CDATWeb: A Client/Server Model for the Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT)



Backgound

UV-CDAT is an integrated framework that provides an endto-end solution for management, analysis, and visualization of ultrascale data sets generated for current and future DOE climate data repositories and for the climate science community at large.

UV-CDAT is based on a client-server architecture and is integrated within the Earth System Grid Federation (ESGF) framework, allowing UV-CDAT to take advantage of the advanced data management mechanisms of ESGF. In this way, UV-CDAT provides regridding, reprojection, and aggregation tools directly as a component of the ESGF data node, eliminating or substantially decreasing data movement. The UV-CDAT server provides a turnkey application for building complex data analysis and visualization workflows by interacting with one or more UV-CDAT backend servers. These workflows may use predefined components for data transformation and analysis, data collection from disparate data sources outside ESGF, visualization, and user-defined processing steps. Extended provenance information can be captured on the backend server for sharing and collaboration.

Data from CADES at ORNL Visualization Server in CADES at ORNL (File server mounted on visualization server) CDATWeb Client running in development environment at LLNL

R FILE BROWSER	QPLOT SELECTORS	PLOT WINDOW
ne120_monthly_ens3	Graphic Methods	
• gridded	• meshfill	heufill /defeult
• 1979	3d_vector	boxfill/default
• 1986	 3d_scalar 	
• 1995	• xvsy	default
 EXTRA 	 yxvsx 	
 gridded_ne120_v0_3_00003.cam.h0.1995.xml 	• xyvsy	TVO
 gridded_ne120_v0_3_00003.cam.h0.1995-05.nc 	 taylordiagram 	TVQ
 gridded_ne120_v0_3_00003.cam.h0.1995-03.nc 	 isoline 	
■ gw	 boxfill 	
 TMQ 	 polar 	
 LANDFRAC 	 a_lambert_boxfill 	
 SHFLX 	 default 	
 QFLX 	 a_boxfill 	
 area 	 a_polar_boxfill 	
 lat_bnds 	 a_robinson_boxfill 	Imation TVQ
 FLNS 	 a_mollweide_boxfill 	
 FLNT 	 quick 	
 Ion_bnds 	 robinson 	60N
 TUQ 	• isofill	
 FSNS 	• vector	40N
 time_bnds 	• 1d	Contraction 2
 LHFLX 	 3d_dual_scalar 	- 20N
 PRECL 	• scatter	1 5 9
 PRECC 	Templates	- Set of F
 FSNT 	ASD	
 TVQ 	 ASD_dud 	- TAR BU
 gridded_ne120_v0_3_00003.cam.h0.1995-06.nc 	 BL_of6_1legend 	F 208 100
 gridded_ne120_v0_3_00003.cam.h0.1995-01.nc 	 BLof6 	- Crime
 gridded_ne120_v0_3_00003.cam.h0.1995-09.nc 	 BR_of6_1legend 	40S
 gridded_ne120_v0_3_00003.cam.h0.1995-07.nc 	 BRof6 	
 gridded_ne120_v0_3_00003.cam.h0.1995-11.nc 	 LLof4 	605
 gridded_ne120_v0_3_00003.cam.h0.1995-08.nc 	 LLof4_dud 	
 gridded_ne120_v0_3_00003.cam.h0.1995-12.nc 	 LRof4 	- 805
 gridded_ne120_v0_3_00003.cam.h0.1995-02.nc 	 LRof4_dud 	
 gridded_ne120_v0_3_00003.cam.h0.1995-10.nc 	 ML_of6 	905 0 30E 60E 90E 120E 150E
 gridded_ne120_v0_3_00003.cam.h0.1995-04.nc 	 ML_of6_1legend 	
 scripts_logs 	 MR_of6 	-380 -340 -300 140 130 80
 famipc5_ne120_v0_3_00003.cam.h0.1983-06.nc 	 MR_of6_1legend 	-280 -240 -200 -160 -120 -80
 famipc5_ne120_v0_3_00003.cam.h0.1988-08.nc 	 UL_of6_1legend 	
 famipc5_ne120_v0_3_00003.cam.h0.1991-08.nc 	 ULof4 	
 famipc5_ne120_v0_3_00003.cam.h0.1993-11.nc 	 ULof4_dud 	
 famipc5_ne120_v0_3_00003.cam.h0.1988-04.nc 	 ULof6 	
 famipc5_ne120_v0_3_00003.cam.h0.1989-12.nc 	• UR_of6 //	



Access

- We are providing three different ways for users to access UV-CDAT web informatics (i.e., CDATWeb):
- (1) import data from ESGF;
- (2) import an existing UV-CDAT project that has been previously saved; and
- (3) import data sets from local machines (prior to ESGF publication).
- **Import Data from ESGF.** A user can simply go to ESGF to find their desired data sets and import these data sets into the CDATWeb informatics interface via the UV-CDAT hyperlink that is associated with either a single data set or the entire data cart.
- **Existing Project.** Our web interface provides the option to open an existing project that has been previously saved. Also, we intend to provide the option of recovering an existing project in the case where a client browser crashes in the middle of an active session.
- Local File. A user can simply open a directory containing all the data sets and upload to the UV-CDAT server for remote data analysis.

