Unidata Activities in "The Cloud"

Julien Chastang, Unidata Staff

Unidata Program Center
University Corporation for Atmospheric Research
Boulder, Colorado USA
http://www.unidata.ucar.edu

UCAR COMMUNITY PROGRAMS

NSF

Unidata 2024 Proposal: Science as a Service

The Science as a Service concept draws together Unidata's ongoing work to provide geoscience data and software for analysis and visualization with access to workflows designed to take advantage of cloud computing resources.

Getting our Feet Wet

Amazon Web Services – 2016

- Modest CentOS-6 VM: 2 vCPUs, 7.5 GB RAM, 400 GB disk
- Experiment in running LDM, McIDAS and GEMPAK applications to produce and distribute IDD products in a “cloud” environment
- So successful that it is still running and creates content for our IDD UNIWISC and FNEXRAD feeds
- Paid service - $250 / month

Microsoft Azure – 2016-2018

- More ambitious experimentation: THREDDS Data Server, McIDAS ADDE, AWIPS EDEX, RAMADDA all fed by LDM/IDD
- Much more robust server instances
- Resources provided from a grant from Microsoft

Lessons Learned

Amazon Web Services

- Very reliable
- More expensive than running same tasks on internal machines
- Egress costs for high volumes would be cost prohibitive

Microsoft Azure

- Reliable, but not as reliable as AWS
- VM instances noticeably slower than equivalent AWS instances
- One year grants: At end of grant period, VMs that we setup were thrown away
- Egress costs for high volumes would be cost prohibitive

Next Steps

Amazon Web Services

- NOAA Big Data project: began moving NEXRAD Level 2 data to S3 bucket in real-time using the LDM/IDD
- THREDDS Data Server instance to serve data from the S3 bucket
- GOES-16 GRB L1b data to S3 bucket using Python-based procedure
- CONDUIT content added to data being uploaded to S3 buckets
- S3 buckets for NEXRAD Level 2, GOES-16 GRB and CONDUIT data provided by Amazon
- VM instances that support the NEXRAD Level 2 data upload provided by Amazon

Next Steps (cont.)

Google Cloud Platform – depending on funding

- NOAA Big Data project, move NEXRAD Level 2 data in real-time using the LDM/IDD
- THREDDS Data Server, McIDAS ADDE, AWIPS EDEX and RAMADDA instances to serve data uploaded using LDM/IDD
- Juypter Hub

NSF Jetstream Cloud Collaboration

What is Jetstream?

- A National Science and Engineering Cloud funded by an $11 million NSF grant
- Data centers at IU and TACC

Attached to fast Internet2 capability

Cloud based on *OpenStack* for creation of VMs, routers, networks, subnets, security groups etc.

Unidata has been running on Jetstream for 3 years through a series of research grants

Once get through granting process, Jetstream is free including egress!

