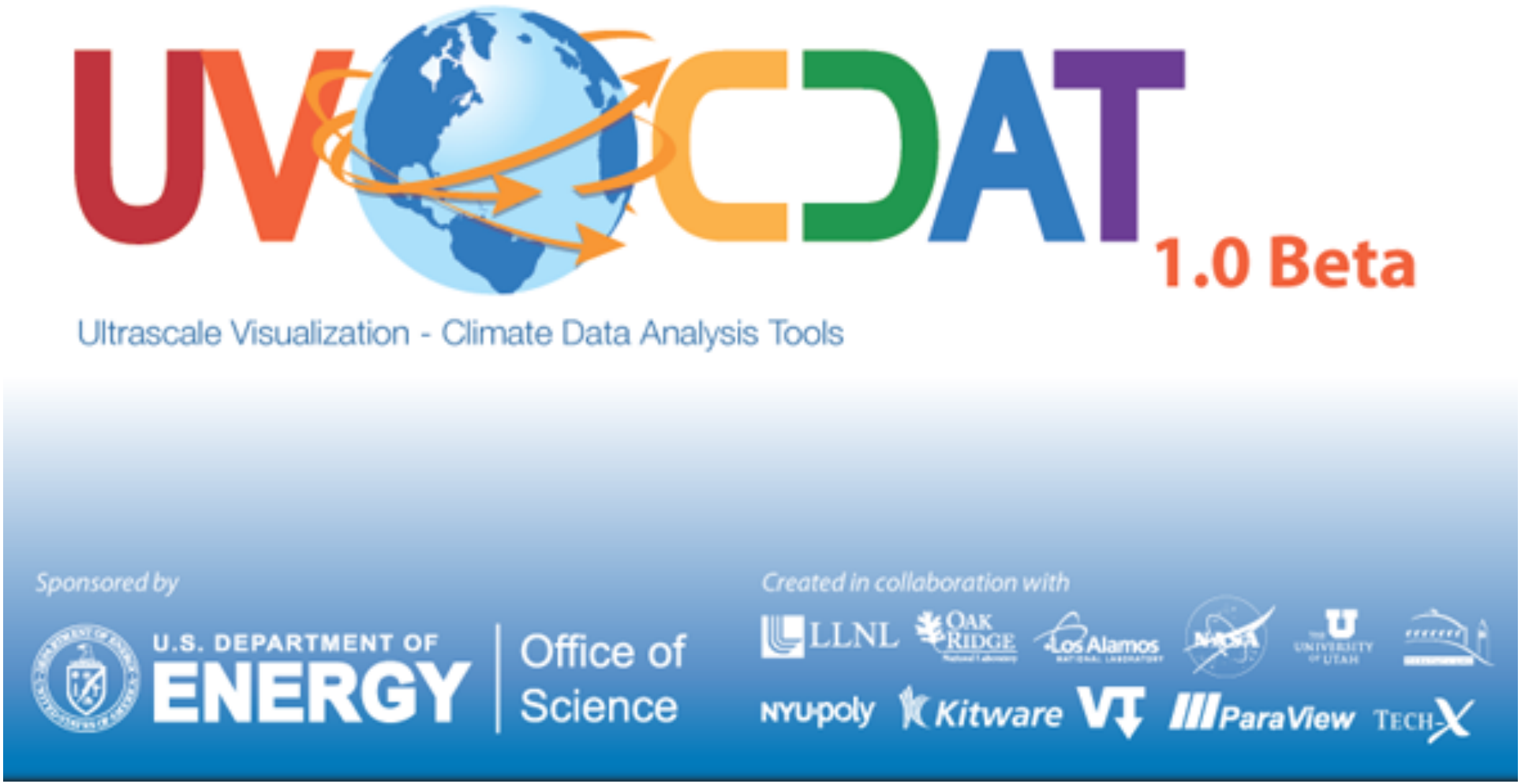


http://uv-cdat.org/

Ultrascale Visualization Climate Data Analysis Tools (UV-CDAT) are targeted for analyzing, diagnosing, and visualizing data for model-intercomparison projects', observation output, and very high-resolution climate-model simulations

Biological and Environmental Research (BER)

- **Climate and Environmental Sciences Division (CESD)**
  - **Earth System Modeling (ESM) Program**

http://science.energy.gov/ber/
http://science.energy.gov/ber/research/cesd/earth-system-modeling-program/

