16 imagery, but we envision that users would like to see more coverages added

Relevant Metrics

- Lots O Data!

  The volume of GOES-16 imagery, space weather and GLM products can be seen in the real-time statistics plot from our ingest machine:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?EXP+ingest.unidata.ucar.edu

  The volume of just the ABI imagery and GLM Level 2 products can be seen in the real-time statistics plot from NCAR/RAL's ingest machine:

  http://rtstats.unidata.ucar.edu/cgi-bin/rtstats/iddstats_vol_nc?EXP+satops3.rap.ucar.edu

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Standing up ADDE and TDS data services for real-time GOES-R/S data will benefit the greater Unidata community.

2. **Develop and provide open-source tools for effective use of geoscience data**
   The IDD is powered by the Unidata LDM-6 which is made freely available to all. The

Unidata NOAAPort ingest package is being used by a variety of university and non-university community members. Both the LDM and NOAAPort ingest packages are bundled in AWIPS.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The community-driven IDDs provide push data services to users an ever increasing community of global educators and researchers.

4. **Build, support, and advocate for the diverse geoscience community**
   Providing access to data in real-time is a fundamental Unidata activity.

Prepared *March 2018*

Status Report: Internet Data Distribution

October 2017 - March 2018

Steve Emmerson, Mike Schmidt, Jeff Weber, Tom Yoksas

Questions for Committee Members

- Suggestions regarding content of data streams like CONDUIT, FNEXRAD, UNIWISC and LIGHTNING? ... We (UPC, the Unidata community and UAlbany for the NLDN component of LIGHTNING) have control of the content of these data streams, so they are open for suggestions.
- The top level IDD relay clusters we maintain were upgraded to support relay of the full volume of GOES-16 imagery and products being received via redundant GOES-16 downlinks. The data volume for this feed can be seen in the SATELLITE (aka DIFAX) feed in the volume listing for a variety of machines operated by Unidata. The question for the governing committees continue to be if real-time delivery of GOES-16 data via the LDM/IDD is desired, **and** if end typical users can handle the data volume.

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:

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

Maximum hourly volume  83304.423 M bytes/hour
Average hourly volume  55799.609 M bytes/hour

Average products per hour     426604 prods/hour

Feed                            Average             Maximum     Products
                       (M byte/hour)             (M byte/hour)   number/hour
FSL2                    14017.099    [ 25.120%]    17451.845    21357.106
NGRID                    8964.431    [ 16.065%]    13636.916    64253.213
CONDUIT                  8039.861    [ 14.408%]    20861.237    90058.872
DIFAX                    6238.162    [ 11.180%]     8319.725     3187.638
NOTHER                   4874.481    [  8.736%]     7504.662     8342.511
EXP                      4582.119    [  8.212%]     6863.535     3238.660
NEXRAD2                  3833.757    [  6.871%]     5631.066    58824.340
FNMOC                    2264.711    [  4.059%]    14722.657     3318.000
NEXRAD3                  1362.992    [  2.443%]     1864.028    87414.170
HDS                      1103.913    [  1.978%]     1417.177    39476.362
```