Exploration of the Jetstream Cloud

Started by containerizing Unidata technology offerings using Docker

- THREDDS Data Server
- LDM
- McIDAS ADDE
- RAMADDA

Deployed containers to create near complete Unidata data center

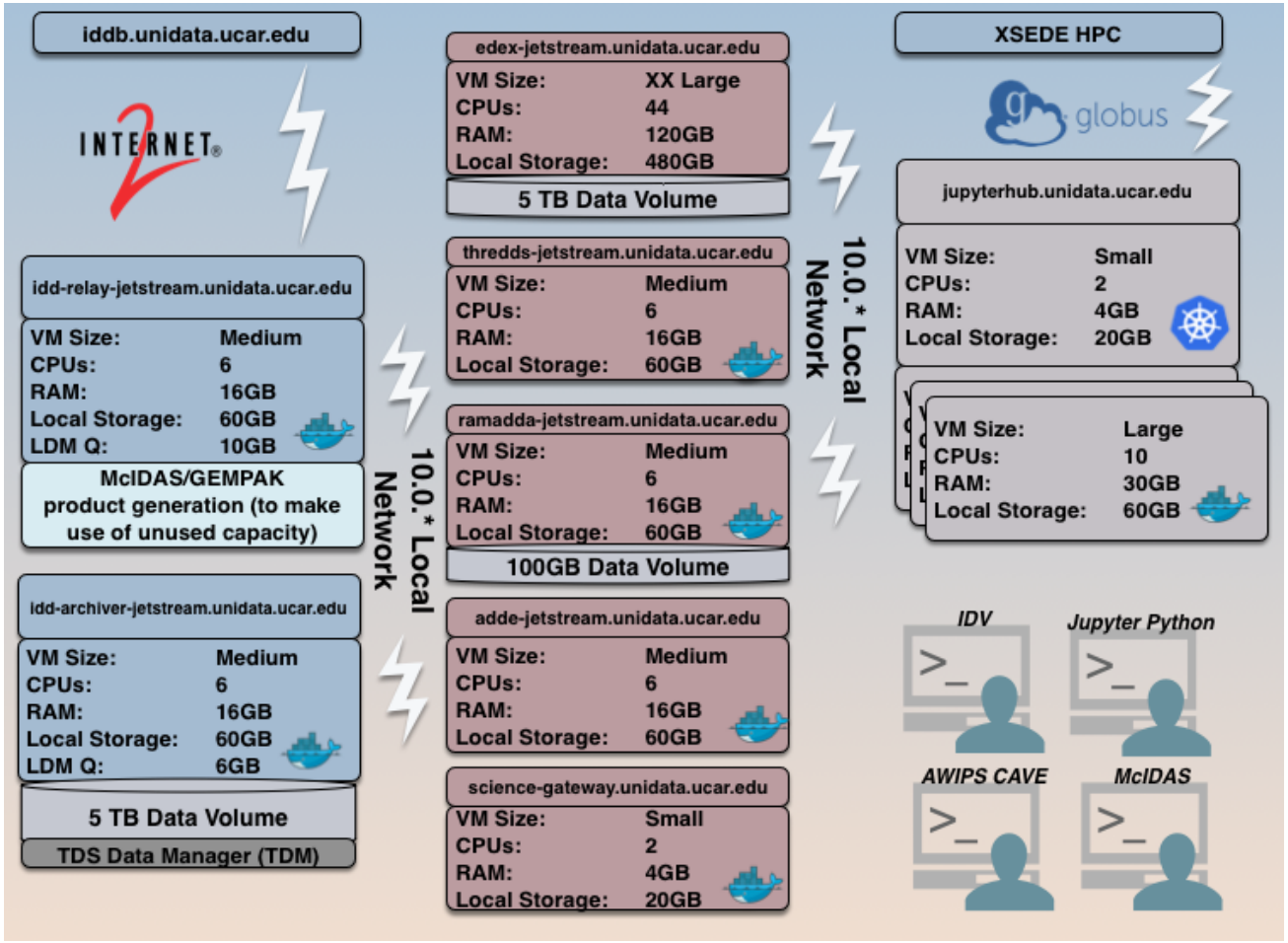
Plenty of NCEP model output at thredds-jetstream.unidata.ucar.edu

Question: What about client-side offerings in cloud?

Next step: "data-proximate" analysis and visualization

Unidata

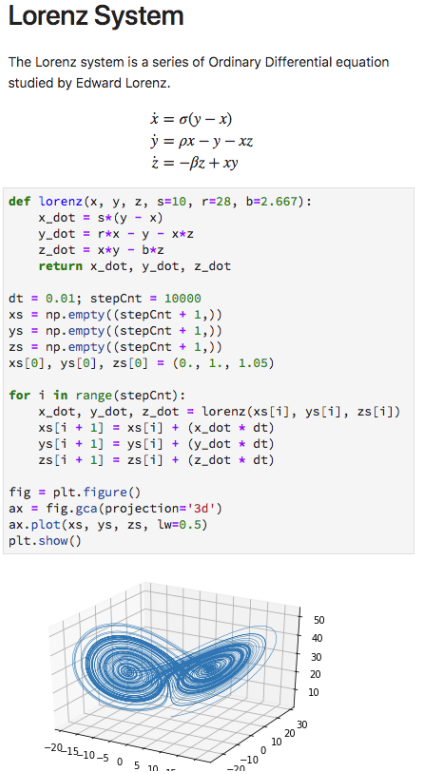
Providing data, tools, and community leadership for enhanced Earth-system education and research



