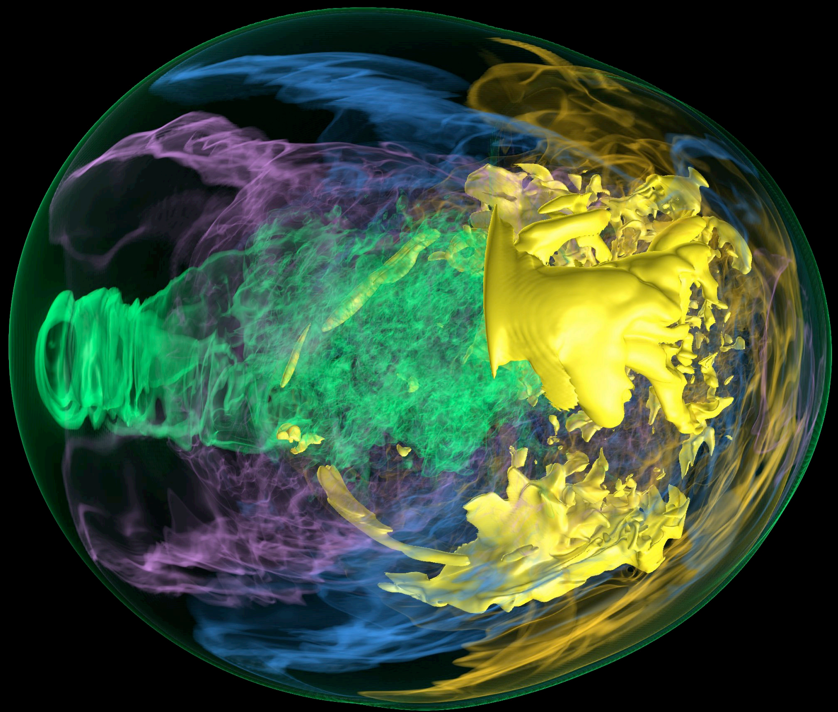
Factors in Large Scale Parallel Visualization



Volume rendering of angular momentum in a simulation of hydrodynamics during the early stages of a core-collapse supernova.

Tom Peterka

tpeterka@mcs.anl.gov

Mathematics and Computer Science Division

Argonne NATIONAL LABORATORY

... for a brighter future

UltraVis SciDAC Institute for Ultrascale Visualization

www.ultravis.org

U.S. Department of Energy

UChicago Argonne LLC

Office of Science U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory managed by UChicago Argonne, LLC

A Growing Rift

We are computing more data, faster than we can analyze.

Paul Ricker: 1 out of 10 timesteps are saved in a 1K^3 cosmological simulation. This consumes 15 TB of data. Storage capacity is the limitation. 4096^3 sampled to 512^3 (512:1) before visualizing

Aleks Obabko: 10K-100K timesteps computed in NEK5000 CFD code. 100-1K restarts are saved.

Rob Jacob: 100:1 data reduction between compute and analysis in CCSM.

Thomas Cram: 10% of computed data from GCRM are looked at.

Janine Bennett: Between 1 and 5 timesteps out of 100 in S3D are saved and analyzed

Causes:
-storage bandwidth and storage capacity
-analysis tools capability
-human perception

The Narrows

Storage is the bottleneck.

10,000:1

1:2000

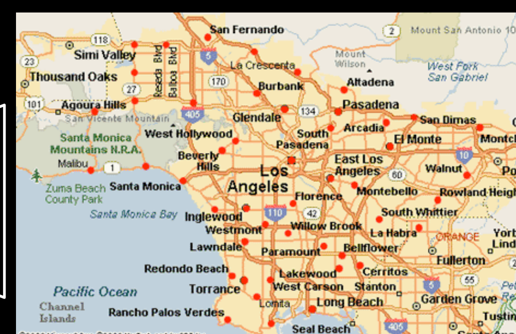
Intrepid BG/P
557 TF/s

Enterprise Storage
50 GB/s

Eureka QuadroPlex Cluster
111TF/s

9,000:1

1:2500



California
Pop. 36.7 M

Tahoe City
Pop. 4 K

Los Angeles County
Pop. 10 M

Effect on Analysis

Large scale parallel visualization on HPC machines

This model will not scale indefinitely:

sim → write → read → analysis/vis → write → read → view

dataset | images

supercomputer | vis cluster | desktop

The increasing demands for analysis and visualization can be met by performing more analysis and visualization tasks directly on supercomputers traditionally reserved for simulation.

