# Interactive Analysis of Petascale Climate Datasets

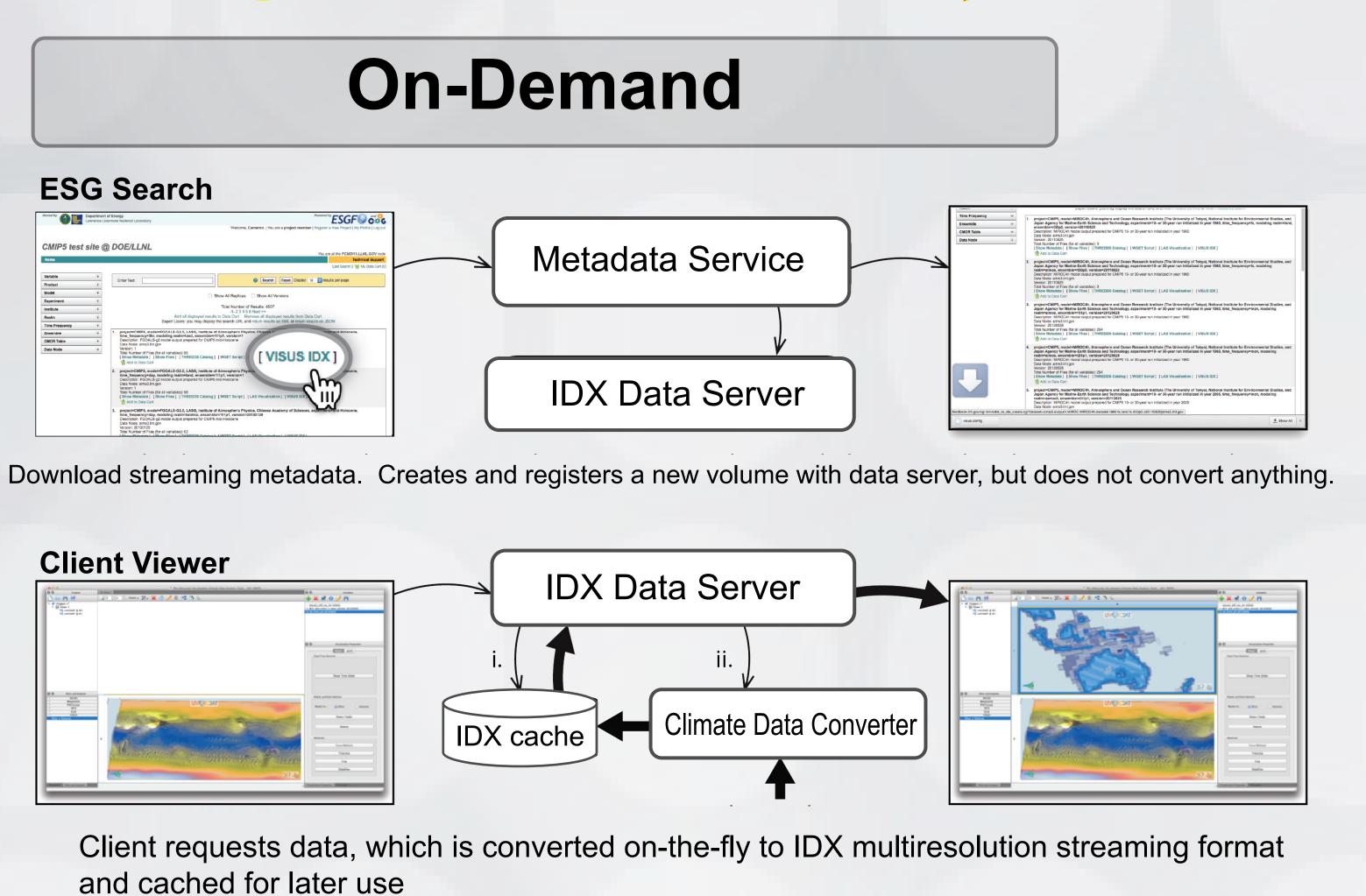
Cameron Christensen, Shusen Liu, Peer-Timo Bremer, Valerio Pascucci