UV-CDAT

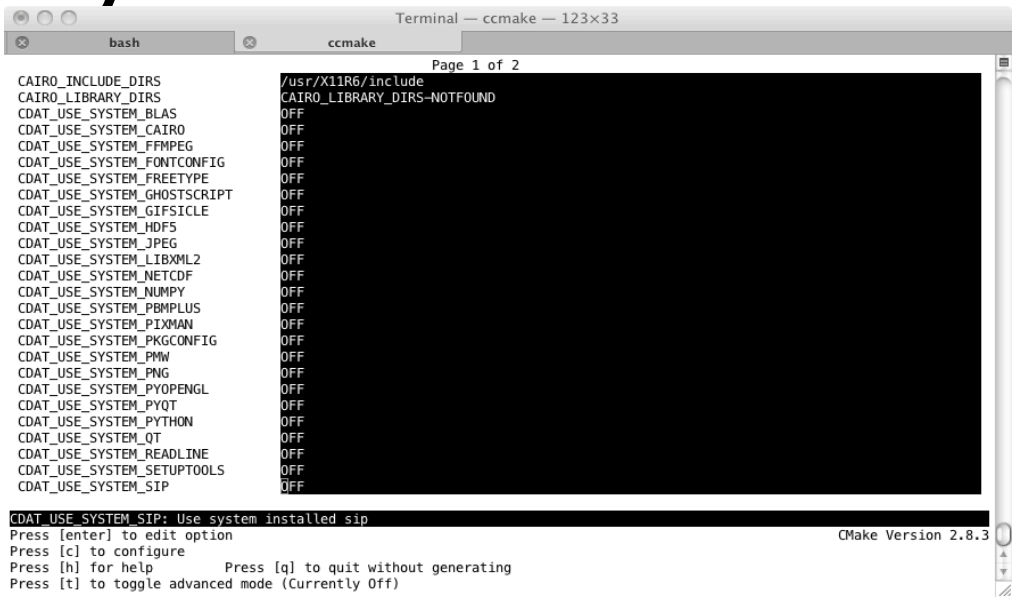
- Delivery Mechanism for both:
  - *Ultrascale Visualization Climate Data Analysis Tools*
  - *Visual Data Exploration and Analysis of Ultra-large Climate Data*
- Observational and Model data
- front-end to a rich set of visual-data exploration and analysis capabilities well suited for climate-data analysis problems

Contributors

- Led by Lawrence Livermore National Laboratory
- Los Alamos National Laboratory
- Lawrence Berkeley National Laboratory
- Oak Ridge National Laboratory
- University of Utah
- Polytechnic Institute of New York University
- NASA
- Kitware
- Tech-X

Building UV-CDAT

Build System



- Problem: UV-CDAT has nearly forty dependencies when fully configured. (CDAT, VisTrails, ParaView,...)
- Solution: Provide a cross platform CMake build system that is easy to extend and maintain as UV-CDAT grows.
  - Modularize sub-packages
  - Uses sub-package existing build systems
- Impact: It will improve and simplify the ability of UV-CDAT developers to configure, build, extend and test UV-CDAT.