Deploying a Geoscience JuypterHub on Jetstream Cloud

A narrative of:

- Explanatory and expository text
- Software code (Python, R, etc.) and output
- Equations (MathJax, LATEX)
- Figures and multimedia



Lorenz System

The Lorenz system is a series of Ordinary Differential equation studied by Edward Lorenz.

$$\dot{x} = \sigma(y - x)$$
$$\dot{y} = \rho x - y - xz$$
$$\dot{z} = -\beta z + xy$$

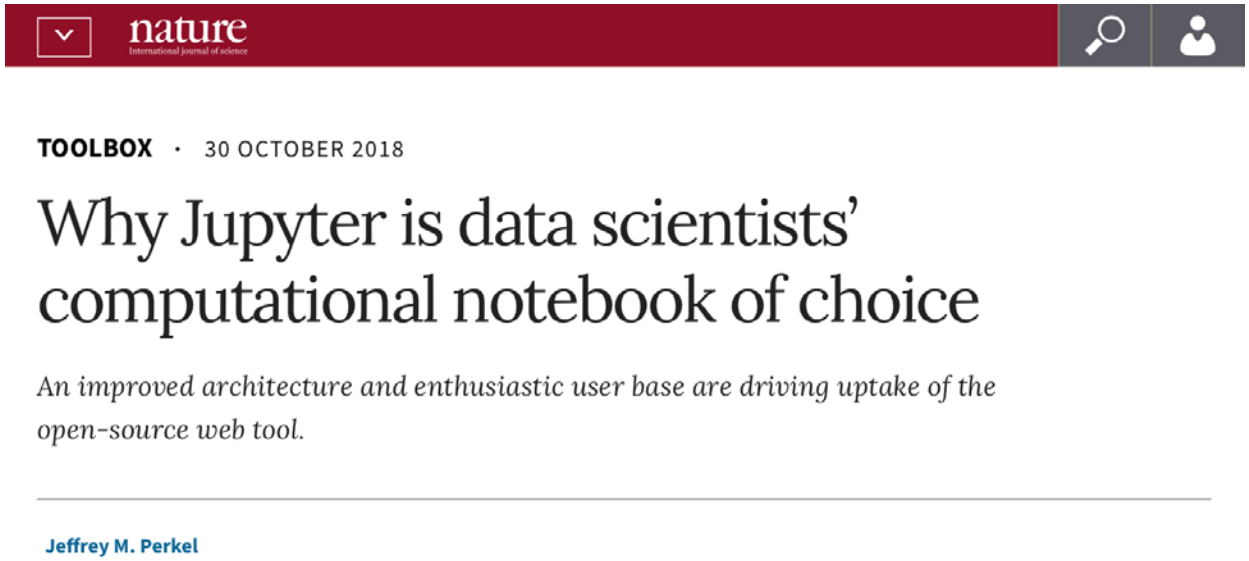
In [10]:

```
def lorenz(x, y, z, s=10, r=28, b=2.667):
    x_dot = s*(y - x)
    y_dot = r*x - y - x*z
    z_dot = x*y - b*z
    return x_dot, y_dot, z_dot

dt = 0.01; stepCnt = 10000
xs = np.empty((stepCnt + 1,))
ys = np.empty((stepCnt + 1,))
zs = np.empty((stepCnt + 1,))
xs[0], ys[0], zs[0] = (0., 1., 1.05)

for i in range(stepCnt):
    x_dot, y_dot, z_dot = lorenz(xs[i], ys[i], zs[i])
    xs[i + 1] = xs[i] + (x_dot * dt)
    ys[i + 1] = ys[i] + (y_dot * dt)
    zs[i + 1] = zs[i] + (z_dot * dt)

fig = plt.figure()
ax = fig.gca(projection='3d')
ax.plot(xs, ys, zs, lw=0.5)
plt.show()
```