#### **Challenge: Interactive Analysis of Petascale Datasets**

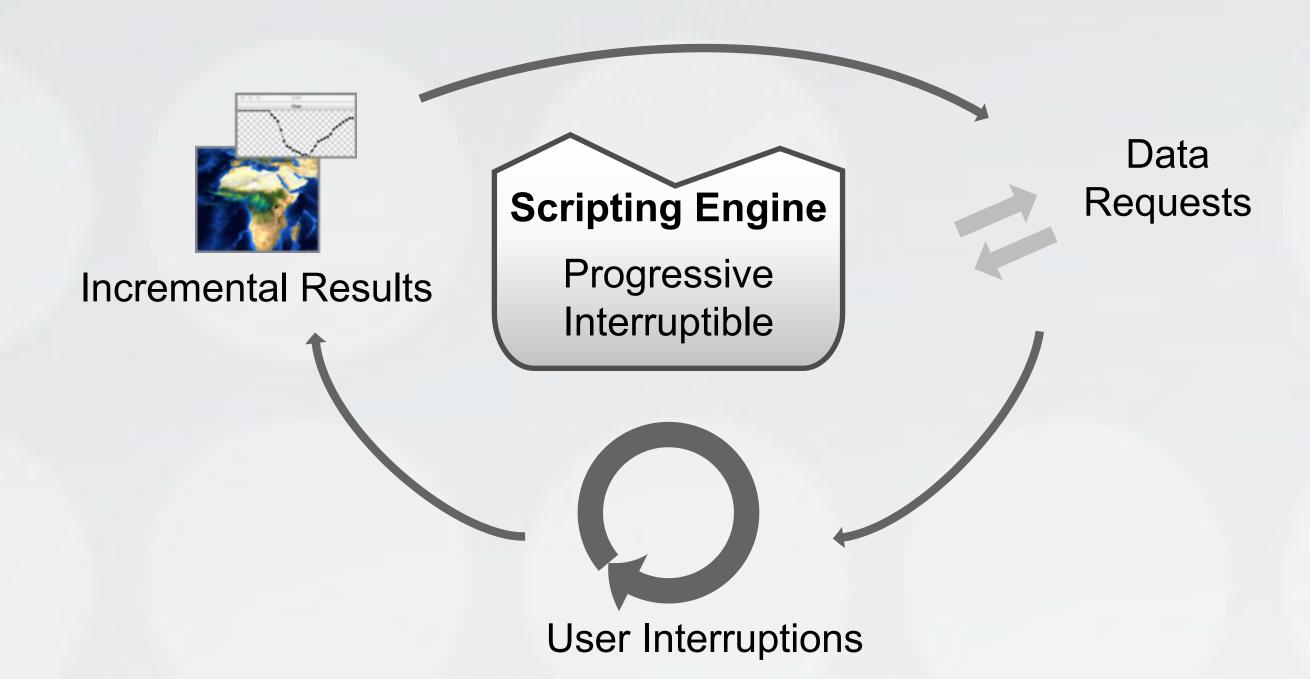
- Spatiotemporal datasets, such as global climate simulations, are large and remotely located.
- Offline batch processing using large computing clusters means analyses can take hours to days.
- Many communities are simply unable in any practical way to analyze large datasets.

Problems result in limiting exploratory investigation.

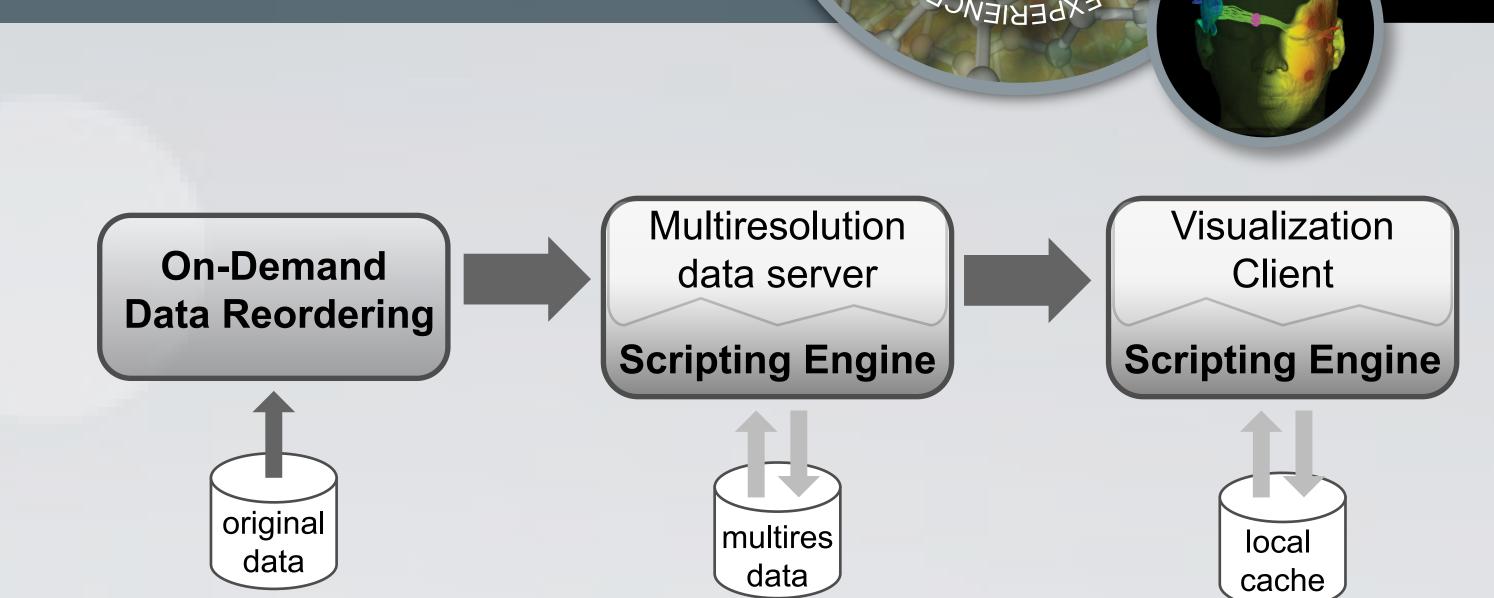
Solution: Embedded Domain-Specific Language (EDSL) for Incremental Multiresolution Processing enables interactive analysis and visualization of massive, disparately-located data.