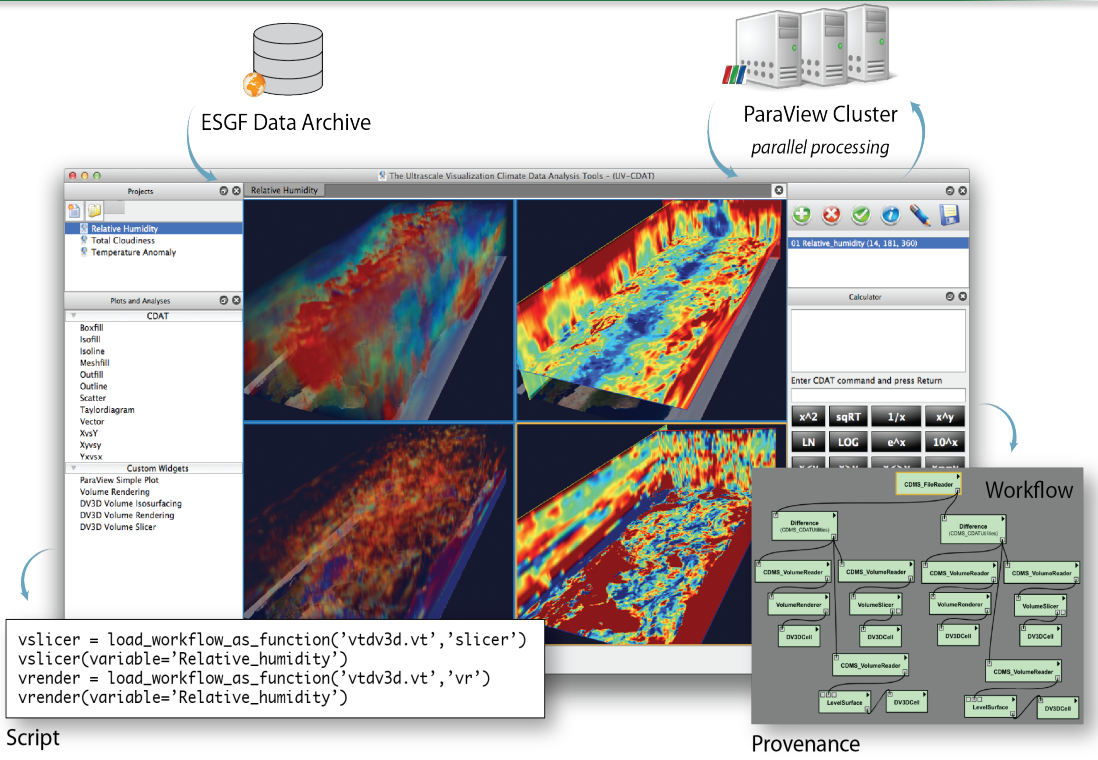
Berk, Dave

Ultra-scale Visualization Climate Data Analysis Tools

Objective: Integrates several existing, widely used open-source data analysis and visualization packages into seamless environment

- CDAT – Climate data analysis/viz
- VTK - Visualization Toolkit
- R – Statistical analysis
- VisTrails – Workflow Provenance
- VisIt, ParaView – 3D Visualization

- Local and remote visualization and data access
- Comparative visualization and statistical analyses
- Robust tools for regridding, reprojection, and aggregation
- Support for unstructured grids and non-gridded observational data, including geospatial formats often used for observational data sets
- Workflow analysis and provenance management