Unidata

Providing data, tools, and community leadership for enhanced Earth-system education and research

Success of Jupyter in Research and Education



nature
International journal of science

TOOLBOX · 30 OCTOBER 2018

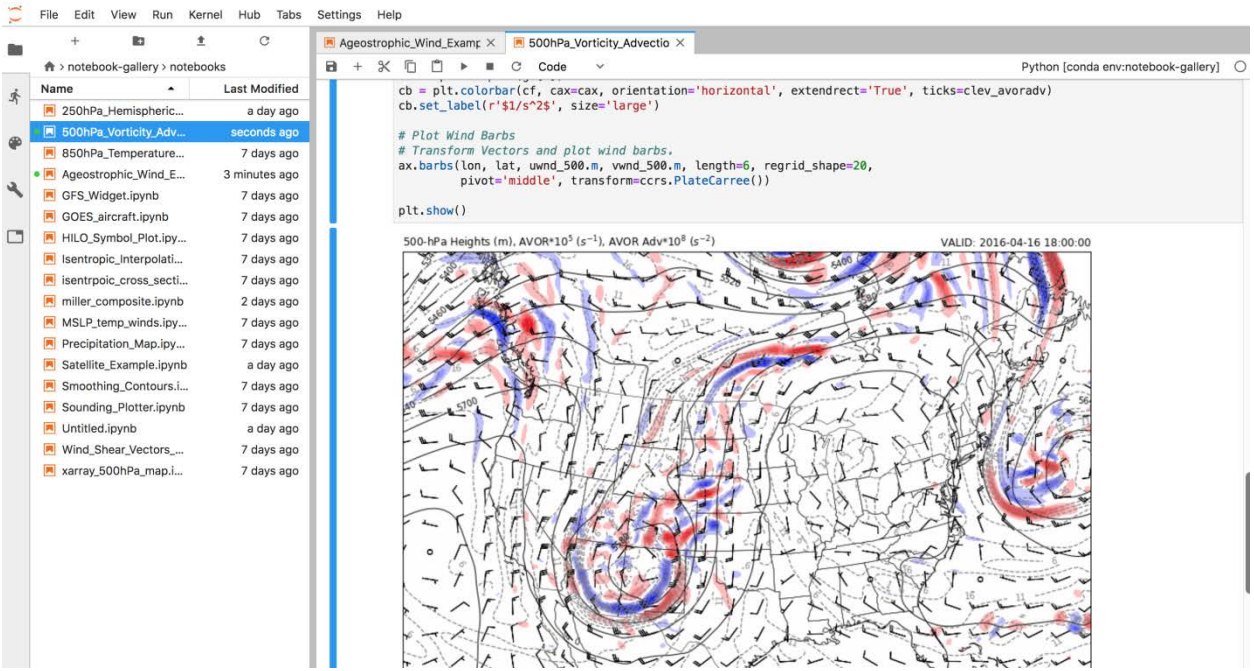
Why Jupyter is data scientists' computational notebook of choice

An improved architecture and enthusiastic user base are driving uptake of the open-source web tool.