```
GEM                      154.485    [  0.277%]      647.338      960.617
UNIWISC                   95.946    [  0.172%]      138.430       49.574
FNEXRAD                   95.338    [  0.171%]      123.683      105.106
NIMAGE                    94.249    [  0.169%]      139.496      122.851
IDS|DDPLUS                76.285    [  0.137%]       87.291    45487.021
LIGHTNING                  1.779    [  0.003%]        7.221      408.234
```

The 6-8 GB/hr value in IDD data volume shown as DIFAX represents the ingestion of GOES-16 data.

Ongoing Activities

We plan to continue the following activities:

- 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-jumbo.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

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

NOAAPort Data Ingest

- Ingest of the upgraded NOAAPort Satellite Broadcast Network (SBN) products and their relay to end-users via the IDD has been “operational” at the UPC since the August 2014.
- The UCAR NOAAPort dish pointing was moved from SES-1 (approx. 101 W) to Galaxy 23 (approx. 89.5 W) in December. Since the move, the indicated Carrier to Noise/EsNo reported by Unidata Novra S300N receivers has improved from the mid-15s to mid-17s

with peaks occasionally being higher than 18.

- 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: Unidata, UW/SSEC, and LSU/SRCC. The NOTHER data stream contains GOES-16 tiles that need to be stitched together to make full image scenes usable to end-user applications. Unidata provides Ryan May's ldm-alchemy package in the Unidat section of Github for this purpose.

  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 v6.13.6, but will soon be updated to v6.13.7.

- Raytheon continues to submit modifications to the LDM for enhanced AWIPS functionality.

Relevant Metrics

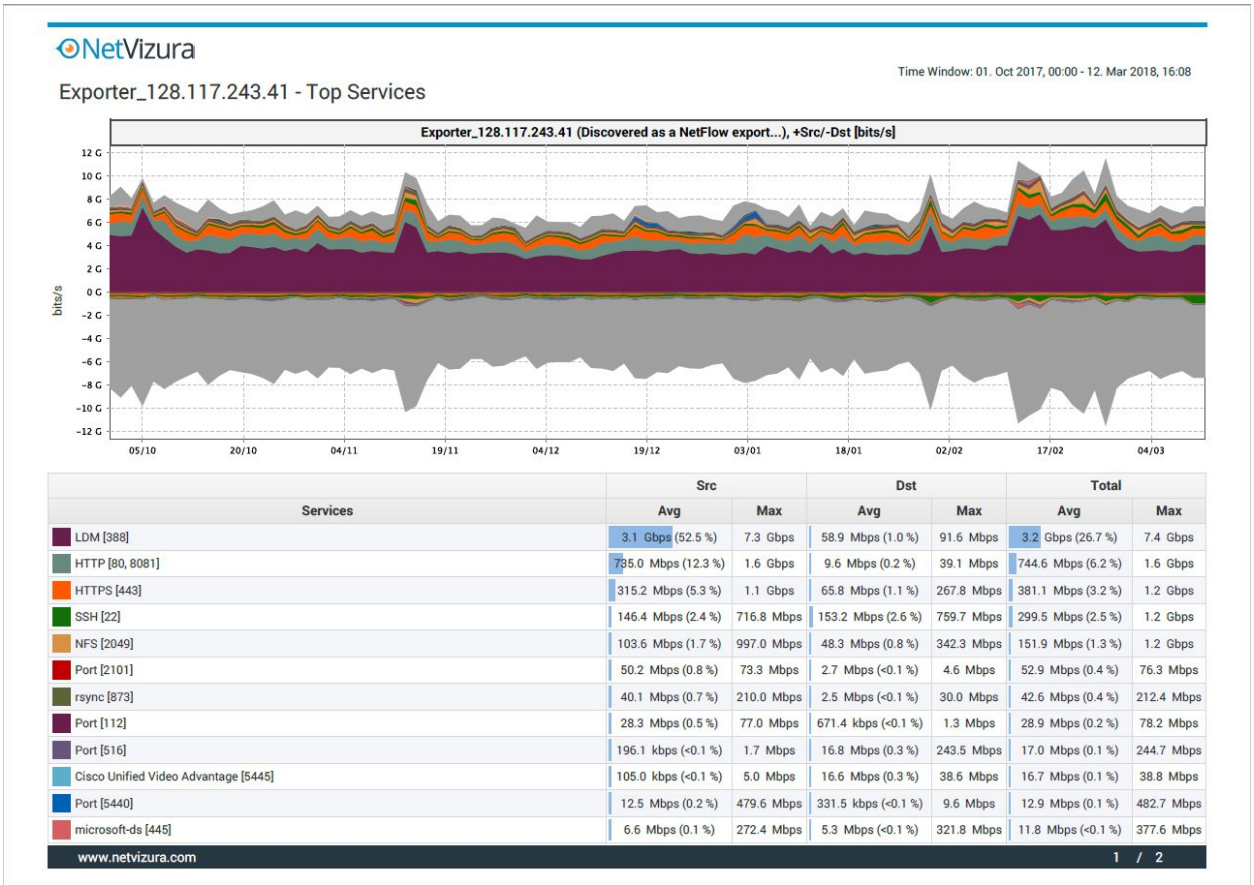
- Approximately **566** machines at **237** sites are running LDM-6 **and** reporting real time statistics to the UPC. Unidata staff routinely assist in the installation and tuning of LDM-6 at user sites as a community service.

- A number of organizations/projects continue 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 1250 downstream connections.

  Over the period from October, 2017 through February, 2018 the average volume of LDM/IDD data flowing from the UCAR/NCAR network averaged around 3.2 Gbps (~34.6 TB/day), and peak rates reached 7.4 Gbps (which would be ~80 TB/day if the rate was sustained).



Services	Src		Dst		Total	
	Avg	Max	Avg	Max	Avg	Max
LDM [388]	3.1 Gbps (52.5 %)	7.3 Gbps	58.9 Mbps (1.0 %)	91.6 Mbps	3.2 Gbps (26.7 %)	7.4 Gbps
HTTP [80, 8081]	735.0 Mbps (12.3 %)	1.6 Gbps	9.6 Mbps (0.2 %)	39.1 Mbps	744.6 Mbps (6.2 %)	1.6 Gbps
HTTPS [443]	315.2 Mbps (5.3 %)	1.1 Gbps	65.8 Mbps (1.1 %)	267.8 Mbps	381.1 Mbps (3.2 %)	1.2 Gbps
SSH [22]	146.4 Mbps (2.4 %)	716.8 Mbps	153.2 Mbps (2.6 %)	759.7 Mbps	299.5 Mbps (2.5 %)	1.2 Gbps
NFS [2049]	103.6 Mbps (1.7 %)	997.0 Mbps	48.3 Mbps (0.8 %)	342.3 Mbps	151.9 Mbps (1.3 %)	1.2 Gbps
Port [2101]	50.2 Mbps (0.8 %)	73.3 Mbps	2.7 Mbps (<0.1 %)	4.6 Mbps	52.9 Mbps (0.4 %)	76.3 Mbps
rsync [873]	40.1 Mbps (0.7 %)	210.0 Mbps	2.5 Mbps (<0.1 %)	30.0 Mbps	42.6 Mbps (0.4 %)	212.4 Mbps
Port [112]	28.3 Mbps (0.5 %)	77.0 Mbps	671.4 kbps (<0.1 %)	1.3 Mbps	28.9 Mbps (0.2 %)	78.2 Mbps
Port [516]	196.1 kbps (<0.1 %)	1.7 Mbps	16.8 Mbps (0.3 %)	243.5 Mbps	17.0 Mbps (0.1 %)	244.7 Mbps
Cisco Unified Video Advantage [5445]	105.0 kbps (<0.1 %)	5.0 Mbps	16.6 Mbps (0.3 %)	38.6 Mbps	16.7 Mbps (0.1 %)	38.8 Mbps
Port [5440]	12.5 Mbps (0.2 %)	479.6 Mbps	331.5 kbps (<0.1 %)	9.6 Mbps	12.9 Mbps (0.1 %)	482.7 Mbps
microsoft-ds [445]	6.6 Mbps (0.1 %)	272.4 Mbps	5.3 Mbps (<0.1 %)	321.8 Mbps	11.8 Mbps (<0.1 %)	377.6 Mbps

Cluster real server backends and accumulator nodes routinely have instantaneous output volumes that approach 2 Gbps. Bonding of pairs of gigabit Ethernet interfaces was needed to be able to support these output data rates. The next generation of cluster machines will need to have 10 Gbps Ethernet capability.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   A project like the IDD demonstrates how sites can employ the LDM to move data in their own environments.

2. **Develop and provide open-source tools for effective use of geoscience data**
   The IDD is powered by the Unidata LDM-6 which is made freely available to all. The Unidata NOAAPort ingest package is being used by a variety of university and non-university community members. Both the LDM and NOAAPort ingest packages are bundled in AWIPS.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The community-driven IDDs provide push data services to users an ever increasing

community of global educators and researchers.

4. **Build, support, and advocate for the diverse geoscience community**
   Providing access to data in real-time is a fundamental Unidata activity.
   The IDD-Brasil, the South American peer of the North American IDD operated by the UPC, is helping to extend real-time data delivery outside of the U.S. to countries in South America and Africa. The Universidad de Costa Rica is actively pursuing IDD relay in the Latin America based IDD-Caribe.

Prepared *March 2018*

Status Report: IDV with RAMADDA

October 2017 - March 2018

Yuan Ho, Cece Hedrick, Julien Chastang

Areas for Committee Feedback

We have no questions at this time.

Activities Since the Last Status Report

IDV System Changes

__Latest netCDF-Java Version__

The version of the netCDF-Java library currently distributed with the latest stable IDV version (5.4) is the 4.6.10 (April 19, 2017). The prior version of netCDF-Java to be distributed with the IDV was 4.6.7. There have been many improvements and bug fixes in that range. The nightly development version of the IDV is currently being distributed with the 4.6.12-SNAPSHOT version netCDF-Java library. The complete release notes for these versions can be found here.

__ISL Changes__

ISL movie and image capture improvements for DRILSDOWN project.

__Java Version__

The IDV is distributed with Java 8 u51. Because of a Windows 10 related bug on some older Intel hardware, this will be the latest version of Java we will be deploying with the IDV for some time.

__Java 3D Version__

The IDV is distributed with Java 3D version 1.6.0 Final. Java 3D 1.7 prerelease is under development and continues to be supported in some capacity by the open source community.

__HTTP to HTTPS__

As the web transitions from HTTP to HTTPS, we update URL references in the IDV (rbi files, plugin files, etc.) to reflect these changes.

__Java Code Refactorings__

Improving Java coding standards in the IDV code base.

IDV Display Changes

__3D Isentropic Analysis__

Isentropic analysis is a very common and effective way to diagnose and visualize vertical motion of the atmosphere, the new isentropic analysis on the constant potential temperature surface helps us better understand the dynamic and stability of weather phenomenons. This new feature allows IDV users to do the isentropic analysis on a regular isobaric gridded dataset without any vertical coordinate conversion.

__GOES 16__

GOES 16 imagery available from ADDE servers on atm.ucar.edu and lead.unidata.ucar.edu. See the IDV Dashboard, Data Choosers Tab, Images node to select these ADDE servers to obtain GOES-16 imagery.

__2D/3D Trajectories__

Continued improvements and enhancements in the underlying VisAD library has enabled a new grid 2D/3D trajectory feature in the IDV. The 2D/3D grid trajectories can be colored by speed or other parameters. The initial area of the trajectory is the entire domain of the dataset. However, regions can be selected with points, rectangles and closed polygons.

There are a variety of grid trajectory types including ribbon, cylinder, deform ribbon, point.

__Latest Version of VisAD__

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

IDV WRF-Hydro Collaboration

Jeff Weber and Yuan are collaborating with David Gochis, and Dave Yates from NCAR-RAL assisting them ensuring WRF-Hydro data is CF compliant regarding Point Feature type. In addition, we are helping them with visualization of this dataset in the IDV.

IDV Release

IDV 5.4 was released in May of 2017.

IDV EarthCube Proposal Awarded

In collaboration, with University of Miami Professor Brian Mapes, Unidata submitted an

EarthCube proposal: “Drilling down from a statistics scatterplot to pre-populated case notebooks”. This proposal was awarded $1.3 million over three years.

Unidata hired Celia Hedrick, a Software Engineer II to work on this project.

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

__2017 American Geophysical Union (AGU) Annual Meeting__

- The Comparison of Point Data Models for the Output of WRF Hydro Model in the IDV
- Understand your Algorithm: Drill Down to Sample Visualizations in Jupyter Notebooks

__2018 American Meteorological Conference (AMS) Annual Meeting__

- The New Trajectory Display in the UNIDATA's IDV
- Synoptic–Dynamic Meteorology in 3D: Introducing an IDV-Based Lab Manual
- Drilling down from Python Statistical Analyses to Rich Interactive Case Study Visualizations, within Jupyter Notebooks

Ongoing Activities

We plan to continue the following activities:

__IDV scatter display__

We will integrate the scatter display feature of the McV into the IDV for imagery dataset and 2D gridded dataset.