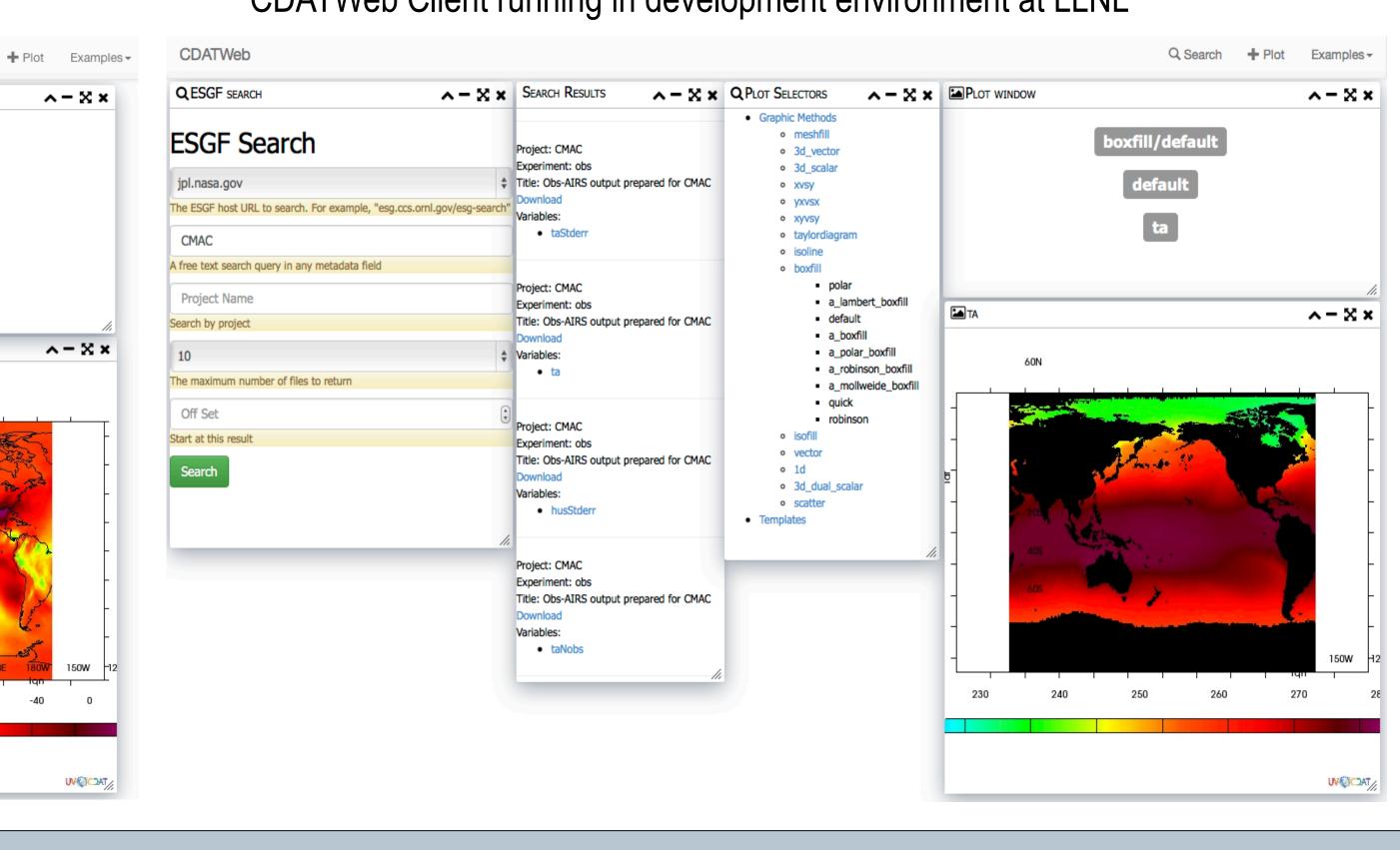
Web Browser

Prototype

Data from JPL Data Node

Visualization Server in CADES at ORNL (File server not mounted on visualization server)