Potential benefits: Increased performance, reduced cost, tighter integration of analysis and visualization in computational science.

Factors in Large Scale Visualization

Not surprisingly similar to simulation

Domain Decomposition

Grid topology
Decomposition strategy
Neighbor cells
Load balance
Static / dynamic distribution

Scalability

Strong scaling
Weak scaling
Max. effective number of processes
Efficiency
Isoefficiency function

Performance

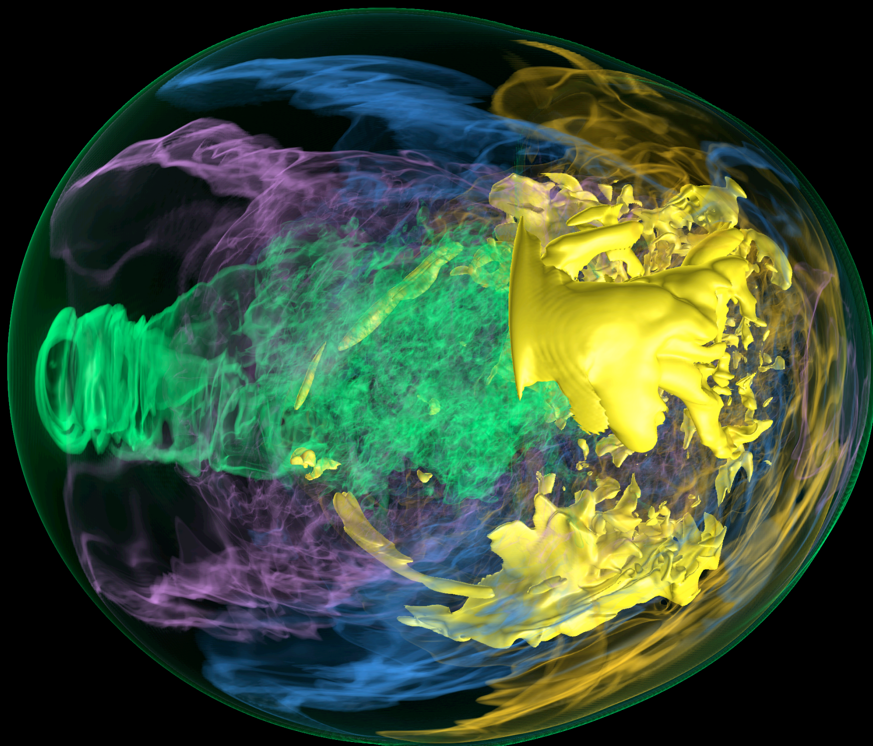
Overall time to completion
Component time
Time distribution

Data Movement

Nature of algorithm
Communication signature
Storage patterns

Applications

Astrophysics, climate modeling, nuclear reactor design (so far)

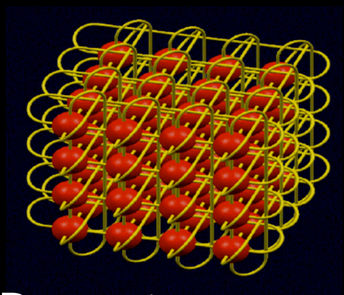


Volume rendering of angular momentum during shock wave formation in core-collapse supernova dataset, courtesy of John Blondin, NCSU. Structured grid of 1120^3 data elements, 5 variables per cell.

Velocity magnitude in MRI simulation from Princeton Plasma Physics Laboratory. The goal is concentric, uniform flow for maximum heat transfer. Dataset courtesy of Paul Fischer and Aleks Obabko, ANL.

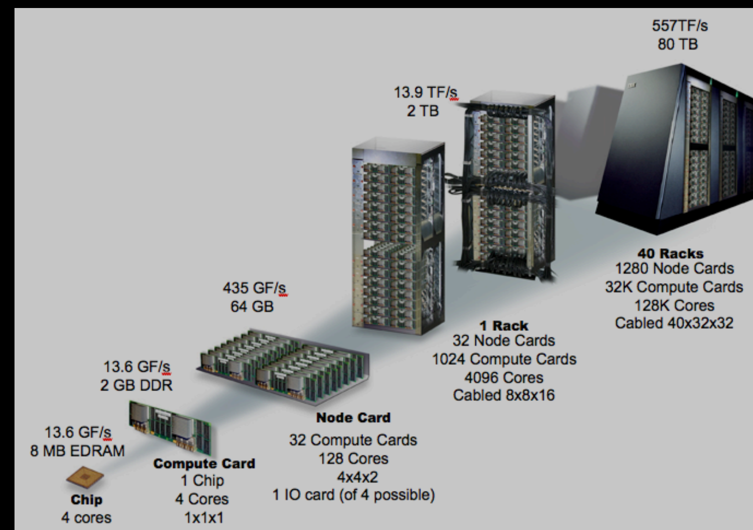
Hybrid of Systems and Visualization Research

We wear many hats.