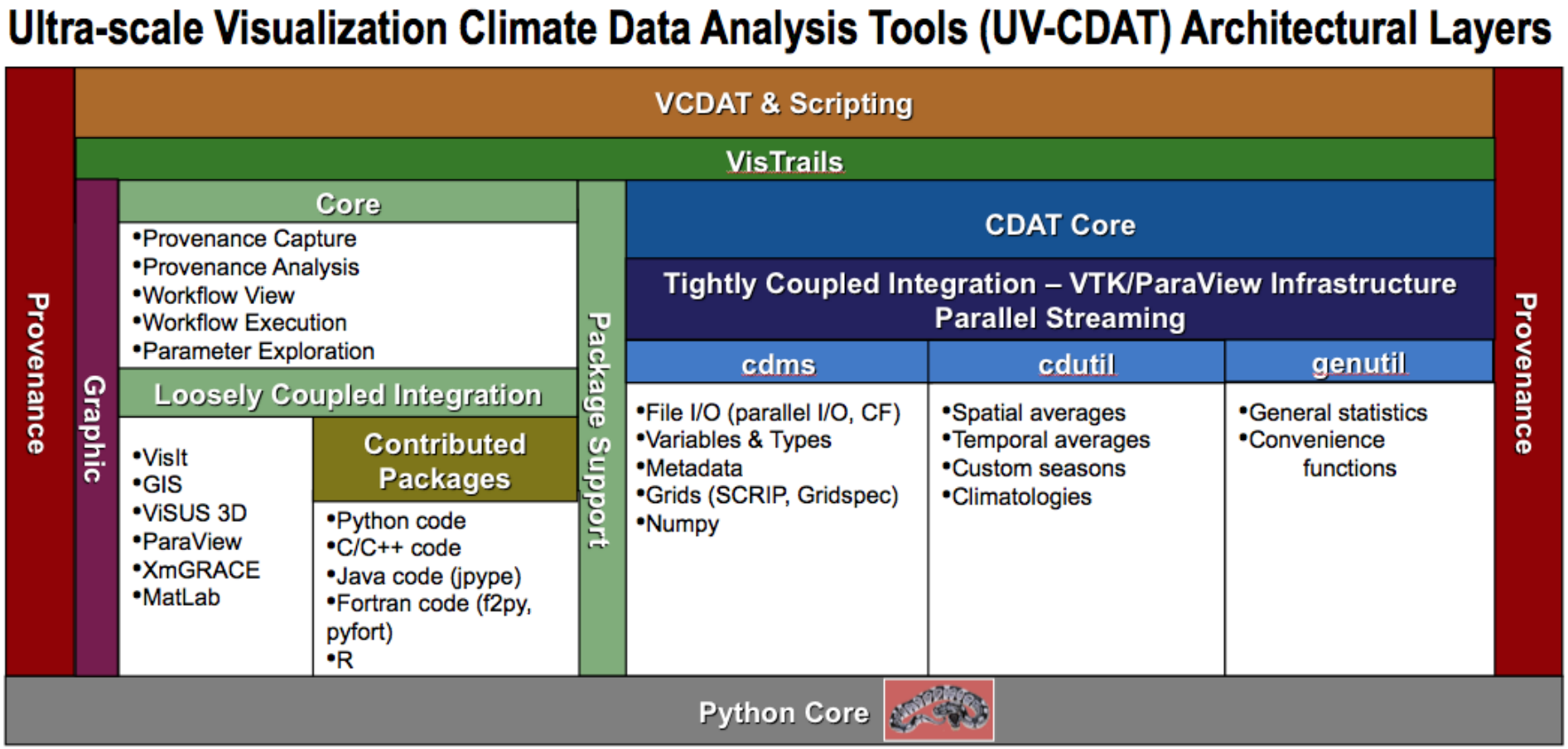


Recent Accomplishment

- ParaView successfully demonstrated the scalability of a new spatio-teomporal pipeline by processing 1/2 TeraByte of data composed of 365 time steps of 1/10 degree POP ocean model in under 2 minutes.

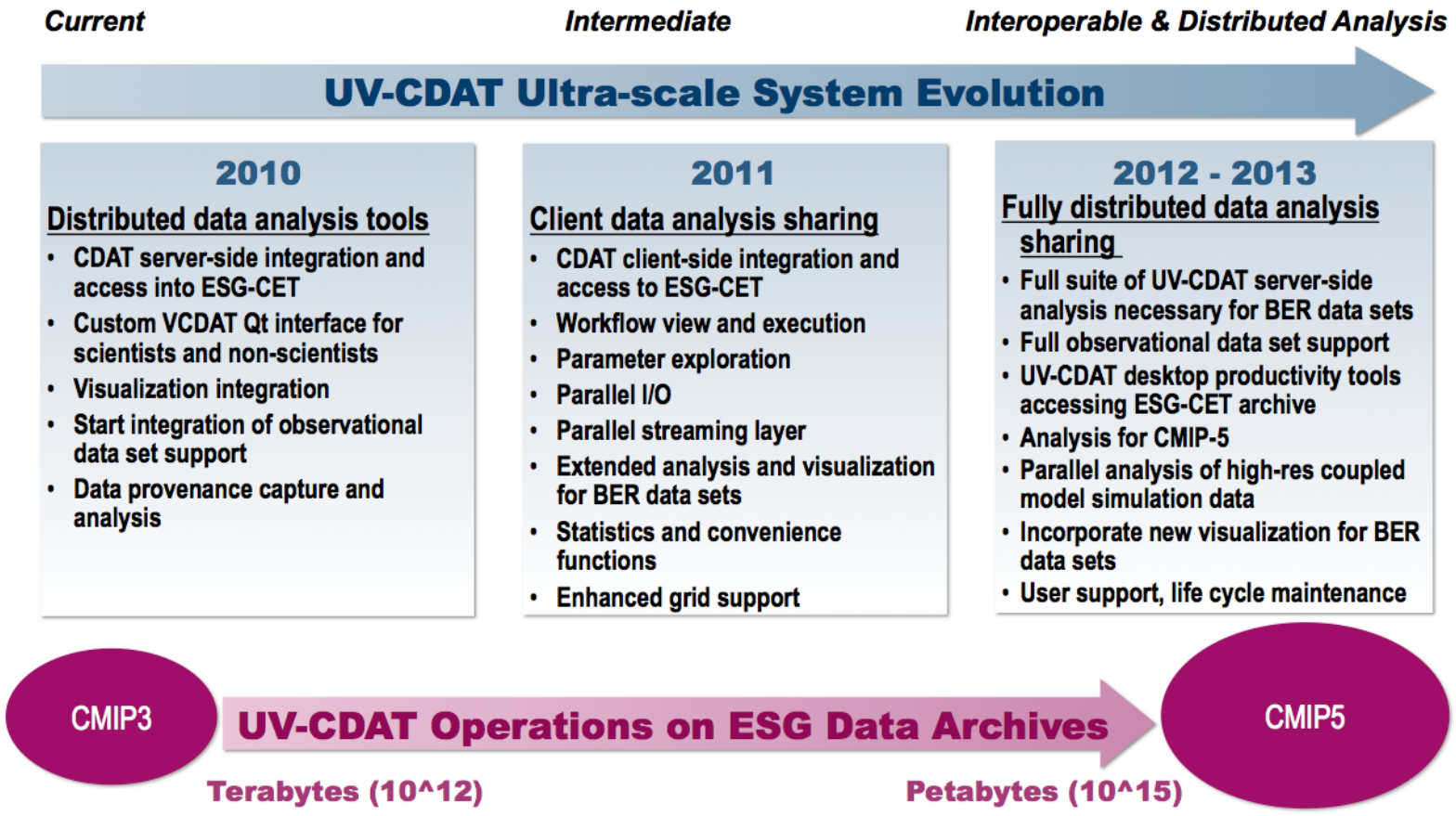
Contact Dean N. Williams (williams13@llnl.gov) for more information or see http://uv-cdat.org/wiki

