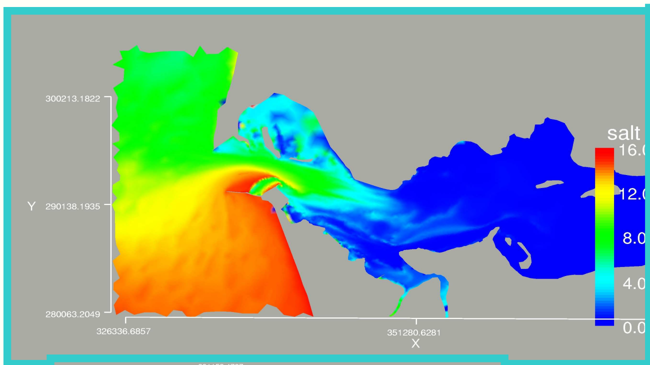
GridFields: Querying and visualizing Gridded Datasets for e-Science

PORTLAND STATE UNIVERSITY

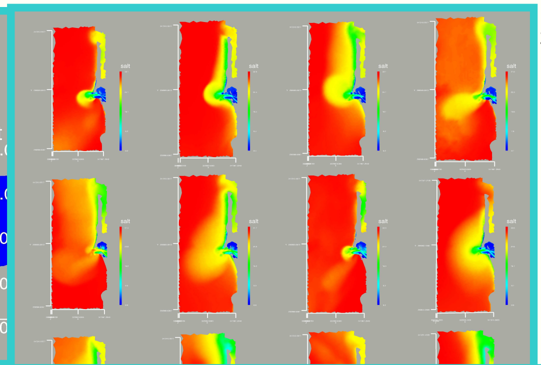
Bill Howe
howe@cs.pdx.edu

David Maier
maier@cs.pdx.edu

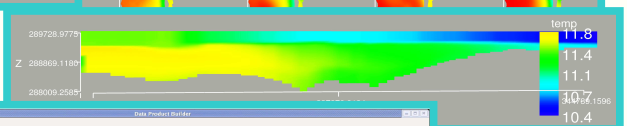
http://datalab.cs.pdx.edu/cormorant/



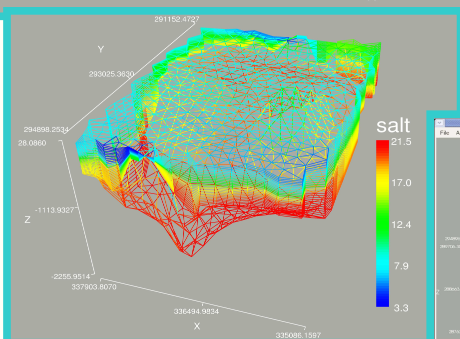
1) Average salinity over depth in the estuary of the Columbia River, Oregon, USA.



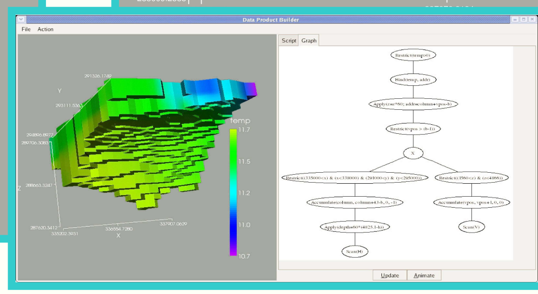
2) The plume of fresh water extending into the ocean from the mouth of the Columbia River, Oregon USA. Each panel represents a different month of 2004. In the summer, the plume swings southward and drifts out to sea. In the winter, the plume swings northward and hugs the shore.



3) A vertical *transect* of temperature along a a deep channel in the Columbia River, Oregon, USA.



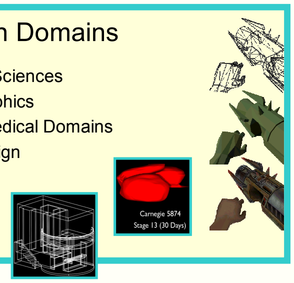
5) Salinity data bound to a 3-dimensional grid.