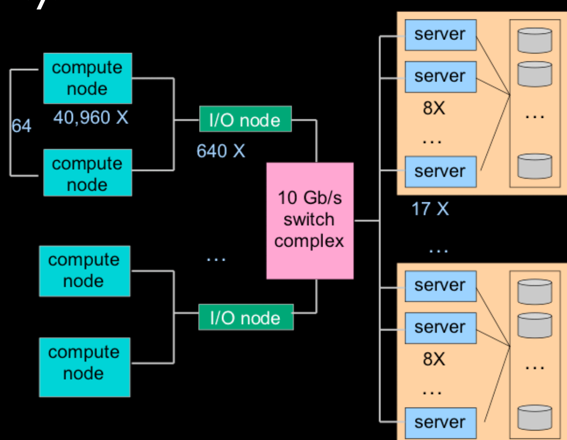


3D torus interconnect offers multiple paths between nodes for high bandwidth and low latency.

Argonne Leadership Computing Facility



The Blue Gene/P features a highly scalable compute architecture composed of 160,000 compute cores. Peak performance is 557 TF

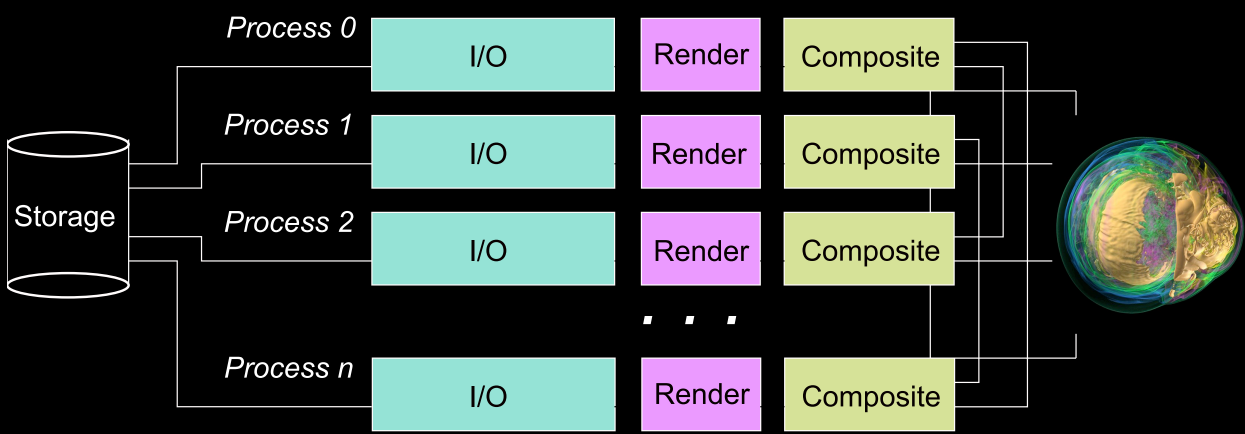


PVFS-2 parallel file system provides 50 GB/s peak aggregate b/w and 5 PB total capacity.

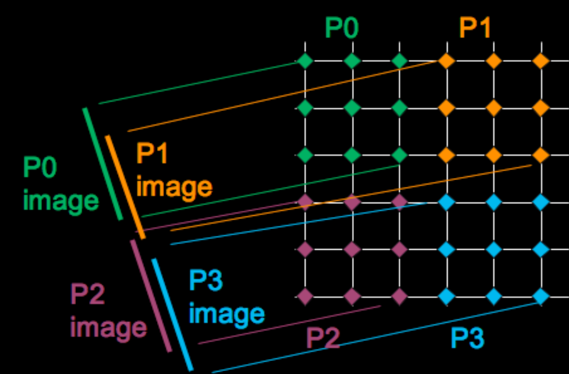
Cray XT4 at Oak Ridge Leadership Computing Facility