__IDV Seam Issues__

We will continue to collaborate with the Unidata netCDF-Java team to devise a solution for long-standing longitude seam issues in the IDV.

__Investigation of Java 3D Alternative__

Because of concerns about the long-term viability 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 few months, we plan to organize or take part in the following:

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. The volume of e-support remains high and constitutes a large fraction of our daily activities. In the last half year the IDV team has closed ~50 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.

__Usage Metrics__

Raw IDV usage metrics, are available here
http://www.unidata.ucar.edu/software/idv/logging/left.html.

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 113 “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 15,000 times compared with 13,200 from six months ago.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to 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. **Develop and provide open-source tools for effective use of geoscience data**
   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. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   RAMADDA allows geoscience specialists the ability to search and publish their IDV bundles on-line. Unidata's RAMADDA installation enables the IDV team to communicate more effectively to our users concerning their IDV issues. Specifically, during support ticket conversations, the IDV team requests that users upload pertinent data to RAMADDA for analysis. One of RAMADDA’s best features is the ability to upload a CDM file and obtain the OpenDAP link from the new entry. The DAP link can be shared and opened in the IDV. RAMADDA also has “server-side view” capability where a specific part of the file system can be made available through the RAMADDA interface. This feature can be helpful to view LDM data feeds, for example. The IDV team also takes advantage of RAMADDA to share instructional IDV screencasts with users.

4. **Build, support, and advocate for the diverse geoscience community**
   Unidata offers yearly multi-day training and occasionally regional workshops for IDV and RAMADDA. The IDV coupled with RAMADDA enables our earth science community partners to distribute geoscience data and metadata through web-based technologies thereby fostering scientific collaborations. Moreover, the IDV's ability to share bundles through RAMADDA creates a scientific social and collaborative network for the geoscience community.

Status Report: Information Technology

October 2017 - March 2018

Mike Schmidt, Matt Perna

Major Activities

GOES-16 issues -- we've been investigating a terrestrial interference (TI) issue with much higher than desirable uncorrectable errors coming from our Mesa Lab (ML) GOES-16 dish. We're working with the receiver equipment manufacturer (Quorum Communications) and they have loaned us a "demo" hardware for installation on a dish outside the Foothills Lab (FL) cafeteria that we used previously for GVAR GOES-East. We've found the GOES-16 signal reception much better at FL, so we're likely to move the GOES-16 ingest to the FL dish and repurpose the ML dish to GOES-S/17 when it becomes operational

motherlode clone upgrade and roadmap -- we've been working on a motherlode clone server refresh and moving machine to machine. We've completed atm.ucar.edu, and should have motherlode complete by the end of the month. After we're done there, we'll be updating equipment hosted at UW SSEC, and then to our R&D server lead.unidata.ucar.edu

100Gb/s networking -- as UCAR upgrades their backbone infrastructure from 10-20Gb/s links to 100Gb/s links, we will be upgrading our data movers (IDD cluster nodes, data aggregators, motherlode clones) from 2 x 1Gbp/s bonded ethernet as required. We're anticipating that the addition of GOES-S data in late 2018 would have mandated the need soon anyway.

UCAR FISMA directions -- 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're having some meetings and generating lots of documents as we move incrementally in that direction.

Security -- as we're all aware, keeping services and systems secure on the network takes consistent attention and occasional herculean efforts (to patch everything all at once). We moved reasonably quickly through the Spectre / Meltdown issue(s) but not all of our systems will have firmware patches for the BIOS. We're keeping track and will be looking to phase those systems out on a more aggressive schedule.

LDM 7 node -- we installed, configured, tested 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.

End user (laptop/desktop) hardware refresh -- we completed the most recent round of hardware refreshes for Unidata staff. We somewhat accelerated the schedule at the recommendation of the NSF Site Visit Team (SVT) , and are very near catching up after being

somewhat behind schedule for several years

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 2018*

Status Report: LDM

October 2017 - April 2018

Steve Emmerson, Tom Yoksas, Mike Schmidt, Yuanlong Tan (UVA)

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:

- Improved building on BSD systems.
- Added support for non-standard installation (i.e., ones that don't have a version-specific installation directory).
- Replaced reliance on the UPC's time-server with the more generic and standard "pool.ntp.org".
- Added software flow control for TTY input to pqing(1), ensured conforming TTY behavior, and improved documentation.
- Improved response to an "ldmadmin stop".
- Added entries to the GEMPAK Grib2 tables for, otherwise, unknown parameters.
- Repurposed the unused DIFAX feed to SATELLITE.

Dependencies, challenges, problems, and risks include:

- Dealing with missing GEMPAK table entries represents a continuing problem area
- The LDM is sometimes held responsible for decisions made by the NWS on how to categorize data products.

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.

Another multi-year grant has been awarded by the NSF for this project. The funds were finally available at the end of September.

Progress has been made on the following:

- Installation of an LDM7 server at the Front Range GigaPop and creation and

verification of an OSI Layer 2 virtual circuit between it and a similar system at the University of Virginia.
- Improved integration of the File Multicast Transfer Protocol into the LDM via 8525 lines of C code and 2395 lines of C++ code. According to the COCOMO model of software development, this represents 2.5 person-years of effort.
- Python code to programmatically create and destroy OSI Layer 2 virtual circuits in Internet 2.
- Documentation of the steps required by participating universities.

Ongoing Activities

We plan to continue the following activities:

- Support and maintenance of the LDM
- Continue adapting the LDM to the Docker container technology to support cloud activities and (perhaps) make life easier for LDM users

New Activities

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

- Install LDM-7 servers at participating universities
- Collect metrics on the LDM7 network

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

- Continued development and deployment of LDM-7

Relevant Metrics

- Data on the LDM package can be found here
- The LDM system at the Unidata Program Center powers the IDD. 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. **Enable widespread, efficient access to geoscience data**
   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.

2. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   By using the LDM to move data into the cloud and developing multicast technologies.

Prepared *March 2018*

Status Report: McIDAS

October 2017 - March 2018

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 investigation has to been the creation of ADDE servers for NOAAPort-delivered GOES-16 satellite imagery.

Current Activities

- Unidata McIDAS version 2017 was released and has been updated 3 times via the **addendum** process

  v2017 includes all SSEC versions up to and including the current McIDAS-X and -XCD releases, both of which are v2017.2.

The latest releases feature the following:

- Updated ADDE servers for Himawari and GOES-R imagery

  A notable feature of core McIDAS support for imagery from platforms such as GOES-R and Himawari is the implementation by SSEC of an Expanded Stretch Feature which includes bit-depth changes to increase the detail shown in certain visible, water vapor and short-wave infrared imagery.

- ADDE server for NOAAPort GOES-R Sectorized Cloud and Moisture (SCMI) imagery

  The core McIDAS ADDE servers do not support the tiled GOES-16 image sectors that are being distributed in the NOAAPort SBN. New ADDE servers have been developed to support these NOAAPort-delivered images.

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.

New Activities

Ongoing Activities

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

- Bandwidth usage by the Unidata McIDAS ADDE servers routinely exceeds 8 TB/month.

  The amount of data served by Unidata ADDE instances increased substantially with the availability of GOES-16 ABI imagery and GLM lightning data.

