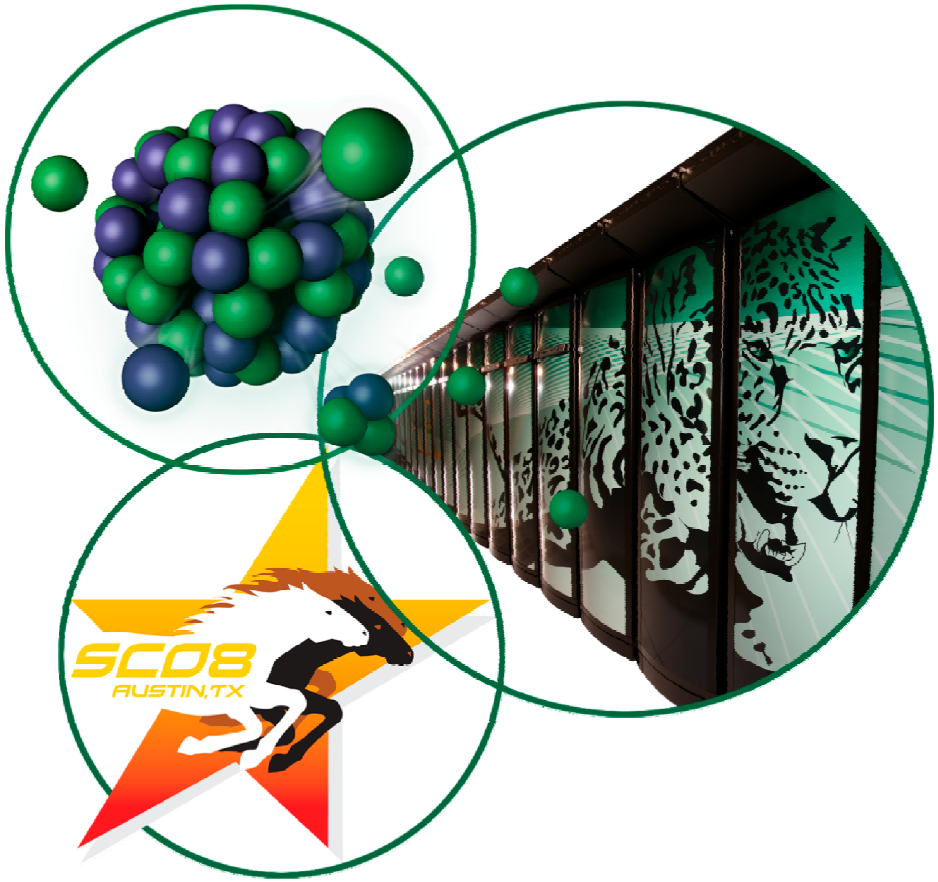
Scalable Data Servers for Large Multivariate Volume Visualization



Presented by

Markus Glatter
Colin Mollenhour
Jian Huang
Jinzhu Gao

The University of Tennessee

UltraVis
SciDAC Institute for Ultrascale Visualization

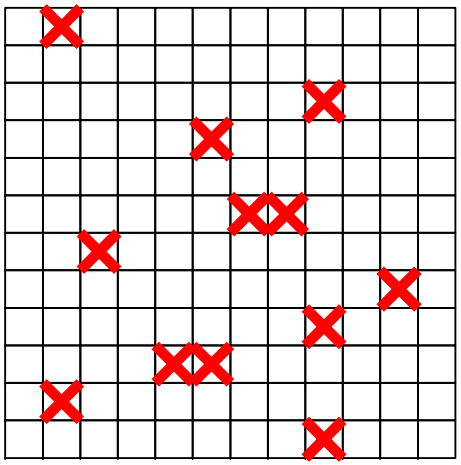
The University of Tennessee
Computer Science Department

Motivation

- Volumetric multivariate time-varying data sets are often complex and massive in size
- Efficient and interactive selection and visualization of subsets is not easily done, as data sets will not fit in core
- Real-word example: Astrophysics simulation of supernova generates 300 time steps of a spatial resolution of 864 × 864 × 864 with multiple attributes stored at each voxel
- Few systems can adequately support (interactive) visualization of such enormous data sets