Parallel Algorithms

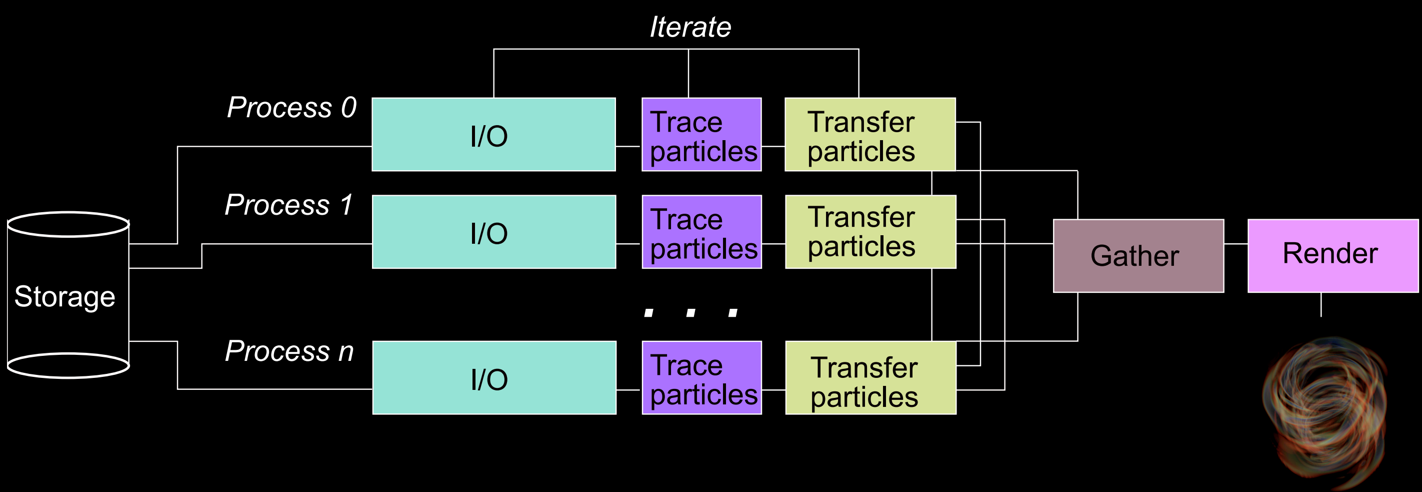
Divide, conquer, reunite



Sort-last parallel rendering requires a merging (compositing) of resulting images into one final image.



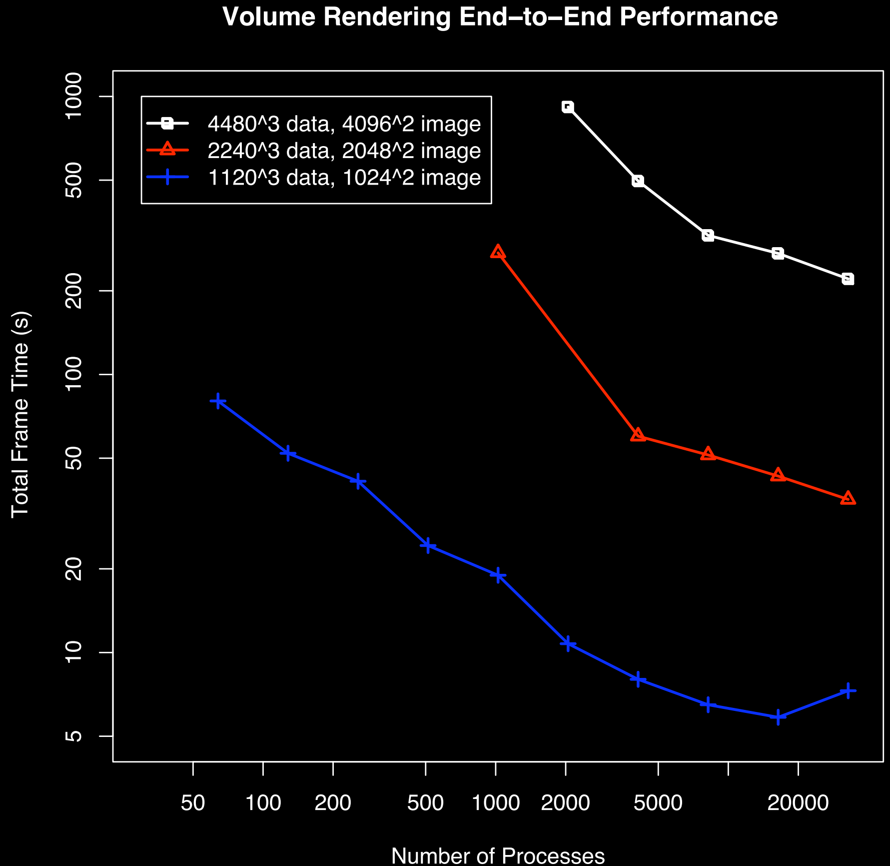
Parallel structure for volume rendering algorithm consists of 3 stages performed in parallel