- 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. This package is in the process of being updated to support changes to various data streams.

Geostationary Satellite Data Ingest and Data Serving

Unidata continues to ingest GOES-East and GOES-West imager data at the UCAR Foothills Lab campus in Boulder.

- Direct, programmatic access to real-time GOES-East (GOES-16) data via McIDAS ADDE has been averaging approximately 5.6 TB/month ***even though this service has not yet been announced yet***.

- Direct, programmatic access to real-time GOES-West (GOES-15) data via McIDAS ADDE routinely exceeds 2.6 TB/month.

Planned Activities

Ongoing Activities

Continued ingest and serving of GOES-East and GOES-West imagery from the current constellation of GOES GVAR platforms. This effort requires maintenance of the satellite ingest and data serving equipment.

New Activities

Finish installing the data ingest hardware for GOES-16 downlink/data distribution facility at the Foothill Lab 2 location. This new activity is in progress.

Continue moving GOES-R imagery and products to an S3 bucket in AWS in near real-time.

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   *McIDAS remains **the** application of choice for the satellite meteorology community.*
   The Advanced 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.

2. **Develop and provide open-source tools for effective use of geoscience data**
   The fifth generation of McIDAS, McIDAS-V, unlike its predecessors, is a fully open source application that is is in wide scale and growing use in the worldwide satellite meteorology community.
   McIDAS ADDE continues to evolve and provide access to increasing volumes of image and

non-image data.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   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 in the SSEC Data Center are currently serving over 1 TB per day.

4. **Build, support, and advocate for the diverse geoscience community**
   McIDAS is sought for use by those interested in satellite meteorology worldwide.

Prepared *March 2018*

Status Report: netCDF

October 2017 - March 2018

Ward Fisher, Dennis Heimbigner

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. To what extent is Amazon S3 used within your organization? Would you benefit from native netCDF support for S3 storage?
2. Are there other cloud-based block storage formats/locations (zarr, Azure, etc) that are actively in use? That we should consider storing.
3. Are there any emergent avenues (stack overflow, etc) for user support which the netCDF team should investigate?
4. 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 108 open issues for netCDF-C, 23 open issues for netCDF-Fortran, and 16 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:

- Addition of user-defined compression filters.
- Work towards enhanced parallel I/O.
- Further extension of the netCDF build-and-test platforms using Docker technology.
- Further enhancements to the netCDF documentation.
- 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 an uptick in the number 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.
- Increased support for native CDF5 on 64-bit platforms, potentially obviating the need for the parallel netcdf library for reading/writing CDF5 files in serial fashion.

Dependencies, challenges, problems and risks include:

- Small group (and shrinking) of developers for supporting large project.
- Dependency on HDF5, controlled by external group.
- Slow progress in user adoption of netCDF-4 features.

Ongoing Activities

We plan to continue the following activities:

- 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.
- Improve organization of Doxygen-generated documentation for netCDF-C and Fortran libraries.
- Extend collaboration as opportunities arise, for increasing the efficiency of parallel netcdf-3 and netcdf-4.

New Activities

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

- Seek out, and prepare material for upcoming, conferences and other outreach opportunities.
- Work on reducing the defects reported by static analysis.
- Release the next versions of netCDF-C, netCDF-Fortran, netCDF-C++.
- Modernize the netCDF documentation to provide easy access to documentation for older versions of netCDF.
- Provide thread-safety for the netCDF C library.

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.
- Participate in development of new CF 2.0 conventions for climate and forecast simulation output and observational data in netCDF-4 form.
- Continue to encourage and support use of netCDF-4's enhanced data model by third-party developers.
- Create and release online educational material in the form of Youtube video tutorials for using netCDF.

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

- Implement support for Amazon S3 in the netCDF C library.
- Improve scalability to handle huge datasets and collections.
- Improve the efficiency of parallel netcdf3 and parallel netcdf4.

Relevant Metrics

There are currently about 183,700 lines of code (up from 142,810 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 been increased slightly from **0.36** six months ago to **0.7** today. 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**. The jump in defect density is a result of the addition of the **DAP4** code. As this is new code, the initial defects are still being worked out.

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:

- **946,000** for netCDF-3
- **937,000** for netCDF-4
- **507,000** for HDF5
- **94,600** for GRIB2

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

- **307** for netCDF-3
- **593** for netCDF-4
- **10,200** for HDF5
- **848** for GRIB2

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**

by developing netCDF and related cyberinfrastructure solutions to facilitate local and remote access to scientific data.

2. **Develop and provide open-source tools for effective use of geoscience data**
   by supporting 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.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   by developing useful data models, frameworks, and protocols for geoscience data; advancing geoscience data and metadata standards and conventions; and providing information and guidance on emerging cyberinfrastructure trends and technologies.

4. **Build, support, and advocate for the diverse geoscience community**
   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 2018*

Status Report: Outreach Activities

October 2017 - March 2018

Ethan Davis, Sean Arms, Jeff Weber, Josh Young, & Tom Yoksas

Areas for Committee Feedback

None at this time.

Activities Since the Last Status Report

Open Geospatial Consortium (OGC) activities

Continue to work with OGC to augment international CF-netCDF standards that have been established over the last several years.

Progress has been made on the following:

- Represent Unidata at OGC Technical Committee meetings.
- Chair OGC NetCDF Standards Working Group
- Serve as UCAR business and technical representative to the OGC
    - UCAR will be co-host of the June 2018 OGC TC mtg, CSU is host (reception event to be held at the NCAR-Wyoming Supercomputing Center)

Dependencies, challenges, problems, and risks include:

- Limited time available to follow and affect relevant OGC activities.

EarthCube Cyberconnector Project

Collaborative project with George Mason University to make Unidata real time datasets available to researchers and educators in other disciplines

Progress has been made on the following:

- Maintained periodic contact with Liping Di, the GMU PI on the project.
- Did mid-course correction to objectives based on one year no cost extension to the project.

Dependencies, challenges, problems, and risks include:

- The one year extension enables us to use the remaining resources available in the project to accomplish tasks that are in line with the strategic goals for the new version of the TDS and to ensure that our data are available via standard interfaces such as those provided by the Data Discovery and Access Broker of the Italian CNR group

Python Workshop Notebooks on alternative platforms.

Due to reduced FTE commitment, no effort is being expended on cloud software development

Ongoing Activities

- Coordination and collaboration with NCAR GIS (e.g. advising the GIS Program on cloud workflows)- Josh Young and Jeff Weber
- Represent Unidata and UCAR in OGC - Ethan Davis
- Represent Unidata in ESIP - Ethan Davis
- Represent Unidata at CUASHI - Jeff Weber
- Continue as co-PI on Earthcube Cyberconnector project - Sean Arms
- Participate in ODIP 2 as resources allow - Mohan Ramamurthy
- Participate in the GOES-16, Pathfinder, and VIIRS data working groups at AWS - Jeff Weber
- Imagery from UCAR GOES East/West ingest systems continue to be routinely accessed by international users in North, Central and South America using McIDAS-X, IDV, and McIDAS-V.
- Even though access to GOES-16 imagery and products has not been announced (we are waiting for GOES-16 data to be declared operational), we are seeing substantial access to the imagery via ADDE on our ingest machine.
- Use of Unidata tools, especially netCDF, the IDV and AWIPS continues to grow internationally. Use of GEMPAK internationally seems to have plateaued.

Renewed effort aimed at creating an IDD-Caribe