UV-CDAT Architecture Layers



Milestones and Timeline

The parallelization of UV-CDAT is needed in order to handle extremely large data sets by using existing software (CDAT, VisTrails, ParaView, etc.). This effort will produce a functionally new software visualization infrastructure for climate science.

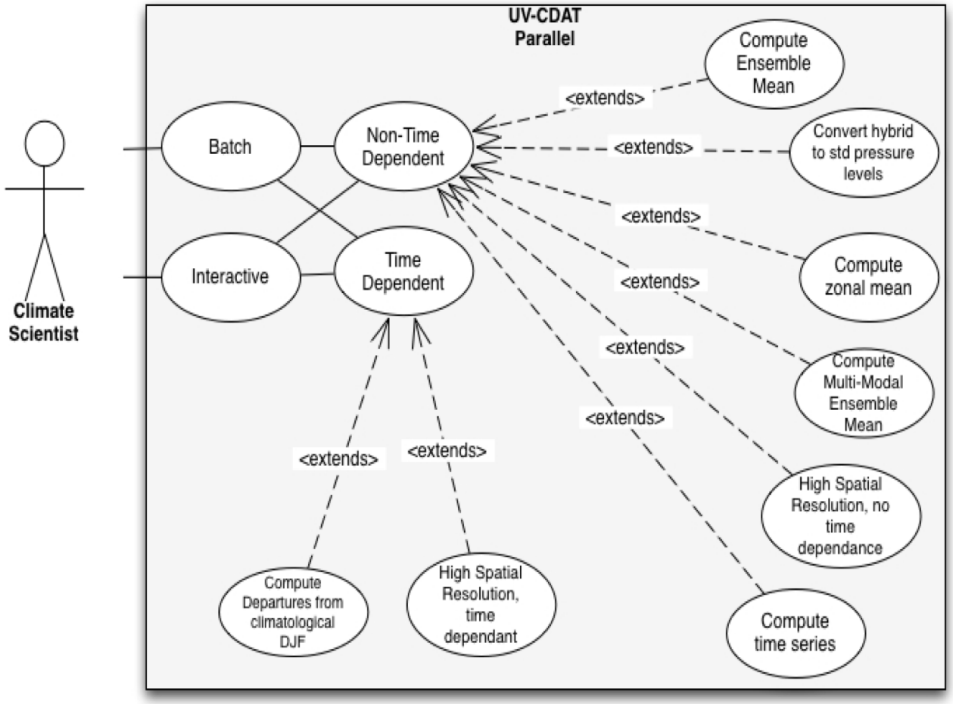


Dean, Dave, Phil

Approach: Use Cases Provide Focus for Effort to Develop More Broadly Applicable Capabilities