# Scripting Engine

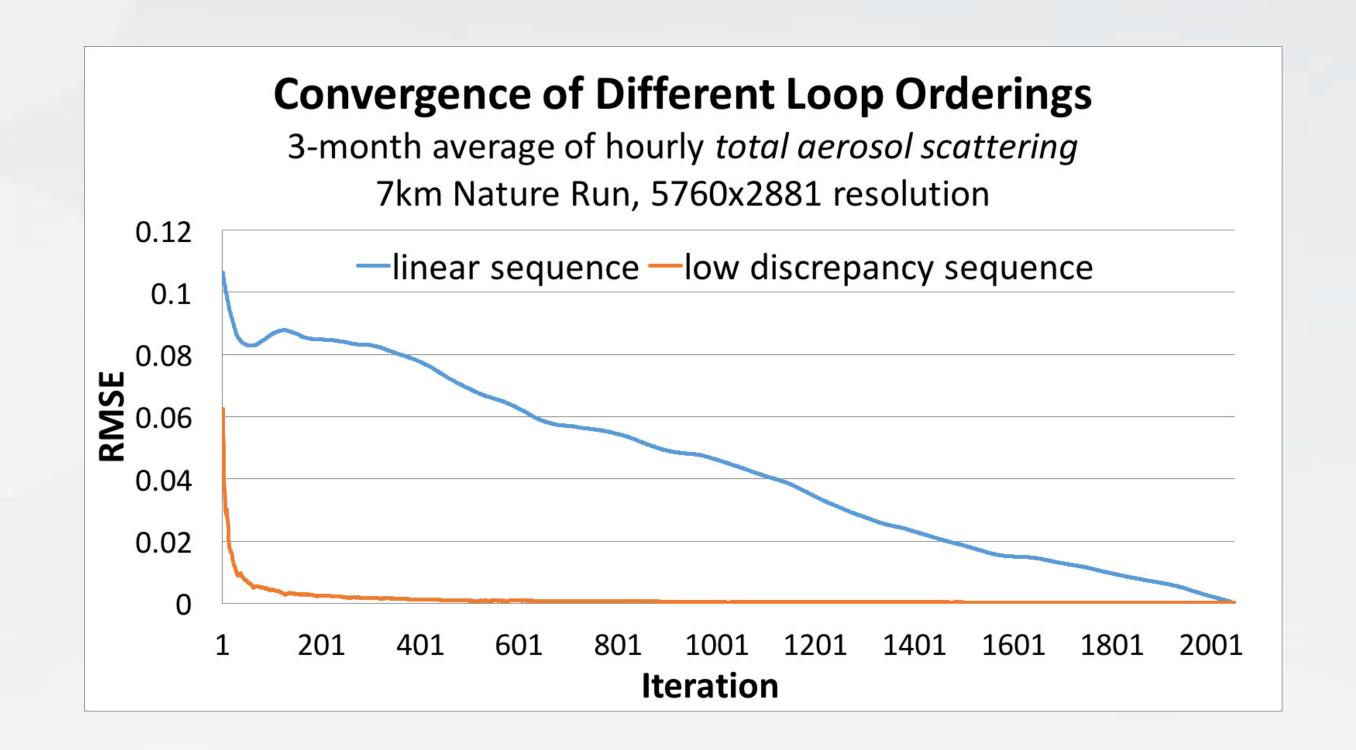


Interactive script interpreter continuously processes data requests, publishes incremental results, and responds immediately to updated user input with *doPublish* extension



## Loop reordering

Faster convergence of incremental results with alternate loop orders



### Case Studies

ii. If not cached, requested data is converted for cacheing

Compare massive disparately located spatiotemporal datasets

. Data server first checks idx cache.



Climate scientists used our EDSL to quickly discover previously unknown errors in a widely-used dataset: the first day of each month of the 3D fields in the dataset was mistakenly repeated for the rest of the days!

