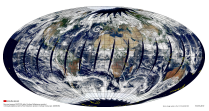
Expressive Feature Characterization for Ultrascale Visualization

Wesley Kendall, Markus Glatter, Jian Huang, Tom Peterka, Rob Latham, and Robert Ross

As datasets grow larger and more complex, leadership computing is becoming more important for analyzing and visualizing data. Through SciDAC, we have designed an end-to-end data visualization system that harnesses leadership computing resources. The system is supplemented with an intuitive programming language interface for feature characterization, addressing the perpetual challenge of succinctly describing and quickly extracting features of interest from ultrascale datasets.

Full Range Global Climate Studies - MODIS



- 31,200 x 21,600
- 416 timesteps
- 2 variables
- 1.1 TB

Our driving application is to extract qualitative climatic events from MODIS and visualize them in real time with no preprocessing. The dataset is stored as time-varying netCDF files.

Driving Application

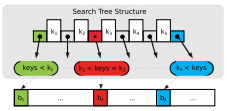
Intrepid – ANL
- 40,960 cores

Jaguar XT4 – ORNL
- 31,328 cores

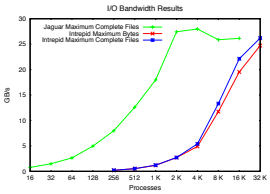
Our system was developed to harness the many characteristics of leadership computing resources, from I/O to communication and computation. The scalable design resulted in applications timings under a minute.

I/O

A data structure similar to a B-tree is used for querying datasets.



During on-the-fly processing, the data items are distributed in such a way to maximize load balance. The data is also filtered to remove "fill" values like ocean points.



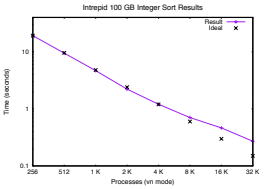
Files can be assigned to processes in various ways to increase I/O bandwidth.

We achieved up to 28 GB/s on Jaguar and 26 GB/s on Intrepid using various I/O heuristics.

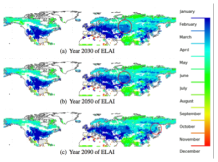
Process

Query

Our system can reduce queried results so that sophisticated analyses can be performed.



In a 100 GB sort benchmark, we obtained excellent scalability up to 32 K processes.

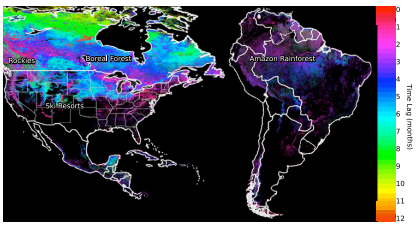


Complex problem spaces are explored by issuing queries similar to regular expressions.

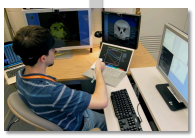
Green-Up: ???[SNO < 0.4]*T[SNO > 0.4]?*
Drought: ?*[WAT < 0.3 & VEG < 0.5]^4?*
Time-Lag: ?*T[WAT > 0.8]?*T[VEG > 0.4]?*

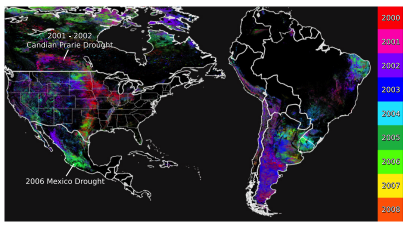
Reduce

Visualize



Time-lag analysis of length of snow season.





Drought analysis colored by worst drought.

U.S. DEPARTMENT OF ENERGY
UltraVis SciDAC Institute for Ultrascale Visualization
Office of Science U.S. DEPARTMENT OF ENERGY

Software Downloads
BIL - http://seelab.eecs.utk.edu/bil.php
SQI - http://seelab.eecs.utk.edu/sqi.php