1. High spatial resolution, parallel, image sequence production
2. High spatial resolution, parallel, time average
3. Compute ensemble mean
4. Compute average multi-model ensemble mean
5. Compute departures from climatological boreal winter
6. Convert from hybrid to standard pressure levels
7. Compute a time series of a regional average
8. Computing a zonal mean
9. Batch processing
10. Interactive processing
11. Time dependent processing
12. Time independent processing

- UV-CDAT use cases URL:
  - http://www.uv-cdat.org/wiki/UseCases



Dave, Phil, John

LANL/Kitware Contribution – 3 Research Papers

- Williams, S., Hecht, M., Petersen, M., Strelitz, R., Maltrud, M., Ahrens, J., Hlawitschka, M. and Hamann, B. (2011), "Visualization and Analysis of Eddies in a Global Ocean Simulation," Computer Graphics Forum, 30: 991–1000. doi: 10.1111/j.1467-8659.2011.01948.x
- Williams, S.; Petersen, M.; Bremer, P.-T.; Hecht, M.; Pascucci, V.; Ahrens, J.; Hlawitschka, M.; Hamann, B.; , "Adaptive Extraction and Quantification of Geophysical Vortices," *Visualization and Computer Graphics, IEEE Transactions on* , vol.17, no.12, pp.2088-2095, Dec. 2011
  doi: 10.1109/TVCG.2011.162
  URL:
  http://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=6064973&isnumber=6064926
- S. Williams, M. Petersen, M. Hecht, M. Maltrud, J. Patchett, J. Ahrens, and B. Hamann, "Interface Exchange as an Indicator for Eddy Heat Transport," accepted to EuroVis 2012

LANL/Kitware Contribution - Development

- **LANL/Kitware Development Work has produced a new parallel pipeline for massive data with 3 separate parallel readers and has identified POP diagnostics as a high value target for further development.**
  - improved parallel Rectilinear POP Reader.
  - Spatio-temporal Parallel Pipeline.
  - Parallel Unstructured POP Reader.
  - Parallel Unstructured CAM Reader.
  - min, max, mean, stddev using Spatio-temporal Parallel Pipeline.

UV-CDAT: Use Case 1

High spatial resolution, time and space parallel, image sequence production

Problem

UVCDAT must be capable of handling data sets with extremely high resolutions in one or more dimensions. Though existing parallel tools can handle high spatial resolutions efficiently, the infrastructure doesn't exist to support high temporal resolution.

Solution

– Add capability to UVCDAT to allow arbitrary allocations of space and time divisions to parallel resources for both interactive and batch processing.

– Readers, Filters, Renderers all have to be aware of "time compartments".

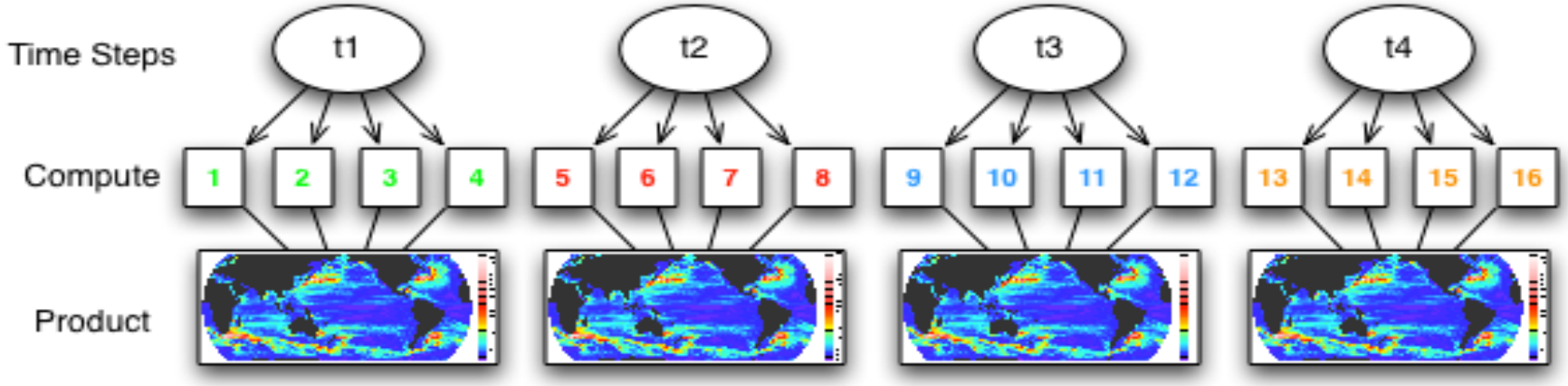
– Ensure increasing computational resources decreases time to solution.