Dr. Marcial Garbanzo Salas of the Universidad de Costa Rica continues to lead an effort aimed at making an IDD-Caribe a reality. Unidata is playing a supporting role in this effort.

New Activities

Explore CloudIDV product generation at the National Water Center

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Work with representatives of other disciplines and serve on their governing boards where appropriate, e.g., NCAR GIS for Geographic Information Systems, CUAHSI (Consortium of Universities for Advancement of Hydrological Science), ODIP (Ocean Data Interoperability Platform), etc.
2. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   Continue to work with Open Geospatial Consortium (OGC) to augment international CF-netCDF standards that have been established over the last several years.

3. **Build, support, and advocate for the diverse geoscience community**
   Serve as Co-Investigator on Earthcube Cyberconnector project which will make Unidata data available to a wide range of research and education communities beyond the traditional Unidata community.

Prepared *March 2018*

Status Report: Python

October 2017 - March 2018

Ryan May, John Leeman, Sean Arms, Julien Chastang

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. Given that the rest of the scientific Python community is starting to drop support for Python 2.7 (core Python developers will stop supporting 2.7 in 2020, matplotlib summer 3.0 release will only support Python 3, etc.), what is your feeling on when we should begin to make the same transition?
2. How rapidly should we be dropping old versions of other dependencies (e.g. numpy, matplotlib, etc.)? Are there suggestions on a time window or other criteria to use?
3. Are there any additions you'd like to make to MetPy's or Siphon's roadmap?

Activities Since the Last Status Report

Python Training Efforts

Python training efforts continue to be a valuable portion 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.

Progress has been made on the following:

- In November 2017, Ryan May and John Leeman taught an intermediate Python workshop sponsored by Eric Bruning at Texas Tech university. This represented an opportunity to teach concepts beyond those normally done. Feedback was overwhelmingly positive.
- Ryan May and John Leeman, together with Kevin Goebbert, taught a short course using MetPy at the 2018 AMS Annual Meeting. This course taught practical application of MetPy (and Siphon) to dynamic meteorology. The course had 15 attendees, including 10 from educational institutions; feedback was very positive. Given the success, we intend to submit another course for AMS 2019.
- Ryan May and John Leeman will be teaching a tutorial on MetPy and Siphon at the UCAR Software Engineering Assembly Conference on April 6th. This will largely be based on the materials from the AMS short course.
- John Leeman continues to lead the "MetPy Mondays" effort. These weekly blog posts and screencasts on the Unidata Developers' blog receive a lot of attention and feedback. Creating these also often uncover improvements for our software.
- John Leeman, together with Ryan May and a team of university collaborators, submitted a proposal to NSF's CyberTraining solicitation to develop a set of online

cyberinfrastructure training materials, focused on Python. The idea is to produce a set of materials suitable for Teaching Assistants to include in their classroom as well as used for “at your own pace” independent learning.

MetPy

MetPy continues to grow, both in features and in community. The pace of externally-driven Pull Requests has become more regular, averaging approximately one per week. Feedback at this past AMS Annual Meeting reflects this trend, as name recognition of MetPy was much greater than in the past. Many people came by the Unidata booth and our posters wanting to hear more about MetPy.

Development going forward will continue to be driven by requirements for our dedicated awards (in addition to bug reports and pull requests from community members). The primary efforts will be focused around improved units support, integration with xarray, and growing the calculation collection. We have already developed some prototype integration with xarray, handling projection metadata, and the success of this work reinforces the idea that xarray integration will be a great benefit to users of MetPy. We hope to release some of this effort in the 0.8 release this spring.

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 471 followers. Involvement and support requests continue to grow as well (e-support, Gitter chat), including quite a few contributions on GitHub.
- MetPy 0.6.1 bug fix release
- MetPy 0.7 released, including quite a few new calculations, as well as adding support for derivative-based calculations on irregularly-spaced grids. We also made quite a few minor changes to the API as we clean up shortcomings from early development. Most importantly, the 0.7 release saw contributions from 9 external contributors.
- We have created special “MetPy Contributor” stickers, and have sent them to all those who have contributed to MetPy in the past. Feedback on social media has been very positive, helping to increase visibility of the project.
- Work towards requirements of MetPy-related NSF awards
- Piotr Flatau from the University of California at San Diego has approached us about supporting the use of MetPy for the Equatorial Line Observations field project. The desire is to use MetPy (and Siphon) to create some basic graphics in support of field project operations. Scripts developed by Unidata in support of this effort will be added to our Python Gallery so that the rest of the community benefits as well.

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 to remote data. Siphon has begun to see an increase in community contributions, which is likely due in part to the inclusion of data sources beyond TDS. We are planning releasing 0.7 this spring, including community contributions as well as support for retrieving upper air data from the Iowa Environmental Mesonet archive and a list

of THREDDS servers (e.g. Unidata, NCEI) accessible programmatically.

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 GOES16 imagery. As we transition more of our internal data processing to Python, this repository will hold those scripts.

Progress has been made on the following:

- Expanding Siphon's scope beyond TDS to a variety of useful atmospheric science datasets.
- Siphon 0.6.1 bug fix release, including fixes for RAMADDA catalogs and for NCSS on TDS 5.0
- Warren Pettee contributed a client for the Applied Climate Information System (ACIS) run by NOAA's Regional Climate Centers
- Daniel Watkins has contributed support for downloading upper air data from the Integrated Global Radiosonde Archive, Version 2 (IGRA2)
- The NOAAPORT GOES16 processing code seems to be leveraged by several members of the community, based on the bug reports received when the feed is modified by NOAA.

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.

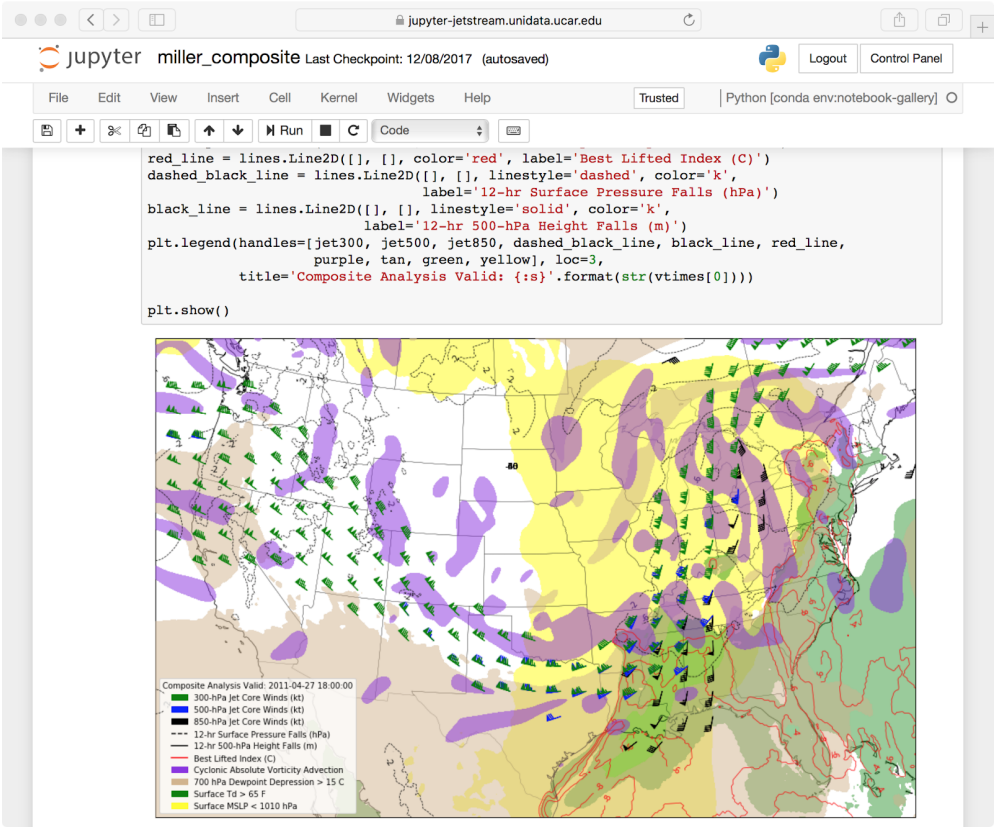
Progress has been made on the following:

- Ryan May was invited to present on MetPy at the "Workshop on developing Python frameworks for earth system sciences" hosted by ECMWF. ECMWF are greatly increasing their use of Python, and it was valuable to engage with them and other users of Python in the international community.
- Ryan May and John Leeman will also be attending SciPy 2018, with Ryan also now co-chairing the Sprints organizing committee. Ryan and John will be speaking about MetPy and (tentatively) presenting a tutorial on open source project infrastructure.
- 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.
- We have been approached by Enrico Fucile from ECMWF about participating in the development of GRIB3. As they finalize the standard, they are very interested in having independent implementations of the specification; we will be developing a pure

Python implementation that will eventually be included in MetPy. This will likely pave the way for read support for other GRIB versions in MetPy as well.

- Ryan May has become a member of the CartoPy steering committee (with commit privileges) and continues participating on the matplotlib core development team.
- We also continue to actively engage with the xarray, numpy, and pint projects
- Enhanced CartoPy to better support parameters from netCDF CF metadata
- Fixed a variety of issues in matplotlib's macOS backend

JupyterHub Server on Jetstream



JupyterHub is a multi-user server for running Jupyter notebooks and other technologies. We deployed a JupyterHub server with pre-installed Unidata Python notebooks on the NSF-funded Jetstream cloud. The aim here is to provide a zero install solution for running Unidata notebooks for our community in a classroom setting, for example. We are improving upon our demonstration server by obtaining invaluable feedback from our Unidata User Committee collaborators: Kevin Goebbert, Kevin Tyle (SAC), Pete Pokrandt, Rich Signell, Shawn Riley, Warren Pettee (UserComm alum) and Unidata collaborator Eric Bruning.