Our approach: Overview

Parallel data management and effective data culling

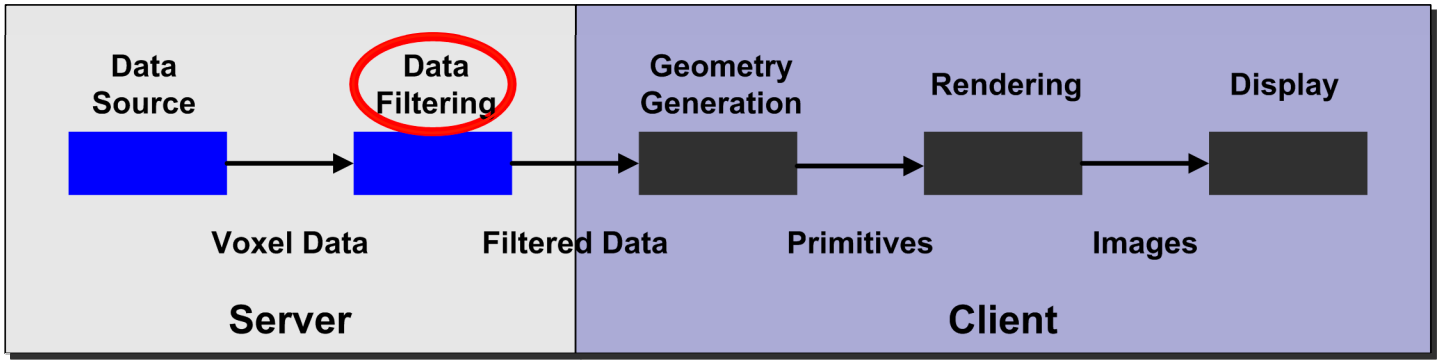


UltraVis
SciDAC Institute for Ultrascale Visualization

The University of Tennessee
Computer Science Department

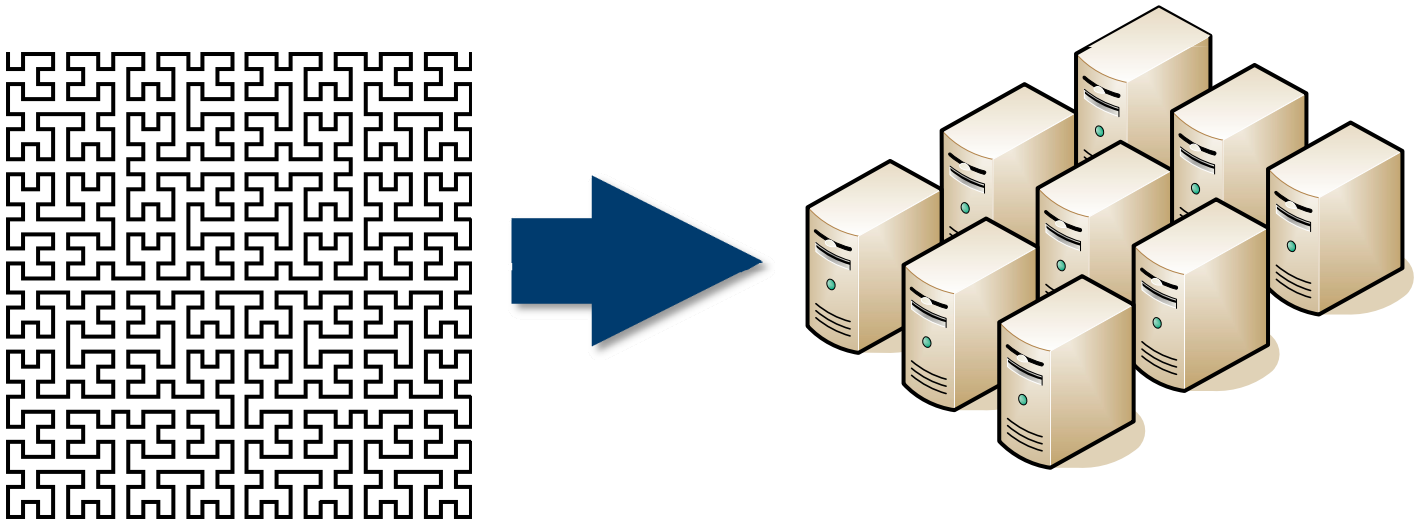
Data culling

- Parallel data culling on the granularity of individual voxels
- Process is driven by compound boolean-range queries
- Cull unnecessary voxels as early in the visualization pipeline as possible (i.e., on the server side)