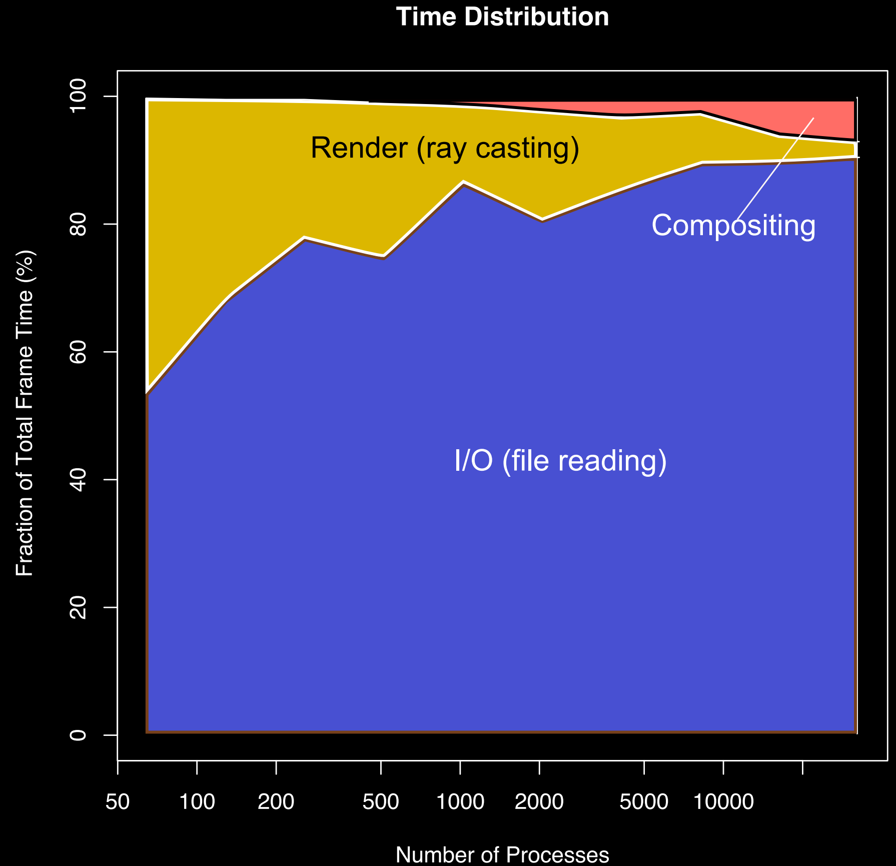
Parallel structure for flow visualization algorithm consists of iterations of particle tracing and transfer, followed by a rendering stage.

Performance

Scalability and time distribution



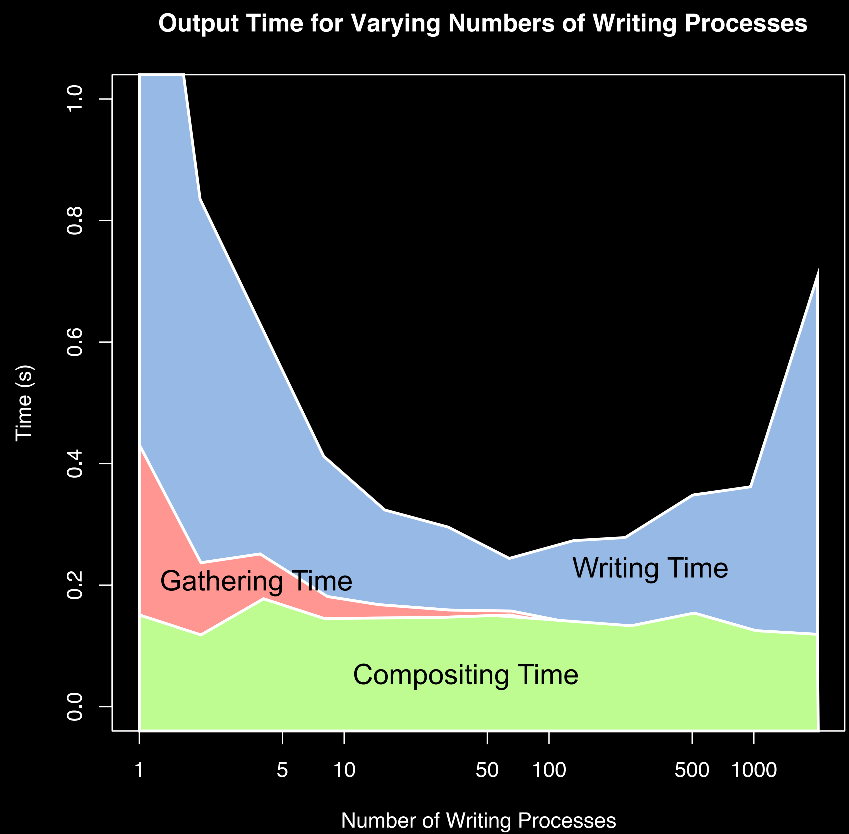
Scalability over a variety of data, image, and system sizes. A number of performance points are possible, depending on system resource availability and other factors.