While this demonstration server is a promising start with, at times, frequent daily use from some testers, it will not scale to a classroom setting or beyond. We must continue researching and employing technologies that allow us to better take advantage of the elasticity of the cloud such as Kubernetes and the Zero to Jupyterhub project. Part of our Jetstream XSEDE allocation includes Extended Collaborative Support Services (ECSS) staffing resources. With

those staffing resources, we are collaborating with Semir Sarajlic (Georgia Tech) and Suresh Maru (Indiana University) to implement a JupyterHub server using of Kubernetes technology to better take advantage of the horizontal scalability of the cloud. We will be presenting our work at the PEARC '18 conference in Pittsburgh, PA.

Recognizing that Unidata time and resources are limited, we have heavily leveraged open-source solutions (e.g., JupyterHub Docker container) for this server and simply deploying, configuring and customizing those open source technologies, resulting minimal in-house development.

In the past six months, we have done two presentations on this work:

- 2017 AGU Fall Meeting | December 11–15, 2017 – New Orleans, LA USA
- 2018 AMS Annual Meeting | January 7–11, 2018 – Austin, TX USA

Python Data Access Framework (python-awips)

An early version of the AWIPS Data Portal has been published. See the AWIPS and GEMPAK report for details.

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
- As resources and time permit, continue growing the Online Python Training project by writing Jupyter notebooks specifically targeted towards teaching the geoscience community programming concepts.

New Activities

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

- Work on native Python support for decoding GRIB3 formatted files
- Attend SciPy 2018

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

- Teach regional workshop at Jackson State University
- Teach another short course on MetPy at AMS 2019

- Present annual update on Python libraries at AMS 2019

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
- Restructure our annual Python workshop to be a full week with introduction to Python/git, intermediate with MetPy/Siphon/etc., and developer hack-day

Relevant Metrics

MetPy

- 97% test coverage
- Watchers: 37
- Downloads for the last year (only Conda for now):
    - 0.5.0: 2103
    - 0.5.1: 2913
    - 0.6.0: 1575
    - 0.6.1: 1289
    - 0.7.0: 1702
- Since 1 October 2017
    - Active Issues: 179 (106 created, 60 closed)
    - Active PRs: 100 (91 created, 91 closed)
    - External Issue Activity: 22 opened, 135 comments
    - External PR Activity: 36 opened, 72 comments
    - Unique external contributors: 40
    - Stars: 50 (246 total)
    - Forks: 3 (113 total)
    - Commits: 280
- Since 1 April 2017
    - Active Issues: 253 (192 created, 117 closed)
    - Active PRs: 219 (196 created, 194 closed)
    - External Issue Activity: 41 opened, 187 comments
    - External PR Activity: 47 opened, 91 comments
    - Unique external contributors: 56
    - Stars: 116 (246 total)
    - Forks: 8 (113 total)
    - Commits: 588

Siphon

- 97% test coverage
- Watchers: 13
- Downloads for the last year (only Conda for now):
    - 0.4.1: 3670
    - 0.5.0: 1987

    - 0.6.0: 1612
    - 0.6.1: 3039
- Since 1 October 2017:
    - Active Issues: 21 (16 created, 7 closed)
    - Active PRs: 17 (17 created, 16 closed)
    - External Issue Activity: 5 opened, 8 comments
    - External PR Activity: 4 opened, 9 comments
    - Unique external contributors: 7
    - Stars: 14 (65 total)
    - Forks: 1 (28 total)
    - Commits: 71
- Since 1 April 2017
    - Active Issues: 51 (34 created, 23 closed)
    - Active PRs: 31 (31 created, 30 closed)
    - External Issue Activity: 7 opened, 10 comments
    - External PR Activity: 6 opened, 9 comments
    - Unique external contributors: 13
    - Stars: 24 (65 total)
    - Forks: 2 (28 total)
    - Commits: 132

Unidata Python Workshop

- Watchers: 19
- Since 1 October 2017
- Watchers: 23
    - Active Issues: 42 (19 created, 25 closed)
    - Active PRs: 25 (24 created, 24 closed)
    - External Issue Activity: 1 opened, 0 comments
    - External PR Activity: 2 opened, 6 comments
    - Unique external contributors: 2
    - Stars: 12 (73 total)
    - Forks: 1 (59 total)
    - Commits: 85
- Since 1 April 2017
    - Active Issues: 83 (62 created, 65 closed)
    - Active PRs: 80 (76 created, 78 closed)
    - External Issue Activity: 1 opened, 2 comments
    - External PR Activity: 2 opened, 6 comments
    - Unique external contributors: 3
    - Stars: 24 (73 total)
    - Forks: 2 (59 total)
    - Commits: 222

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   Python can facilitate data-proximate computations and analyses through Jupyter Notebook technology. Jupyter Notebook web servers can be co-located to the data source for analysis and visualization through web browsers. This capability, in turn, reduces the amount of data that must travel across computing networks.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Our current and forthcoming efforts in the Python arena will facilitate analysis of geoscience data. This goal will be achieved by continuing to develop Python APIs tailored to Unidata technologies. Starting with the summer 2013 Unidata training workshop, we developed an API to facilitate data access from a THREDDS data server. This effort has been encapsulated with the new siphon project, which is an API for accessing remote data, including the THREDDS data server. Moreover, Python technology coupled with the HTML5 Jupyter Notebook technology has the potential to address "very large datasets" problems. Jupyter Notebooks can be co-located to the data source and accessed via a web browser thereby allowing geoscience professionals to analyze data where the data reside without having to move large amounts of information across networks. This concept fits nicely with the "Unidata in the cloud" vision and the goals outlined Unidata 2018 Five-year plan. Lastly, as a general purpose programming language, Python has the capability to analyze and visualize diverse data in one environment through numerous, well-maintained open-source APIs. The additional development of MetPy fills the need for domain-specific analysis and visualization tools in Python.

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The TDS catalog crawling capabilities found in siphon will facilitate access to data remotely served by the Unidata TDS, as well as other TDS instances around the world.

4. **Build, support, and advocate for the diverse geoscience community**
   Based on interest from the geoscience community, Unidata, as part of its annual training workshop, now hosts a three day session to explore Python with Unidata technology. Also, to advance the use of NetCDF in Python, Unidata has promoted Jeff Whitaker's NetCDF4-python project, including hosting its repository under Unidata's GitHub account. Unidata is initiating a project to provide online Python training specifically targeting geoscience students. Unidata is also fostering some community development of meteorology-specific tools under the MetPy project.

Prepared *March 2018*

Status Report: Support

October 2017 - March 2018

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 your needs?

Activities Since the Last Status Report

Training

- In 2018 The UPC will be focusing its in-person training efforts on regional workshops
- 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

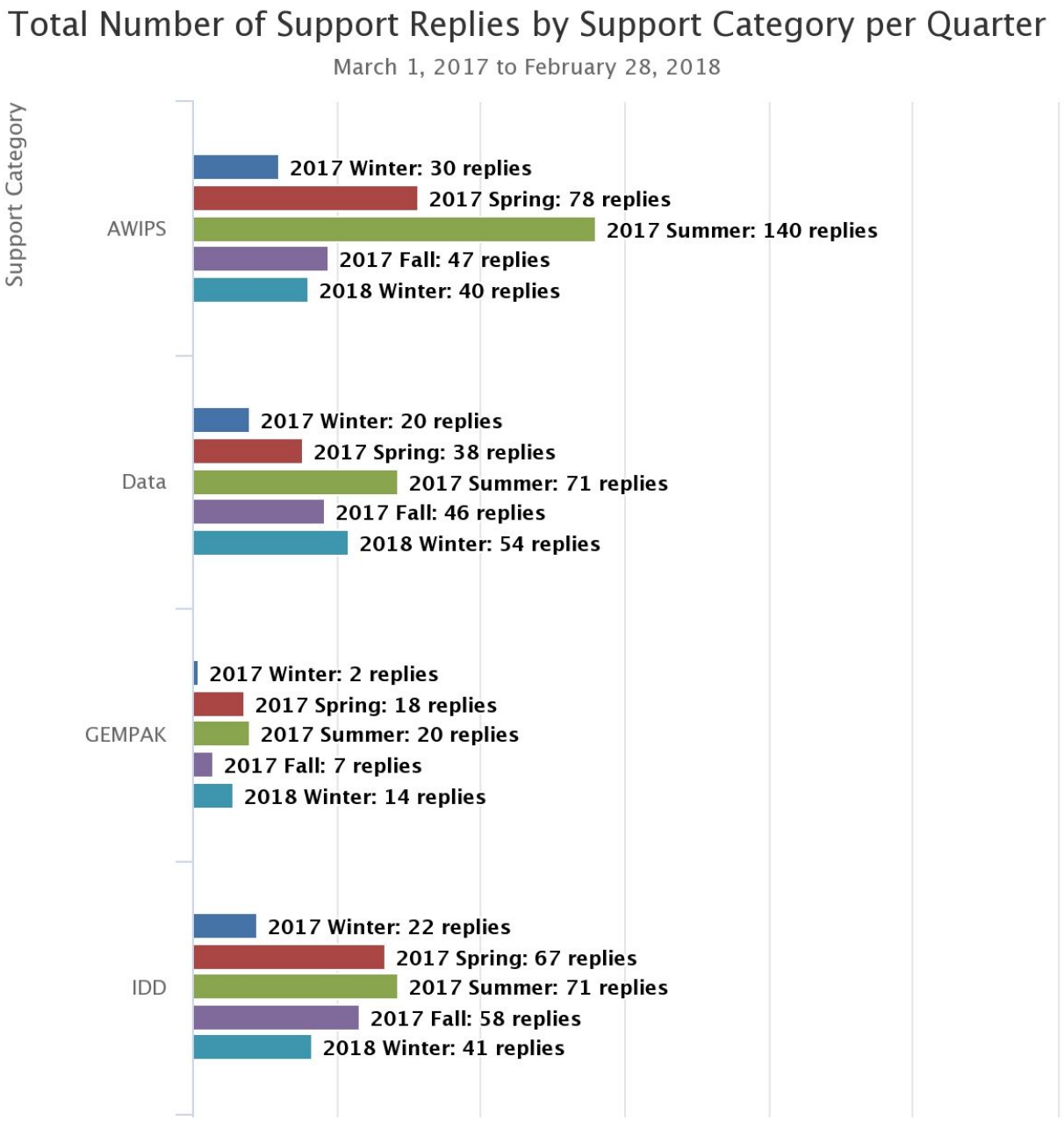
Relevant Metrics

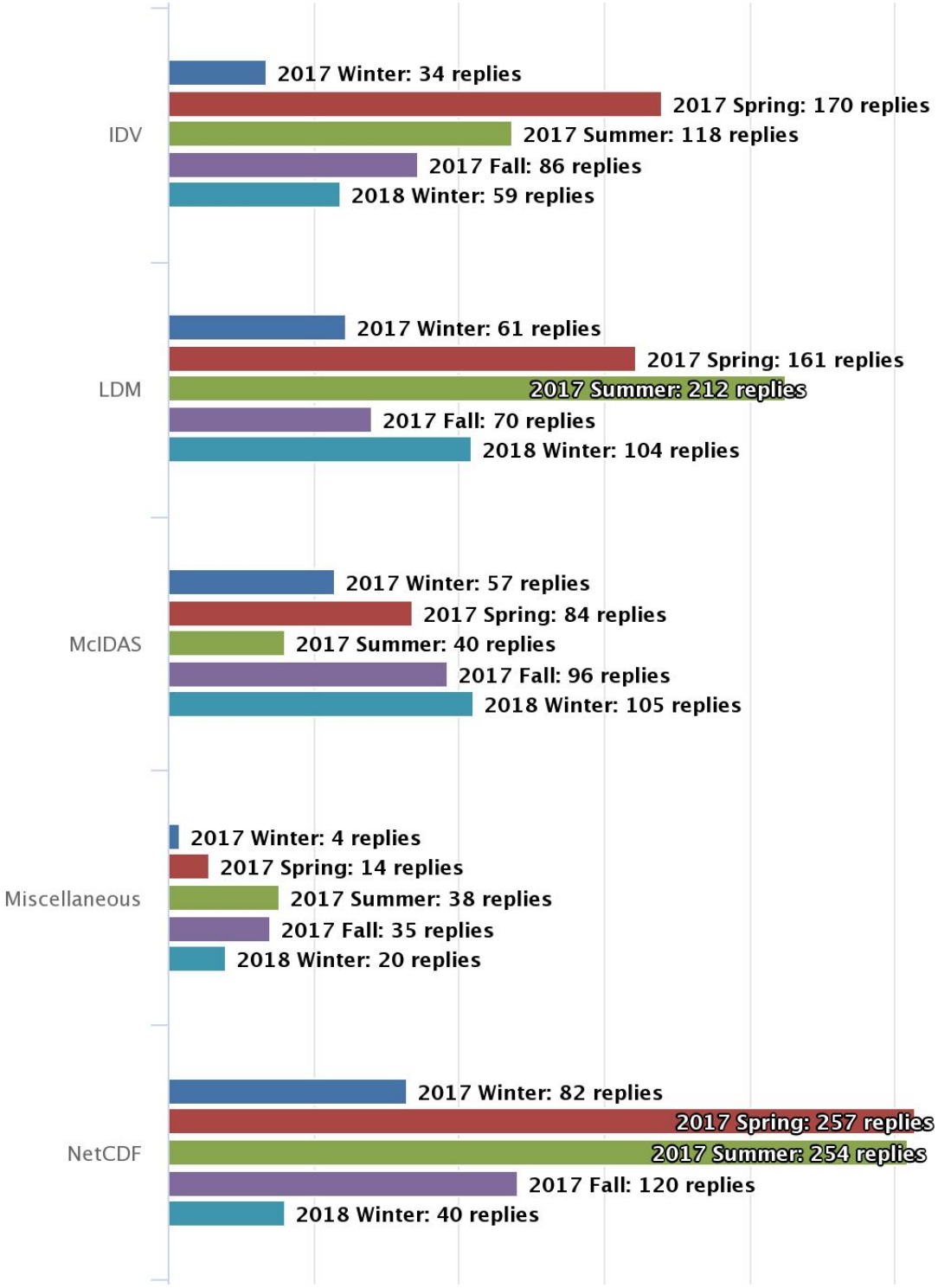
Since January 26, 2006 over 55350 user support "transactions" (new inquiries and follow-ups) have been processed through the Unidata inquiry tracking system. Other methods of providing answers to questions posed (e.g., Github, Stack Overflow, Jira, mailing list replies, etc.) add substantially to the support load.

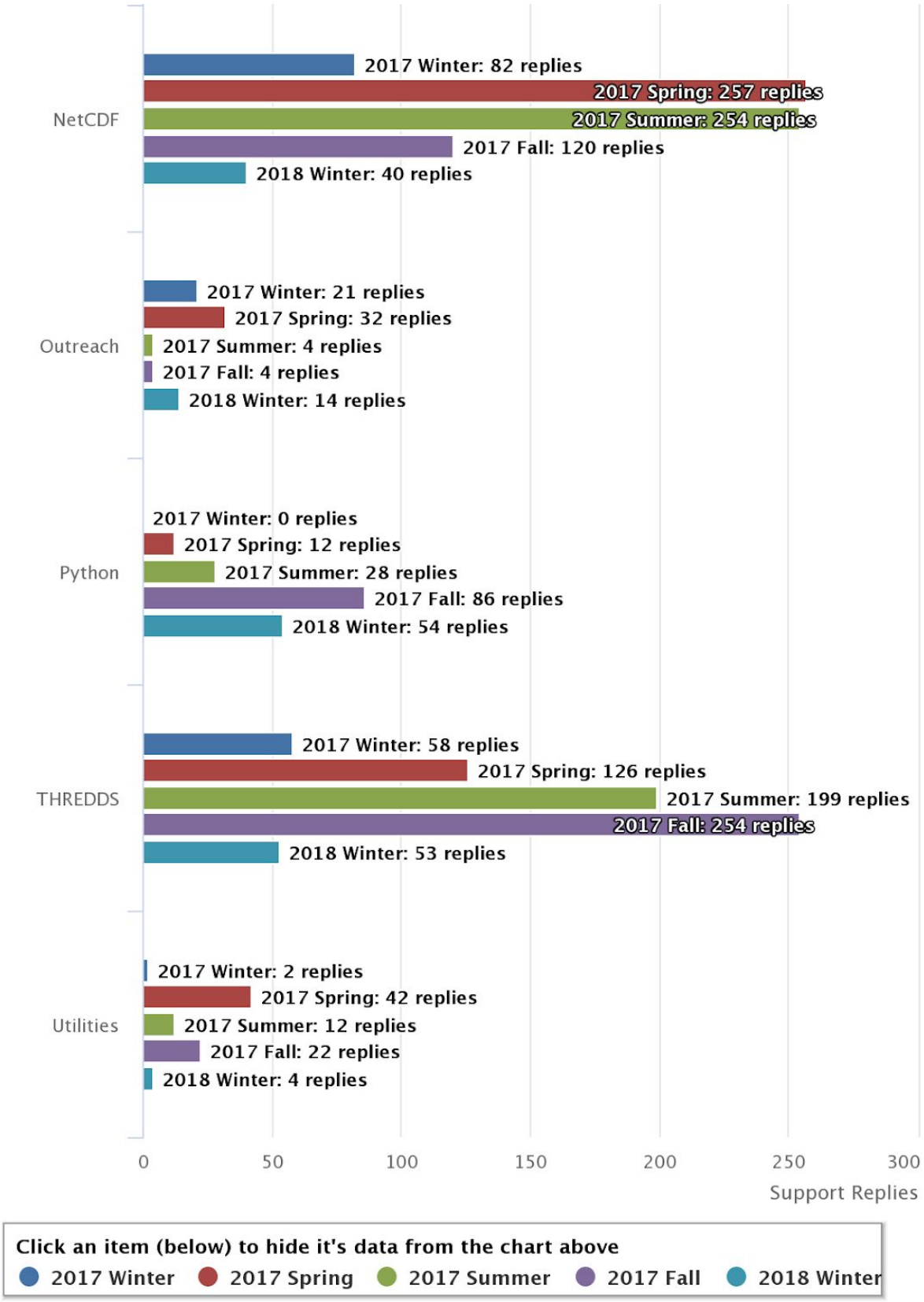
Additional metrics may be found in the Comprehensive Metrics Data portion of this meeting's agenda.

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 14 month period from March 1, 2017 until February 28, 2018.

The quarters shown are defined as: Winter: January, February, March; Spring: April, May, June; Summer: July, August, September; Fall: October, November, December.


Total Number of Support Replies by Support Category per Quarter
March 1, 2017 to February 28, 2018
Support Category
AWIPS
2017 Winter: 30 replies
2017 Spring: 78 replies
2017 Summer: 140 replies
2017 Fall: 47 replies
2018 Winter: 40 replies
Data
2017 Winter: 20 replies
2017 Spring: 38 replies
2017 Summer: 71 replies
2017 Fall: 46 replies
2018 Winter: 54 replies
GEMPAK
2017 Winter: 2 replies
2017 Spring: 18 replies
2017 Summer: 20 replies
2017 Fall: 7 replies
2018 Winter: 14 replies
IDD
2017 Winter: 22 replies
2017 Spring: 67 replies
2017 Summer: 71 replies
2017 Fall: 58 replies
2018 Winter: 41 replies


IDV
2017 Winter: 34 replies
2017 Spring: 170 replies
2017 Summer: 118 replies
2017 Fall: 86 replies
2018 Winter: 59 replies
LDM
2017 Winter: 61 replies
2017 Spring: 161 replies
2017 Summer: 212 replies
2017 Fall: 70 replies
2018 Winter: 104 replies
McIDAS
2017 Winter: 57 replies
2017 Spring: 84 replies
2017 Summer: 40 replies
2017 Fall: 96 replies
2018 Winter: 105 replies
Miscellaneous
2017 Winter: 4 replies
2017 Spring: 14 replies
2017 Summer: 38 replies
2017 Fall: 35 replies