CDATWeb Client running in development environment at LLNL



Components

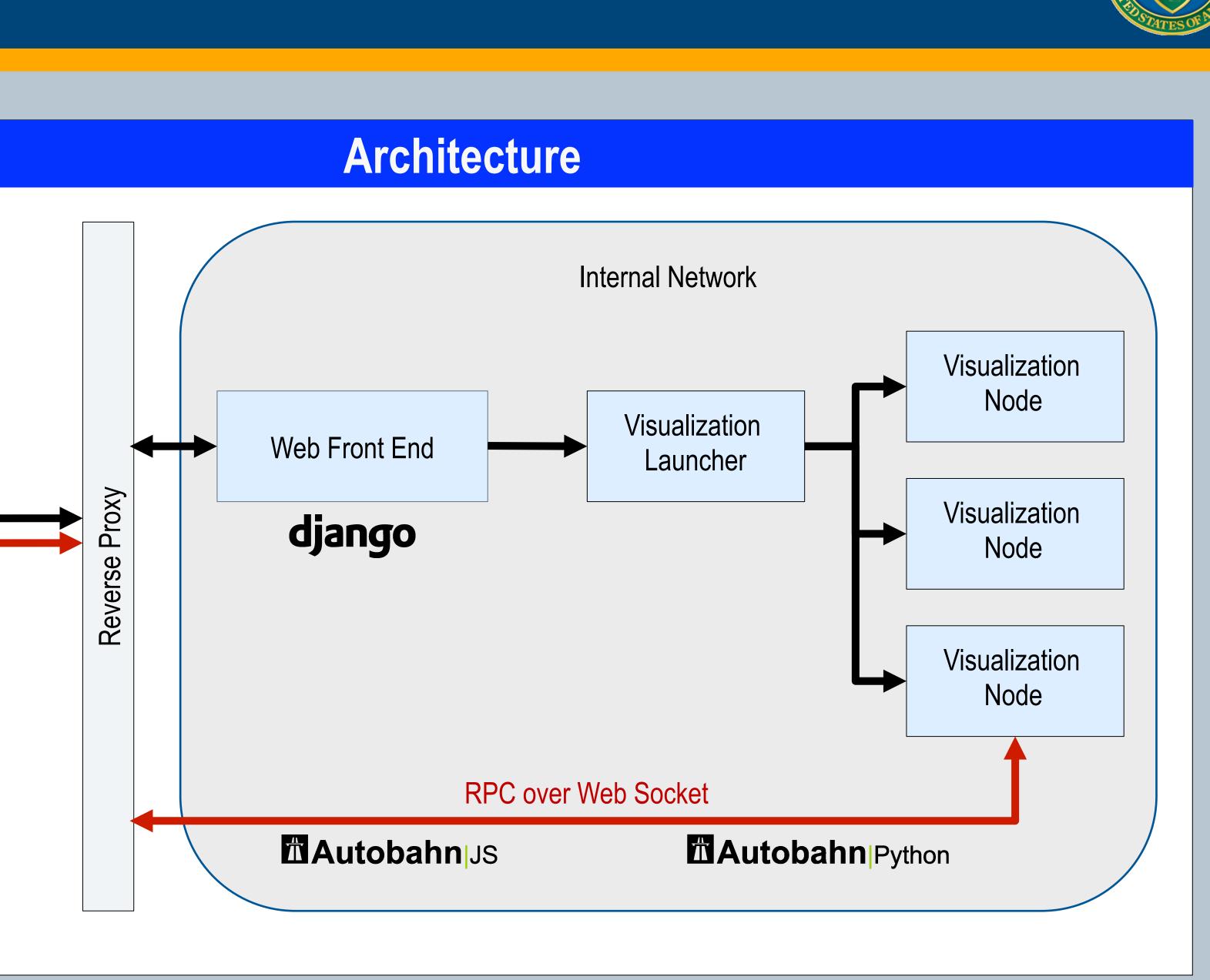






Jonathan Beezley¹, Aashish Chaudhary², Matthew Harris³, Charles Doutriaux⁴, Dean N. Williams⁵

¹Kitware ²Lawrence Livermore National Laboratory



Next Steps

High-resolution models, ensemble analysis, derived data product generation, and intercomparison of model results and observations require substantial data resources regarding storage infrastructure (e.g., high-performance parallel input/ output (I/O) environments) and server-side (or remote) data processing. Remotely manipulating these data sets—from generation to transformation and fusion to archiving—requires a robust infrastructure that supports multiple models, storage systems, operating systems, and compute environments.

Test bed simulations, such as the DOE's Accelerated Climate Modeling for Energy (ACME) project, require parallel I/O and parallel compute environments that provide a structured data model suitable for parallelizing and automating ensemble analysis and intercomparison of models and observational data sets remotely. Observational data sets come in a variety of formats and use numerous metadata models. Supporting the server-side requirements of these data sets and data products derived from these data sets—requires decoupling the data format from data access and exploratory/parallel analysis.

To accommodate ACME better in the future, the following next steps must occur:

- Data centers provide better compute and visualization hardware for ESGF/CDATWeb server-side computation;
- CDATWeb visualization server runs on a local cluster or node at any of the DOE leadership compute facilities (i.e., ALCF, OLCF, NERSC);
- Mount point for data be near data center's storage facilities for fast access and data movement;
- CDATWweb Django front-end installed alongside and or integrated with the ESGF CoG interface;
- Produce PNG files on full rotating 3D plots;
- CDATWeb fully intergraded with ESGF's Solr search engine for better search capability; and
- Mimic the UV-CDAT desktop application's graphical user interface, so users become accustomed to similar widgets.