Parallel data management

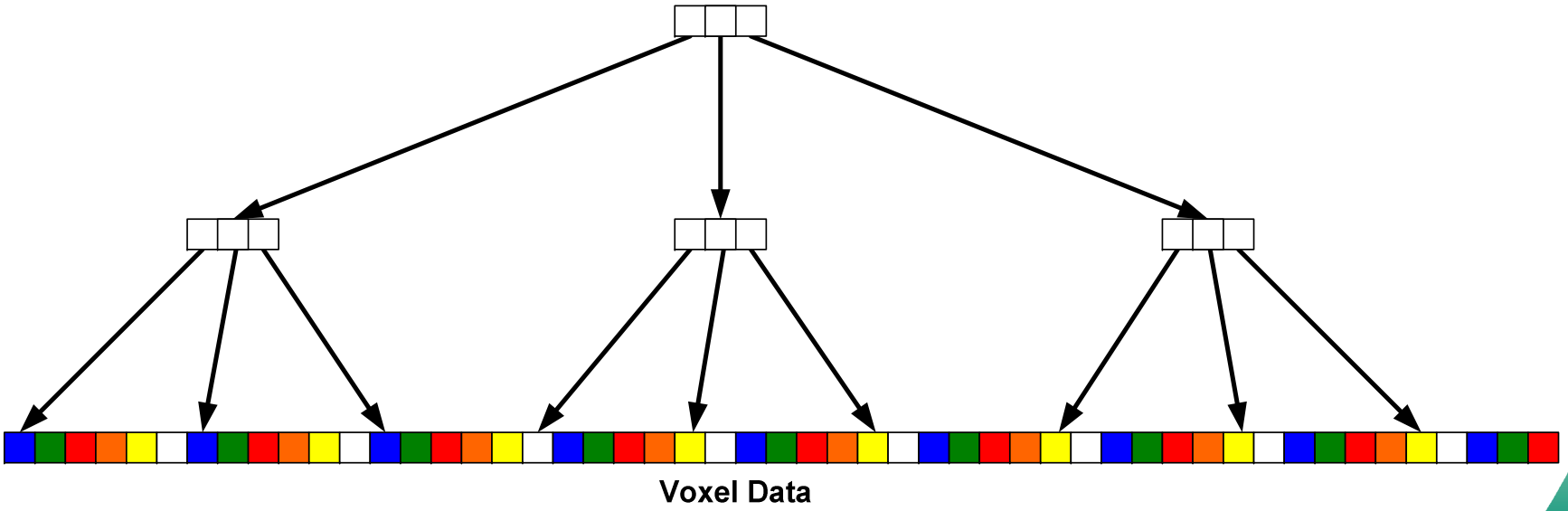
- The entire data set is distributed among a number of networked independent computers according to space-filling curve order in the high-dimensional attribute space
- Load balancing can be achieved among data servers, independent of the range queries



ULTRAVIS
SciDAC Institute for Ultrascale Visualization
THE UNIVERSITY of TENNESSEE
Computer Science Department

Parallel data management

- Each data server leverages a very compact data structure similar to a B-tree to maintain all voxels distributed to it
- Queries to a server can be responded to quickly with a query rate of 4.2 million relevant voxels per second across a local area network