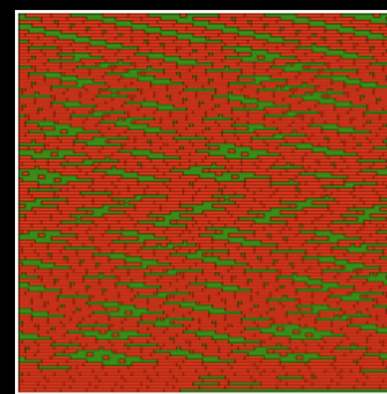
The relative percentage of time in the stages of volume rendering as a function of system size. Large visualization is primarily dominated by I/O and secondarily by communication.

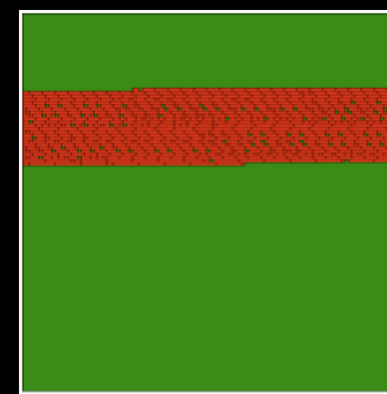
I/O Matters

Use of parallel I/O techniques and both system-level and application-level optimizations can produce drastic speedups.



Different numbers of processes can affect the overall output time for a visualization. Currently, we can control this more accurately in the application than with MPI-IO.

