

High resolution regional climate downscaled data transfer using Globus



Jiali Wang^{1*}, V. Rao Kotamarthi¹, Gagan Agrawal², Raj Kettimuthu³, Roe Ebenstein², Yu Su², Torrance I. Leggett⁴

¹ Environmental Science Division, Argonne National Laboratory

² Department of Computer Science and Engineering, The Ohio State University

³ Mathematics and Computer Science Division, Argonne National Laboratory

⁴ Computing, Environment and Life Sciences, Argonne National Laboratory

*Contact: jialiwang@anl.gov



❖ Why we need a tool

- Pre-process
- Transfer between locations
- Back-up model outputs
- Share with collaborators
- Publish to the world
- Download by users

❖ Data size

- >200TB, 2TB per model year
- Historical/future periods
- Different emission scenarios

❖ Challenges

- Transfer speed/performance
- Low cost
- Security

❖ Machines

- Fusion, Argonne
- BG/Q (Mira), Argonne
- NERSC, Lawrence Berkeley
- Any machines used by users

❖ Globus

- Enable intensive analysis
- ~10x improvement over sftp
- Minimum firewall requirements
- No need to VPN/on-site login
- Mass storage (HPSS)

❖ SDQuery DSI plug-in

- Flexible server-side data subset
- Queries based on dimensions/coordinates/values
- Allow users directly perform particular analysis
- Support NetCDF and HDF

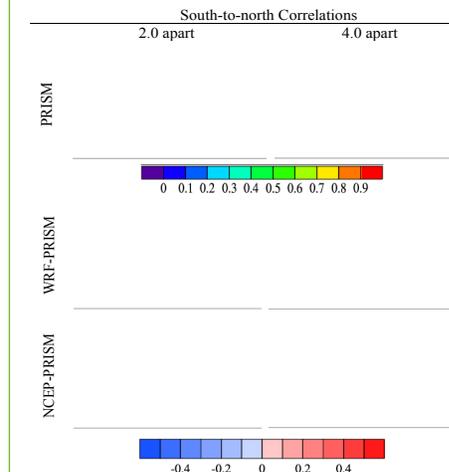
<p>Logical Layout Descriptor</p> <pre>varname = "TEMP", varid = 3 datatype = float ndims = 4 dim = {time, depth, lat, lon} Coordinate Variables: lat = (-78.47, -78.43, ..., 89) lon = (-179.95, -179.85, ..., 179.95) Value Distribution Descriptor Min/Max Value: (-21.1, 33.1) Value Distribution: (-21.1, -20.1): 201 (-20.1, -19.1): 233 (32.1, 33.1): 120</pre>	<p>SQL Query</p> <pre>SELECT TEMP FROM temp.nc WHERE TEMP >=5 AND lat>24 AND lat<30 AND lon>98 AND lon<82 AND depth=50</pre> <p>GridFTP Data Retrieve Request</p> <pre>globus-url-copy "ftp://127.0.0.1:5000/tmp/server/temp.nc SELECT TEMP FROM temp.nc WHERE TEMP >=5 AND lat>24 AND lat<30 AND lon>98 AND lon<82 AND depth=50)" file://tmp/client/netcdfsubset/</pre>
--	---



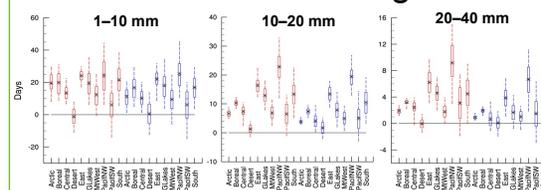
Find the publication here by Su et al. 2013.

❖ Climate model data

- Why high resolution?



- Future climate changes?



Acknowledgement: Strategic Environmental Research and Development Program; Argonne Leadership Computing Facility; Argonne Laboratory Computing Resource Center; National Energy Research Scientific Computing Center.