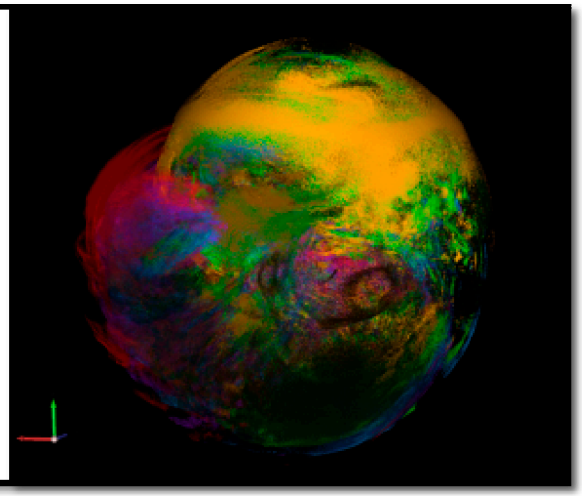


Requirements

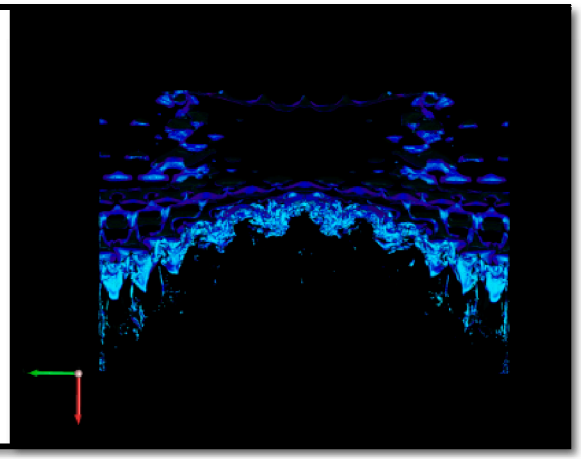
- System of networked workstations without any special hardware or software
- All data servers combined need to hold the entire data set in main memory in a compressed form
- Queries will need to fit into client's main memory to render
- If requirements cannot be met, the system will still be operational, but performance will be hampered

Results—Data sets

TSI—Time-varying (6 time steps)
Multivariate (11 attributes)
864 × 864 × 864
105 GB of raw data