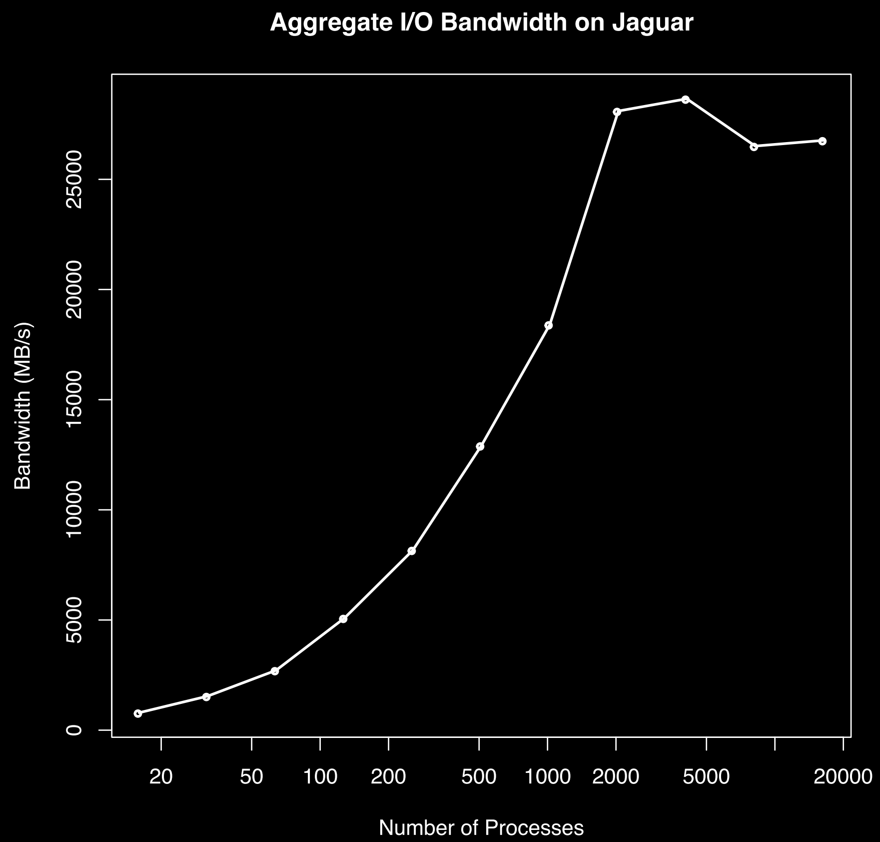


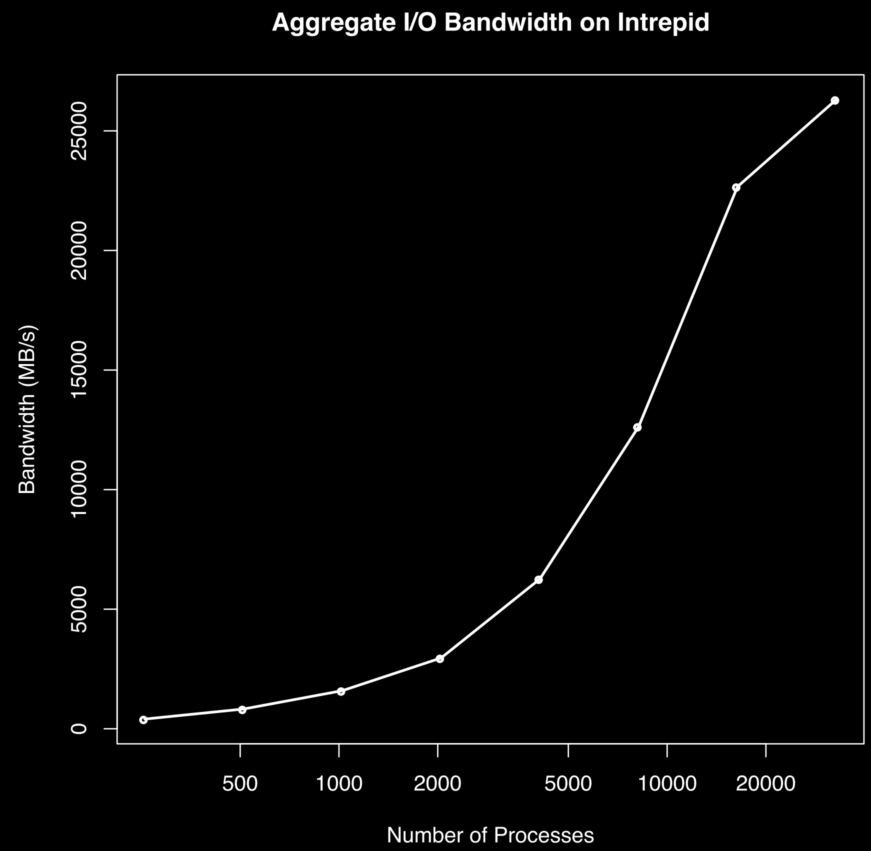


Changing data file layout can drastically improve I/O performance. In a visual representation of reading one variable out of five in a data file, red blocks show physical file blocks that are accessed. Left, most file blocks are read, while a different file organization at right leads to fewer blocks read, and they are contiguous.

Aggregate I/O Bandwidth

Critical for performance of time-varying data visualization.





Aggregate bandwidth for reading 416 time steps of MODIS full resolution data on Jaguar's Lustre and Intrepid's PVFS file systems. 28 GB/s is 75% of IOR benchmark.

