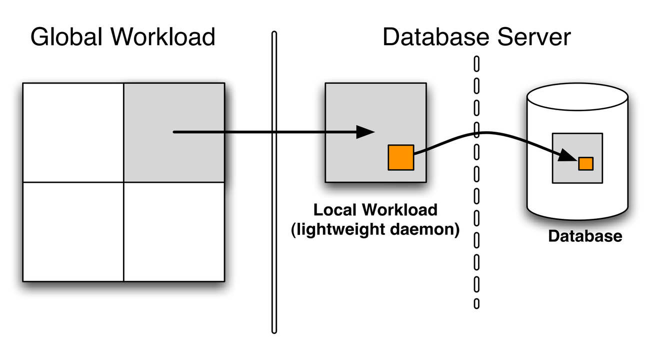
Engineering the 100 Terabyte Turbulence Database

or How to Track Particles at Home

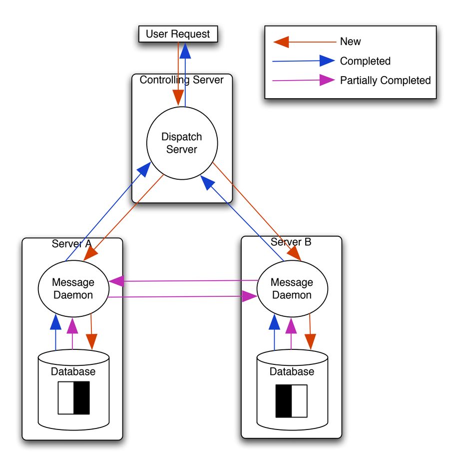
Eric Perlman and **Randal Burns**
Johns Hopkins University
{eric,randal}@cs.jhu.edu
http://hssl.cs.jhu.edu/

Workflow

- Minimize data transfer
  - Execution within the database
- Eliminate overhead
  - Batch schedule requests
  - Multiple time-step propagation in one request
  - DB to DB transfer for completion
- Increase locality
  - Spatial-order execution of points within DB
  - Increases cache performance



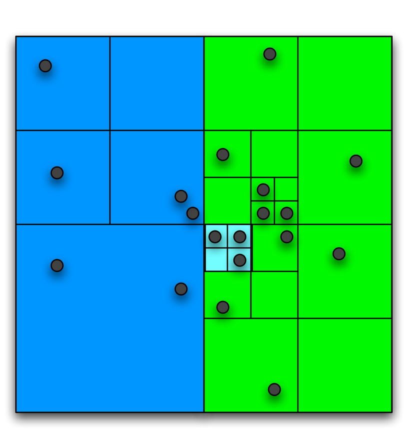
Breakdown of where computation occurs



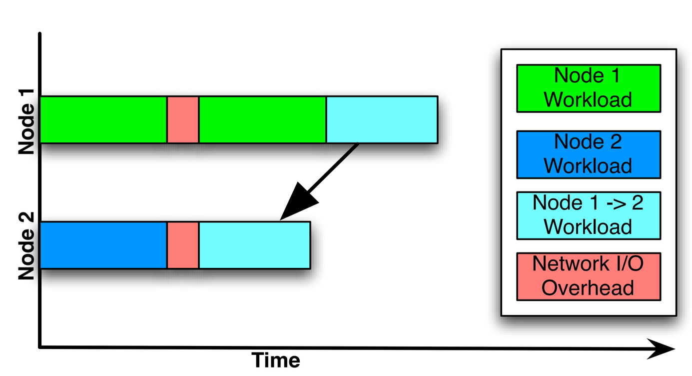
System messaging

Heterogeneous Spatial Workload

- Highly-skew spatial distributions
  - Associated with turbulent structures
- Needs dynamic load balancing
  - Must balance cost of both I/O and workload
- Region quad-tree on workload
  - Identify highest workload/data ratio



Quad-tree partitioning of workload



Execution of workload including data movement costs

The Turbulence Database

We created an environment for large-scale turbulence simulations that uses a cluster of database nodes to store the complete space-time history of fluid velocities. This allows for rapid access to high resolution data that were traditionally too large to store and too computationally expensive to produce on demand. We perform the actual experimental analysis inside the database nodes, which allows for data-intensive computations across a large number of nodes with little network traffic.

We currently have a limited-scale prototype system running turbulence simulations and are in the process of establishing a production cluster with high-resolution data. We allow jobs to be submitted through a Web service for computation on the cluster enabling one to *track particles at home*.

Evolving Direct Numerical Simulation (**DNS**) Studies

Current Methods

- Traditional cluster-based simulation
- Only a few time steps are stored
- Designed to be sequential, limited flexibility for exploration

Space-Time Histories

- Simple computations over large quantity of data
- Temporal Search
- Creates some restrictions on techniques
  - DB model is poorly suited for all-space FFTs

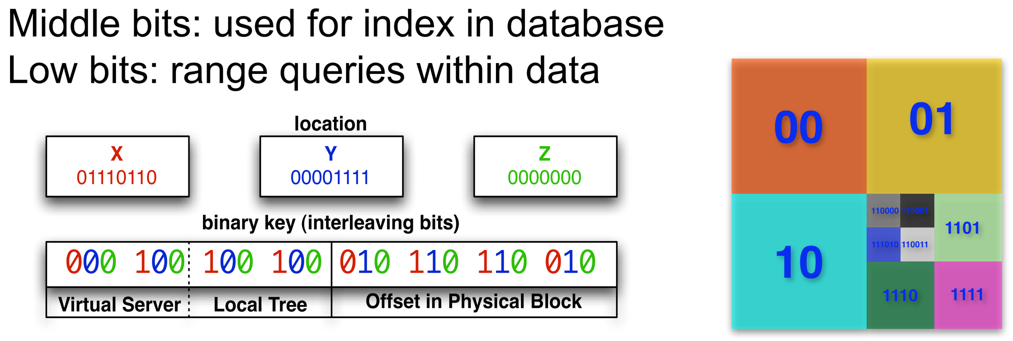
Data Organization

Dataset

- Regular and discrete physical-space mesh
  - Velocity (V_x,V_y,V_z) and pressure for all space-time coordinates
- Current database
  - 64^3 × 1000 steps and 512^3 × 100 steps
  - Segmented into 32- or 64-sided 3-d cubes
  - Index by {time,location}
- Target 100TB database
  - 1024^3 × 1000 generated at JHU
  - 8192^3 × 1000 generated at LANL