Progress

successfully run a proof of concept on use-case 1 using VTK and reading, extracting surface, and rendering 360, 1.4 GB files, across 4 processor groups (operating on 4 time steps simultaneously) using 40 cores to operate on each time step. This took less than 3.5 minutes. Our user was taking more than an hour to do a similar pipeline.

Future

The progress thus far shows that the algorithm and design will work. We are now integrating this into the UVCDAT framework. The ability to manage this decomposition for interactive use also needs exploration.



Galen, John

UV-CDAT: Use Case 2

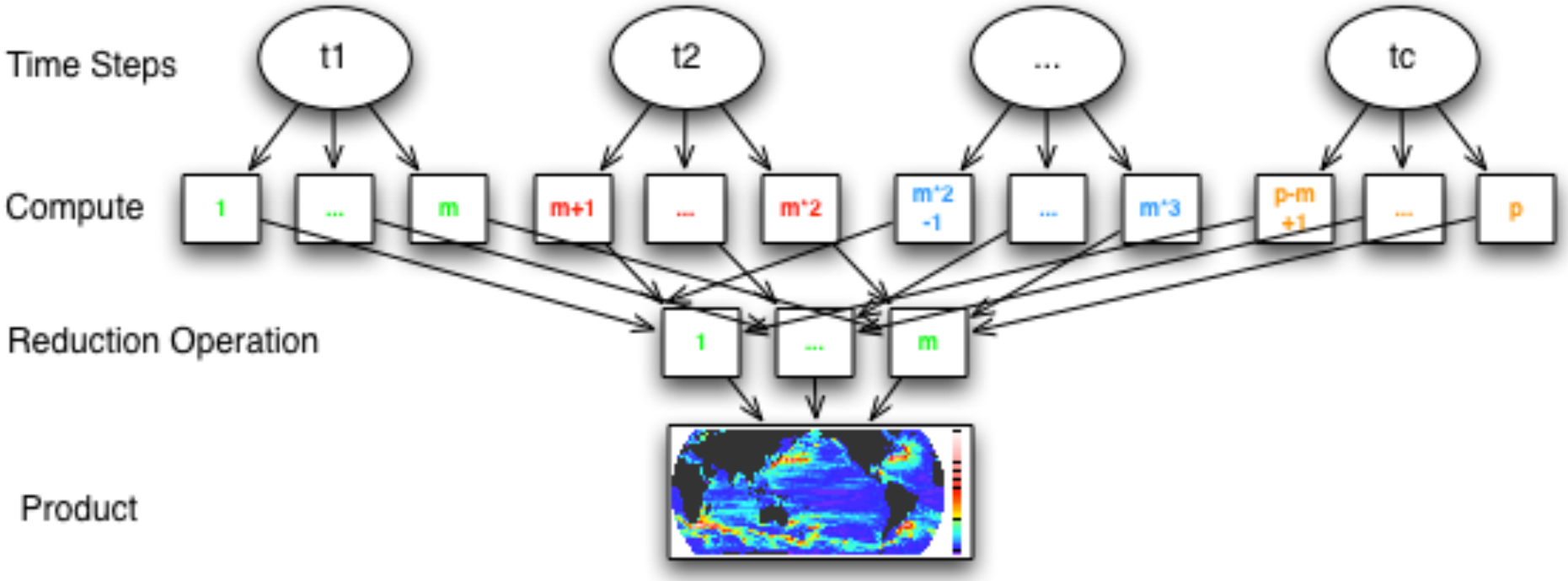
High spatial resolution, time and space parallel, time average

Problem

UVCDAT must be capable of averaging data across multiple time steps to produce a data product based on the result.

Solution

Add framework capability to UVCDAT that allows reduction operations between time compartments. This builds on Use Case 1 implementation which creates time compartments without time dependence.



Let p be the number of processors available. Let m and n be factors of p such that n*m =p, n is the number of processor groups and m is the size of each group.

John, Galen


UVCDAT
LLNL
Oak Ridge
Los Alamos
The University of Utah
NYU-poly
VT
ParaView
Kitware