Recap

Lessons learned and the road ahead

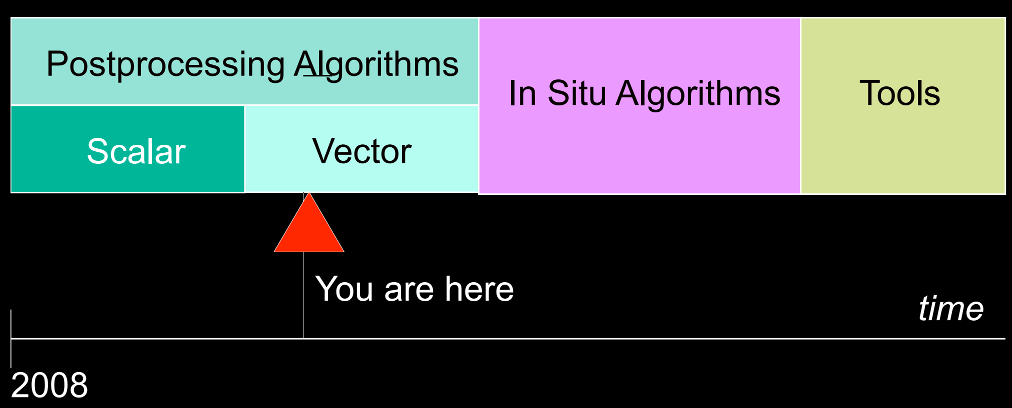
Successes

- Demonstrated scaling on large data and images
- Improved compositing
- Improved and benchmarked I/O

Ongoing

- Other grid topology
- In situ
- Adoption into tools and libraries
- Heterogeneous systems

Road Map



Take-away

- HPC has appropriate resources for visualization: massive parallelism, storage, and interconnect capability.
- Visualization algorithms can be developed that scale with the machine and problem size.

Further Reading

References

Peterka, T., Goodell, D., Ross, R., Shen, H.-W., Thakur, R.: A Configurable Algorithm for Parallel Image-Compositing Applications. Proceedings of SC09, Portland OR, November 2009.

Kendall, W., Glatter, M., Huang, J., Peterka, T., Latham, R., Ross, R.: Terascale Data Organization for Discovering Multivariate Climatic Trends. Proceedings of SC09, Portland OR, November 2009.

Peterka, T., Yu, H., Ross, R., Ma, K.-L., and Latham, R: End-to-End Study of Parallel Volume Rendering on the IBM Blue Gene/P. Proceedings of ICPP'09 Conference, Vienna, Austria, September 2009.

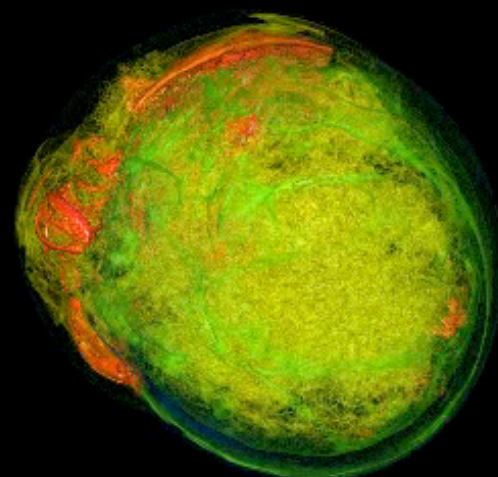
Peterka, T., Ross, R. B., Shen, H.-W., Ma, K.-L., Kendall, W., Yu, H.: Parallel Visualization on Leadership Computing Resources. Journal of Physics: Conference Series SciDAC 2009, June 2009.

Peterka, T., Ross, R., Yu, H., Ma, K.-L.: Assessing Improvements to the Parallel Volume Rendering Pipeline at Large Scale. SC08 Ultrascale Visualization Workshop, Austin TX, November 2008.

Ross, R. B., Peterka, T., Shen, H.-W., Hong, Y., Ma, K.-L., Yu, H., Moreland, K.: Parallel I/O and Visualization at Extreme Scale. Journal of Physics: Conference Series SciDAC 2008, July 2008.

Peterka, T., Yu, H., Ross, R., Ma, K.-L.: Parallel Volume Rendering on the IBM Blue Gene/P. Proceedings of Eurographics Symposium on Parallel Graphics and Visualization 2008 (EGPGV'08) Crete, Greece, April 2008.

Factors in Large Scale Parallel Visualization



Acknowledgments:

Rob Ross, Kwan-Liu Ma, Hongfeng Yu, Rob Latham, Wes Kendall, Han-Wei Shen

John Blondin, Tony Mezzacappa, Paul Fischer, Aleks Obabko

Argonne and Oak Ridge Leadership Computing Facilities

Tom Peterka

tpeterka@mcs.anl.gov

Mathematics and Computer Science Division

Argonne NATIONAL LABORATORY

... for a brighter future

UltraVis
SciDAC Institute for Ultrascale Visualization
www.ultravis.org

U.S. Department of Energy

UChicago Argonne LLC

Office of Science
U.S. DEPARTMENT OF ENERGY

A U.S. Department of Energy laboratory managed by UChicago Argonne, LLC