Hierarchical Indexing

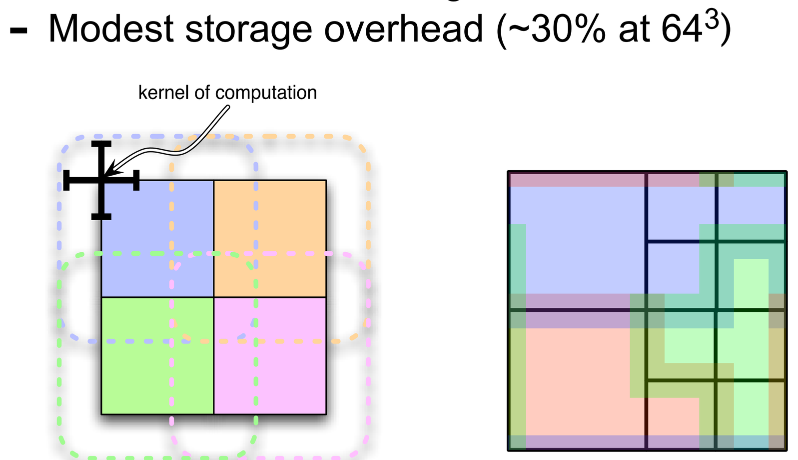
- Bit-interleaved addressing over entire space
- Segmented addressing
  - High bits: server selection
  - Middle bits: used for index in database
  - Low bits: range queries within data



Z-curve indexing (interleaving bits)

Overlapping Edge

- Velocities are interpolated over a kernel
  - currently 6^3 for particle tracking
- Block size is $(n+3)^3$ for each n^3
  - Every point within a cube can be interpolated from local data, avoiding DB to DB interactions
  - Modest storage overhead (~30% at 64^3)

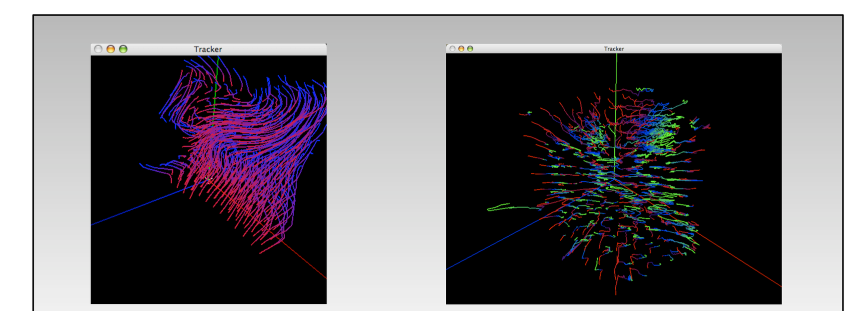


2-d overlapping regions to keep the kernel of computation local

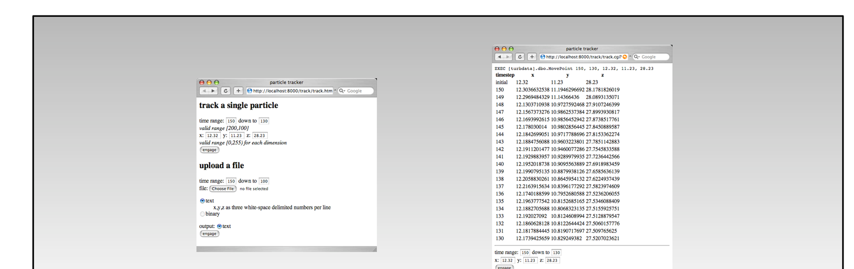
Data Exploration

Particle Tracking

- Study the evolution of turbulent structure
  - e.g. vortex tubes and current sheets
- Database allows efficient study of pre-history
- Enables data exploration
  - Iterate back and forth in time
  - Search for structures
  - Dynamic search criteria



Client visualization utility written in Python (PyOpenGl) using Web services to request particle tracking

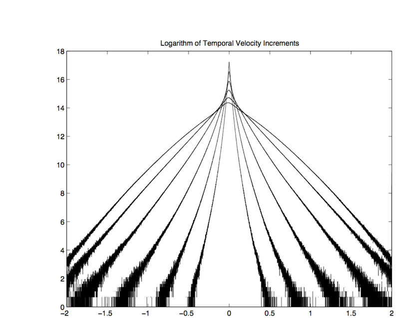


Basic web interface for tracking particles

Spatio-Temporal Analysis

$$\delta_{\ell,\tau} u(x,t) = u(x+\ell, t+\tau) - u(x,t)$$

- "Exotic" diffusion
  - Statistics on velocity field over arbitrary temporal and spatial ranges
  - Implications for skew-ness and non-Gaussian effects
- Requires the complete space-time history for large values of *t*



Temporal Velocity Increment PDFs for τ = 1,2,4,8,16,32

JOHNS HOPKINS ENGINEERING

Funded by NSF AST-0428325