Jeffrey M. Perkel

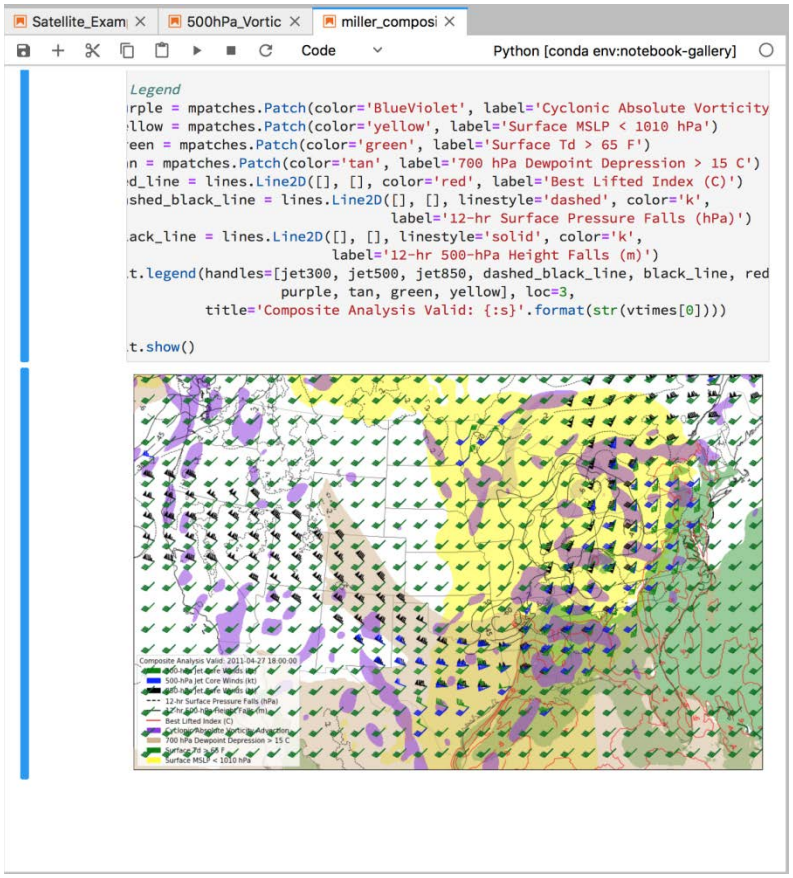
[Jupyter] notebooks are really a killer app for teaching computing in science and engineering – Lorena Barba, Engineering Professor, GWU