2018 Winter: 20 replies
NetCDF
2017 Winter: 82 replies
2017 Spring: 257 replies
2017 Summer: 254 replies
2017 Fall: 120 replies
2018 Winter: 40 replies


NetCDF
2017 Winter: 82 replies
2017 Spring: 257 replies
2017 Summer: 254 replies
2017 Fall: 120 replies
2018 Winter: 40 replies
Outreach
2017 Winter: 21 replies
2017 Spring: 32 replies
2017 Summer: 4 replies
2017 Fall: 4 replies
2018 Winter: 14 replies
Python
2017 Winter: 0 replies
2017 Spring: 12 replies
2017 Summer: 28 replies
2017 Fall: 86 replies
2018 Winter: 54 replies
THREDDS
2017 Winter: 58 replies
2017 Spring: 126 replies
2017 Summer: 199 replies
2017 Fall: 254 replies
2018 Winter: 53 replies
Utilities
2017 Winter: 2 replies
2017 Spring: 42 replies
2017 Summer: 12 replies
2017 Fall: 22 replies
2018 Winter: 4 replies
0
50
100
150
200
250
300
Support Replies
Click an item (below) to hide it's data from the chart above
2017 Winter
2017 Spring
2017 Summer
2017 Fall
2018 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, Systems
NetCDF	Support LibCF, Support netCDF
Outreach	Outreach, Polcomm, Support Egrants, Support News, Support Outreach, Support Workshop, Usercomm
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 remains substantial, and yearly totals have been relatively constant for the past several years. Overall support activities vary by somewhat by quarter. Spikes in support for individual packages is largely correlated with the release of new distributions.
- 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 year, probably due to GEMPAK users investigation use of AWIPS and Python/MetPy.
- Support for AWIPS has steadily increased and has exceeded that for GEMPAK over the past year.
- 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. **Enable widespread, efficient access to geoscience data**

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

2. **Build, support, and advocate for the diverse geoscience community**
   The user support provided by the UPC is recognized throughout the atmospheric science community. Unidata's outreach efforts are routinely called out in surveys of the NCAR/UCAR community.

Prepared *March 2018*

Status Report: THREDDS

October 2017 - March 2018

Sean Arms, Ethan Davis, Dennis Heimbigner, Ryan May, Christian Ward-Garrison

Areas for Committee Feedback

We are requesting your feedback on the following topics:

1. If we were to upgrade thredds.ucar.edu to TDS 5 at the beginning of August, how would this impact your Fall 2018 classes? Would you have time to tests your course resources that utilize the server?

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 a 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 / TDS version 4.6.11 (Stable)

Progress has been made on the following:

- The 4.6.x line of development is now in maintenance mode so that the team can focus on v5.0. “Maintenance mode”, which includes user support and bug fixes, continues to take up quite a bit of resources.
- The THREDDS team now conducts automated security scans on our dependencies to ensure that we are not using external libraries with open vulnerabilities.

Focus on netCDF-Java / TDS (~~Soon-to-be~~ Beta) v5

We have hoped to have the beta out *real soon now*™ for quite some time, and are happy to say a beta is set to be released on March 16th. While there are known bugs in this beta, as well as both unknown and “unknown unknown” bugs, this represents a big step forward for the project. It is our intention that TDS v5 will be released by the end of summer.

Progress has been made on the following:

- The Nexus Repository Manager at https://artifacts.unidata.ucar.edu has been

upgraded from version 2 to version 3 and it will now host **all** build artifacts. For users, this means:
    - Developers should update their software builds to pull netCDF-Java/TDS artifacts from the Maven repository at https://artifacts.unidata.ucar.edu/repository/unidata-all/.
    - Documentation for versions 5 and later of netCDF-Java/TDS will reside at https://artifacts.unidata.ucar.edu/repository/thredds-doc/ and—more conveniently—at https://docs.unidata.ucar.edu/thredds.
- The configuration management tool Ansible has shown great promise as a way for users to be able to deploy TDS and other Unidata software in an automated fashion.
- DAP4 in the TDS has been updated to be consistent with the specification and to successfully allow the netCDF-C DAP4 and NetCDF-Java libraries to read DAP4 responses from the TDS.
- New Coverage data type allows for subsetting across array boundaries (often called the "seam" problem).
- Uses the new edal-java based ncWMS 2.0 server, as well as javascript client Godiva3.
- CatalogScan feature allows for incremental updating of TDS catalogs without the need to restart Tomcat.
- Upload/Download support has been added to TDS. This now includes an upload web form accessible as http://.../thredds/upload.
- Unit and Integration tests are passing in 5.0. This is a big step towards releasing a beta.
- ncSOS has been integrated into the TDS distribution (as part of the OIIP project—see the Rosetta section for more details)
- Access to the netCDF-C library via JNI is now thread-safe so that the HDF5 library no longer needs to be built with thread-safe support.
- The license for netCDF-Java and the TDS has been updated to a BSD-3 clause licence. See https://www.unidata.ucar.edu/blogs/developer/entry/thredds-licence-change for more information.

Dependencies, challenges, problems, and risks include:

- Maintenance of the 4.6.x line of netCDF-Java and TDS continues to have a large impact on the team. 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 will be critical, and so far we have had several community members offer their help (special thanks to Rich Signell!).

Rosetta

Rosetta continues to progress thanks to support from a NASA ACCESS grant (the Oceanographic In-situ data Interoperability Project, or **OIIP**), in which Unidata is partnering with the PO.DAAC at JPL and UMASS-Boston. We are currently in our final 6 months of funding for the project.

Progress has been made on the following:

- Support for the NCEI NODC netCDF v2.0 templates (metadata standards)
- Extension of the NCEI templates to support metadata critical to the use of electronic tagging datasets
- Support automated transformation of output from electronic animal tagging datasets in the Electronic Tag Unified File Format (eTUFF) format via Rosetta.
- Working to create a unified workflow for the gui wizard interface that allows for selection of which metadata standards to use when determining recommended/required metadata
- Engaging with the netCDF Linked Data initiative to define best practices identifying netCDF metadata to a particular metadata standard.

Dependencies, challenges, problems, and risks include:

- Two of the core javascript libraries used by Rosetta have been abandoned by their original creators. This has been a major roadblock in updating the front end interface to use a more recent version of jQuery, which is badly needed due to security concerns. One of these libraries has been picked up by the community (SlickGrid), while the other has been in limbo (jWizard). We are happy to announce that with the support of Jen Oxelson, we have transitioned away from jWizard and are using our own wizard interface.

Ongoing Activities

We plan to continue the following activities:

- Documentation updates - We are reworking the tutorial material for TDS v5.0 with the goal of enabling asynchronous training. The material will undergo a major overhaul to include the use of Docker containers, video snippets, and other new forms of training tools. The first pass at the overhaul is now complete.
- Maintain thredds.ucar.edu and keep up with the addition of new datasets to the IDD.
    - GOES-16 data, with tiles stitched together using python, is available on our test TDS.
- Continue development of the TDS python client siphon, as well as extend its functionality to interface with other web services and servers.

The following active proposals directly involve THREDDS work:

- The NASA ACCESS award with JPL is entering into the second year of the two year award. The award is titled: "Leveraging available Technologies for Improved interoperability and visualization of Remote Sensing and in-situ Oceanographic Data at the PO.DAAC" and was submitted with JPL/PO.DAAC. [Rosetta]
- EarthCube award: "Advancing netCDF-CF for the Geosciences". This two-year, Unidata lead project will work to extend netCDF-CF conventions in ways that will broaden the range of earth science domains whose data can be represented.
- Finished the second and final year of EarthCube award: "CyberConnector: Bridging the Earth Observations and Earth Science Modeling for Supporting Model Validation, Verification, and Intercomparison" with George Mason University.

- Thanks to Rich Signell, we, along with Axiom Data Science, submitted a proposal to NOAA IOOS entitled “A Unified Framework for IOOS Model Data Access” with 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). We have not received word on the the status of the proposal at this time.

New Activities

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

- Officially advertising a public TDS 5.0 Test Server [currently found at http://thredds-test.unidata.ucar.edu/thredds/catalog.html]