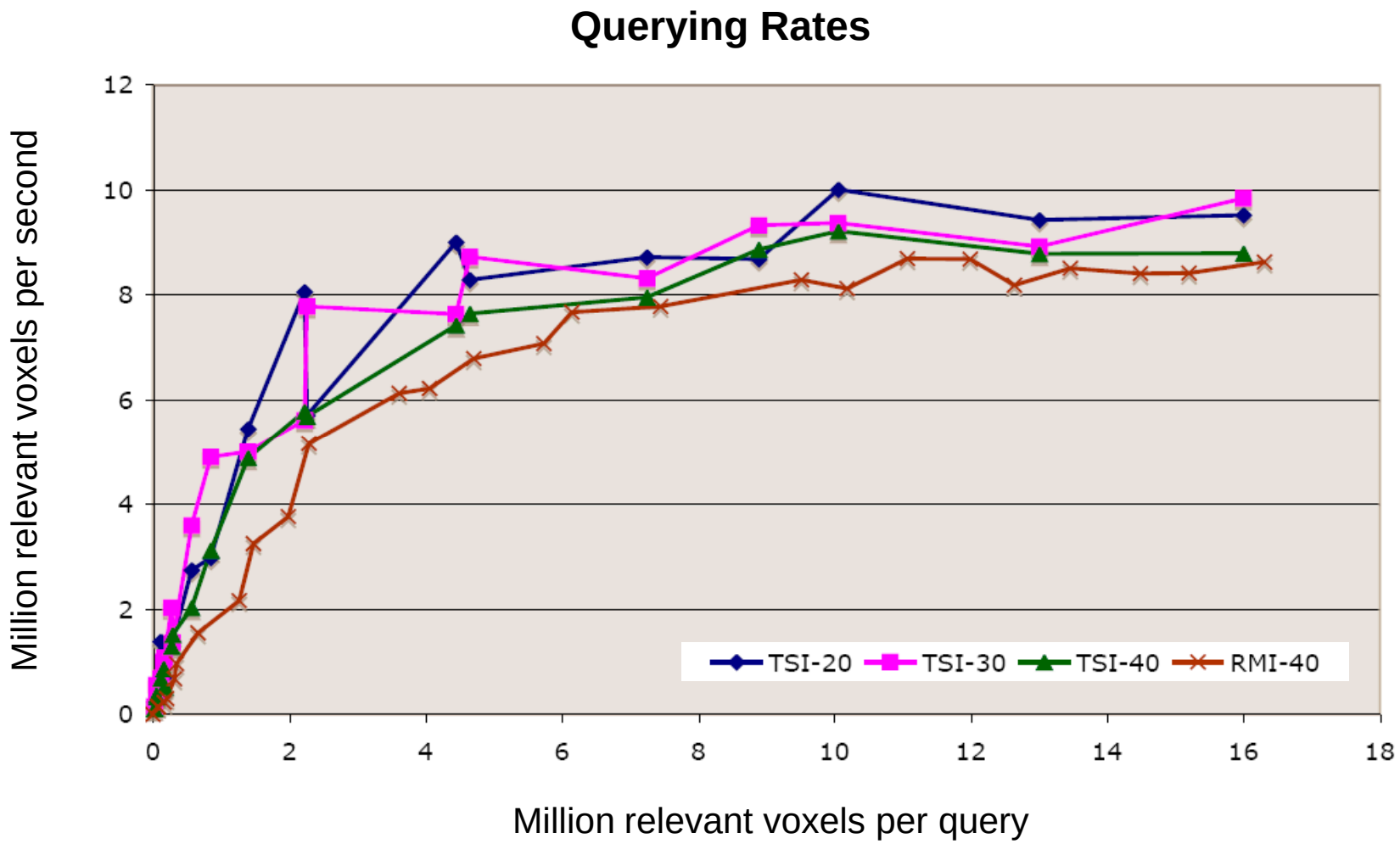


RMI—Time-varying (3 time steps)
Multivariate (11 attributes)
1024 × 1024 × 960
82 GB of raw data



Tests run with 20, 30, and 40 servers

Results



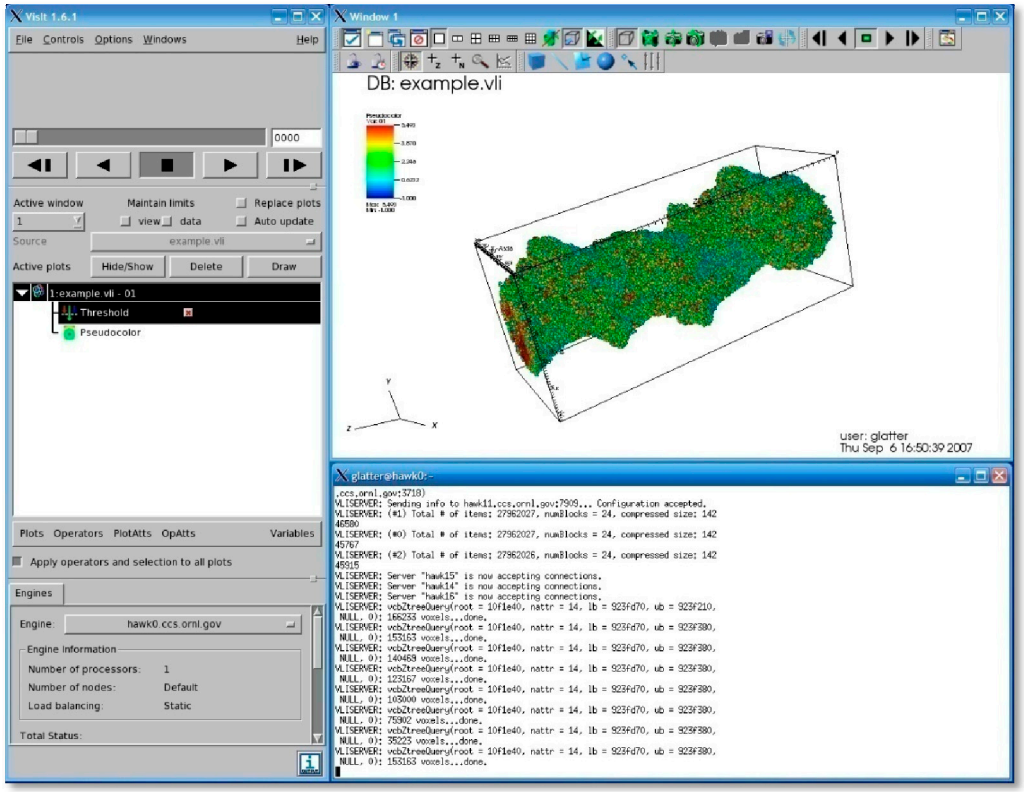
UltraVis
SciDAC Institute for Ultrascale Visualization
The University of Tennessee
Computer Science Department

Results

- **Network seems to be the bottleneck of the system**
  - **Even with a Quadrics Elan3 network (single-direction bandwidth of 400 MB per second)**
- **We can query at 144 MB/s or about 9 million voxels per second in parallel**
- **The load imbalance among servers ranged between a mere 0.012% to 0.155% for medium and large queries**

Integration into VisIt

Scalable data servers have been integrated into the free interactive parallel visualization and graphical analysis tool VisIt as a module in version 1.7



http://www.llnl.gov/visit

Contact

Markus Glatter

The University of Tennessee–Knoxville
(865) 974-5400
glatter@cs.utk.edu