JupyterLab: Next Generation UI

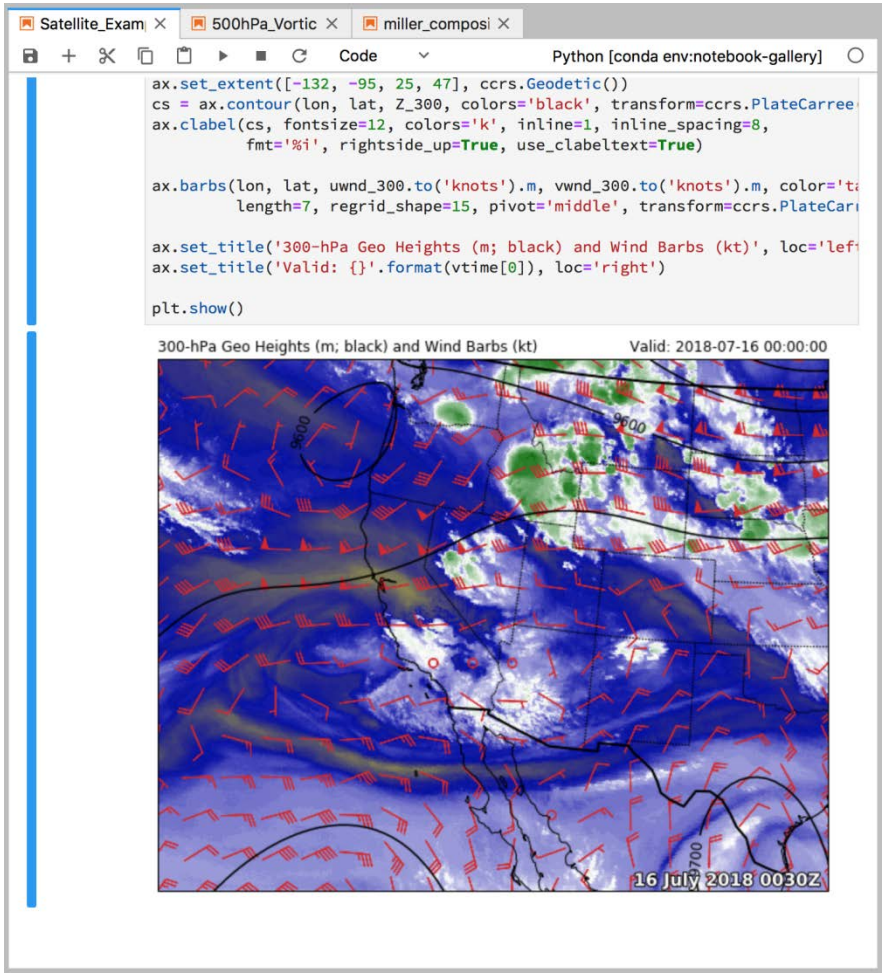


- Terminal (git, conda, etc.)
- Text Editor

Notebooks: Miller Composite



Notebooks: Satellite + GFS Model Output

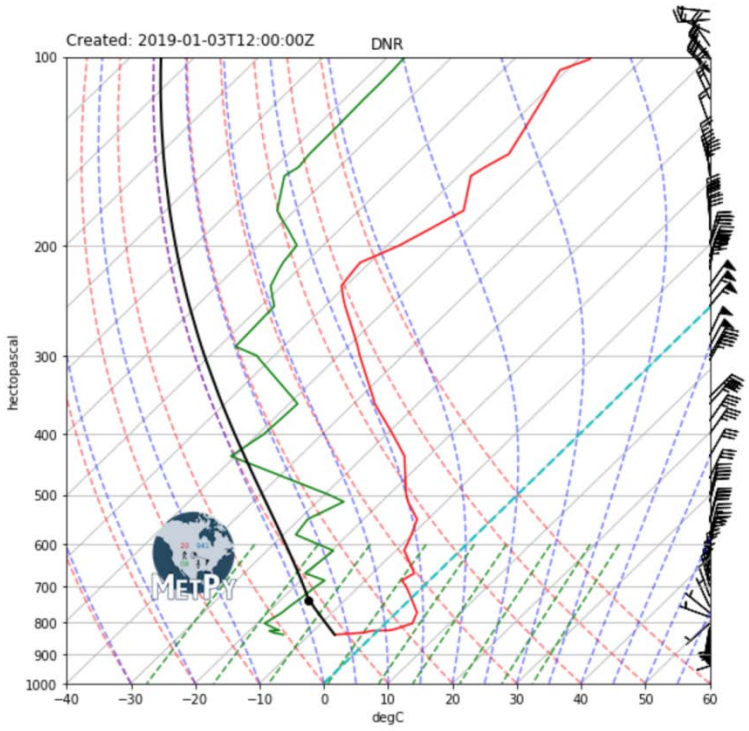


Notebooks: Skew-T

Sounding_Plotter.ipynb

Code | Python [conda env:notebook-gallery]

```
        with tempfile.NamedTemporaryFile(suffix='.png') as f:
            skewt.ax.figure.savefig(f.name)
            uploader.upload_to(f.name, posixpath.join(args.gdrive, fname))
    else:
        skewt.ax.figure.savefig(make_name(args.site, request_time))
```



Unidata Jetstream Resources

IDD	Idd-relay-jetstream.unidata.ucar.edu
IDD Storage	Idd-archiver-jetstream.unidata.ucar.edu
AWIPS EDEX	edex-jetstream.unidata.ucar.edu
TDS	thredds-jetstream.unidata.ucar.edu
ADDE	adde-jetstream.unidata.ucar.edu
RAMADDA	ramadda-jetstream.unidata.ucar.edu
Science Gateway	http://science-gateway.unidata.ucar.edu/
JupyterHub	https://jupyterhub.unidata.ucar.eduu
AMS 2020	https://js-168-90.jetstream-cloud.org

Unidata Information

HomePage	www.unidata.ucar.edu
Support	support@unidata.ucar.edu
Software	www.unidata.ucar.edu/software
LDM	www.unidata.ucar.edu/software/ldm
netCDF	www.unidata.ucar.edu/software/netcdf
IDV	www.unidata.ucar.edu/software/idv
AWIPS	www.unidata.ucar.edu/software/awips2
McIDAS	www.unidata.ucar.edu/software/mcidas
TDS	www.unidata.ucar.edu/software/thredds/current/tds

UCAR COMMUNITY PROGRAMS

NSF