- Getting TDS v5.0 to a stable release
- Ryan May and Sean Arms are officially involved with the GRIB-3 effort at the WMO. Work is being done to create native java and python decodes for the new version as independent implementations to validate the GRIB-3 specification.

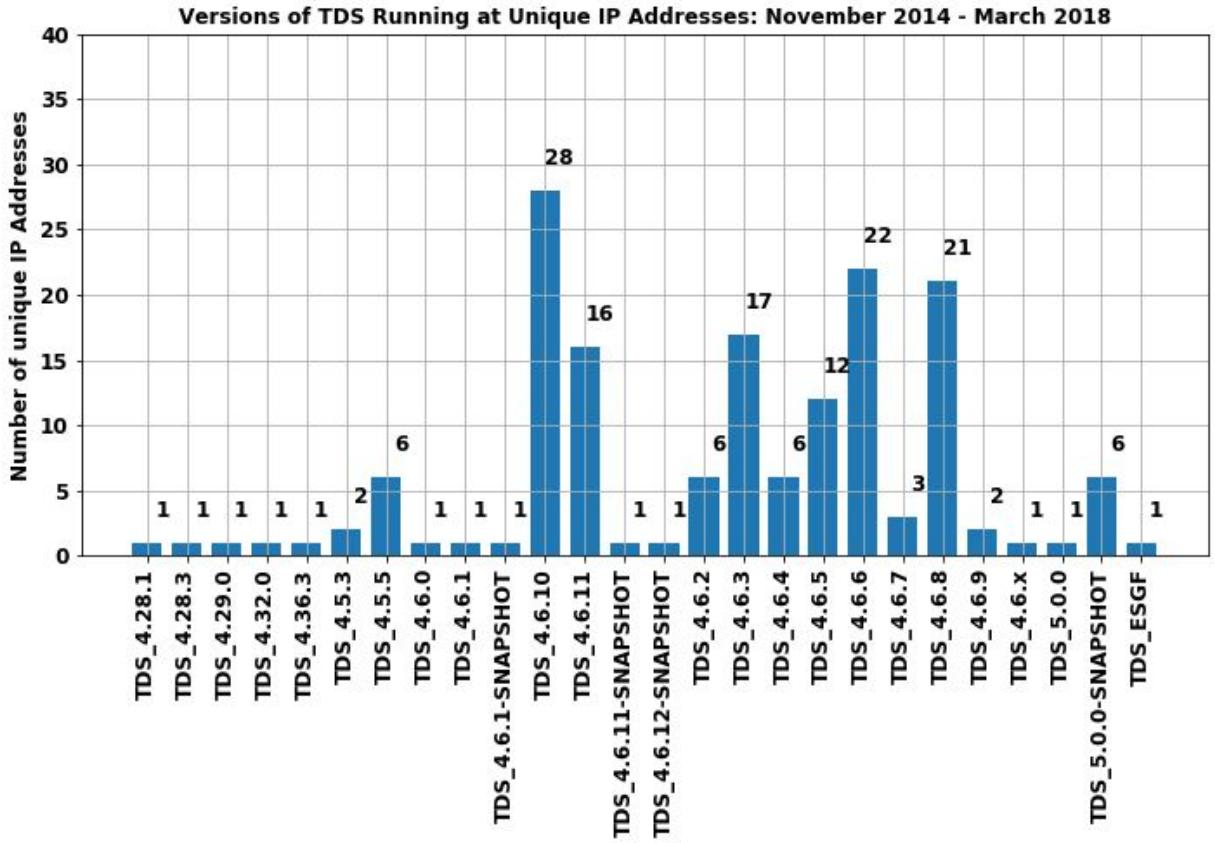
Over the next twelve months, we plan to organize or take part in the following:

- Support the UGRID spec within the CDM
- Support the netCDF-CF simple-geometry spec within the CDM
- Create a TDS plugin layer for external services (potentially use Project Jigsaw from Java 9)
    - Upgrade the ncWMS, ncISO, and other plugin services to use the new TDS 5.x plugin layer
    - Incorporate ncSoS into TDS
- Transitioning thredds.ucar.edu to TDS 5.x
- Getting netCDF-Java v5.x to a stable release

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

- NcML and FMRC collections have many problematic warts, including unreliable caching. Tackling these issues will be critical once netCDF-Java and TDS 5 are released.
- The scalability of a serverless architecture is very attractive, but would require a re-architecture of the TDS, likely to something resembling a microservice architecture. While the time horizon on this kind of transition is long, we plan on exploring options with some TDS capabilities, such as the catalog service or the netCDF Subset Service.

Relevant Metrics



10,832 unique IPs started up thredds from November 2014 through March 2018, **160** of which are publicly accessible servers. Publically accessible is defined as the following URL patterns being accessible with an HTTP GET requests with a return status less than 400 as well as content that contains xml:

http(s)://<ip address>/thredds/catalog.xml
http(s)//<ip address>:8080(8443)/thredds/catalog.xml

You may notice that the number of publicly accessible TDSs decreased just over half since our last report (now at 160). This is due to a new check that, in addition to being resolvable, the server response is actually an xml file. Many of the of the previously counted “publicly accessible” TDSs that are now excluded are AWS 404 html pages (and in some case, not so “PG” ad pages).

This information is only known for servers running v4.5.3 and above. There are many reasons why these number are so different. The differences could be due to:

- People testing the TDS on their local machine, but not actually running a server (most likely the cause for the majority of the difference)
- A TDS running behind a proxy server may not be “seen” in this analysis as publicly reachable at the tested url pattern (<server>/thredds/catalog.xml). For example, a TDS running behind a proxy might be configured to respond to

mytds.<server>/catalog.xml.
- The TDS server may be running behind a firewall that does not allow for public access.
- A TDS running in the past is no longer running today.

Note that the vast majority of the publicly accessible servers are running v4.6.3 or above (v4.6.11 was the most current release during this period, and was released on 11 December 2017, and is the most commonly run version of the 4.6.x line of the TDS). This indicates that users and organizations running the TDS tend to follow along closely with the current releases of the TDS.

Note that there are some odd looking versions of the TDS being reported in the log files, such as TDS_4.28.x. It is likely these version numbers are actually generated by software that is being built on top of the TDS or applications that bundle the TDS as part of a deployment package (perhaps ESGF nodes?).

Strategic Focus Areas

We support the following goals described in Unidata Strategic Plan:

1. **Enable widespread, efficient access to geoscience data**
   The work of the THREDDS group is comprised of two main areas: the THREDDS Data Server (TDS) and the Common Data Model (CDM) / netCDF-Java library. The TDS provides catalog and data access services for scientific data using OPeNDAP, OGC WCS and WMS, HTTP, and other remote data access protocols. The CDM provides data access through the netCDF-Java API to a variety of data formats (e.g., netCDF, HDF, GRIB). Layered above the basic data access, the CDM uses the metadata contained in datasets to provide a higher-level interface to geoscience specific features of datasets, in particular, providing geolocation and data subsetting in coordinate space. The CDM also provides the foundations for all the services made available through the TDS.

   The data available from the IDD is a driving force on both the TDS and netCDF-Java development. The ability to read all the IDD data through the netCDF-Java library allows the TDS to serve that data and provide services on/for that data.

2. **Develop and provide open-source tools for effective use of geoscience data**
   Unidata's Integrated Data Viewer (IDV) depends on the netCDF-Java library for access to local data, and on the THREDDS Data Server (TDS) for remote access to IDD data. At the same time, the CDM depends on the IDV to validate and test CDM software. Many other tools build on the CDM / netCDF-Java library (e.g. ERDDAP, Panoply, VERDI, etc) and on the TDS (ESGF, LAS, ncWMS, MyOcean, etc).

3. **Provide cyberinfrastructure leadership in data discovery, access, and use**
   The Common Data Model (CDM) / netCDF-Java library is one of the few general-purpose implementations of the CF (Climate and Forecast) metadata standards. Current active efforts in CF that we are involved with include use of the extended netCDF-4 data model (CF 2.0) and for point data (Discrete Sampling Geometry CF-DSG).

The TDS has pioneered the integration of Open Geospatial Consortium (OGC) protocols into the earth science communities. Strong international collaborations have resulted in WCS and WMS services as part of the TDS.

The CDM and TDS are widely used implementations of the OPeNDAP DAP2 data access protocol. Unidata has worked with the OPeNDAP group to design, develop, and implement a new version of the DAP specification, DAP4, which is now available in the TDS server and the netCDF-Java client software stack.

4. **Build, support, and advocate for the diverse geoscience community**
   The THREDDS project is involved in several international standardization efforts (CF, OGC, etc.) which cross-cut a multitude of disciplines, both inside and outside of the geoscience community. The netCDF-Java client library, as well as the TDS often serve as incubators for new pushes in these efforts.

Prepared *March 2018*

Special Status Report: NSF Core Proposal Planning

March 2018

On Monday, March 5 2018, Unidata Program Center staff gathered at the NCAR Mesa Lab's Fleischmann Building for a day-long brainstorming retreat. The goal for the day was to generate and refine ideas for ways that Unidata can better serve its community of geoscience researchers and educators, and to expand its reach to new communities that could benefit from Unidata's activities but currently are not involved.

The brainstorming session took place in the context of the Unidata strategic plan as approved by the Strategic Advisory Committee during the Fall 2017 committee meeting. That plan groups Unidata's contributions to the community into three broad categories:

- Managing Geoscience Data
- Providing Useful Tools
- Supporting People

Staff began by writing down ideas they felt would provide benefits to the Unidata community and to the program, and grouping the ideas into the strategic plan's broad categories. This exercise was followed by additional discussion in smaller groups focusing on each of the three areas. The discussion shifted between high-level concepts like

- Unidata serving as a "maker space" or "workbench" that combines geoscience data with tools to analyze that data
- Promoting data access standards for cloud-hosted geoscience data
- Designing general-purpose data transformation tools, to easily convert data from "raw" formats to "useful" formats
- Enhancing Unidata training offerings to better support a 21st century geoscience workforce
- Encouraging more community participation in Unidata activities through internships and "visitor" positions at the Program Center
- Promoting greater community involvement in software development efforts
- "Democratizing" geoscience by creating workflows that are easy for anyone to adopt, lowering resource and experience barriers to doing science

and more technical topics such as

- Which technologies would be best to create web-accessible workflows
- Whether/how existing data services could be recast as "microservices" better suited for cloud environments
- Ways to standardize different web APIs used by different cloud services
- Whether providing pre-built virtual machine images would be an appropriate solution for platform/operating system problems

The ideas generated during the staff workshop are intended to help define the main areas of

new activity for next NSF core funding proposal, which we plan to submit in early Fall 2018. The next steps will be to define specific activities and areas of investigation, to suggest the hoped-for outcomes, and concretely outline why these are important for the Unidata community.

__Word-cloud created from the workshop's initial brainstorming activity__

Prepared *March 2018*