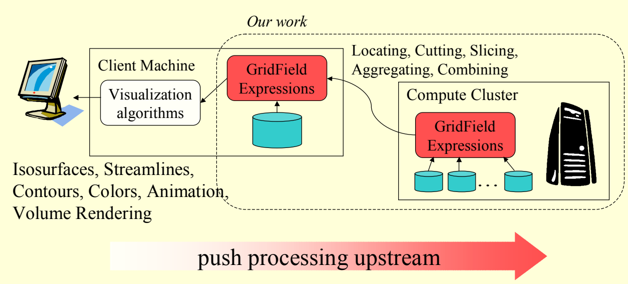
4) A screenshot showing a visualization and corresponding query plan

Motivating Application Domains

- Simulation in the Physical Sciences
- Model Libraries for 3D Graphics
- Numerical Studies in Biomedical Domains
- Architecture, Industrial Design



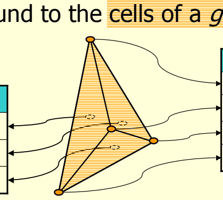
Theme: Separate Data Manipulation from Data Visualization



GridFields: Modeling Scientific Datasets

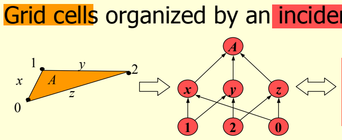
Attributes bound to the cells of a *grid*

flux	area
11.5	3.3
13.9	5.5
13.1	4.5



x	y	salt	temp
13.4	9.0	29.4	12.1
13.9	9.0	29.8	12.5
14.2	9.0	28.0	12.0
13.4	9.4	30.1	13.2

Grid cells organized by an incidence relationship



{0,1,2,x,y,z,A},

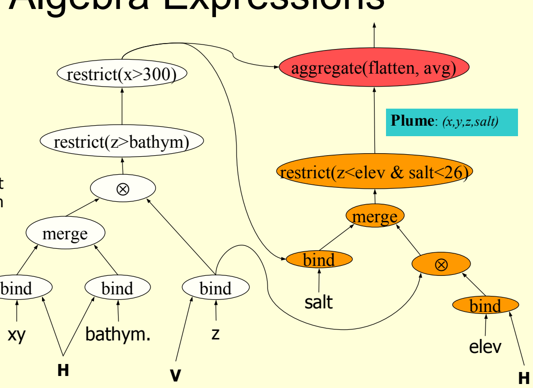
{(0,0), (0,x), (0,z), (0,A), (1,1), (1,x), (1,y), (1,A), (2,2), (2,y), (2,z), (2,A), (x,x), (x,A), (y,y), (y,A), (z,z), (z,A), (A,A)}

where (*x*,*y*) ≡ "*x* is incident to *y*"

GridField Algebra Expressions

Plan to compute the plume of fresh water that juts out into the ocean (Figure 3, above right), then aggregate over depth.

The left-hand side is not time-dependent and can be computed offline.

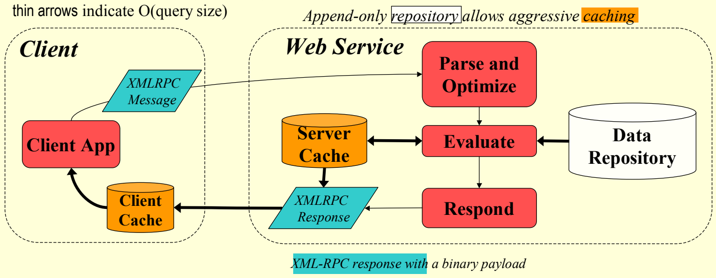


GridFields: Querying and visualizing Gridded Datasets for e-Science

http://datalab.cs.pdx.edu/cormorant/

Architecture

Thick arrows indicate O(result size); thin arrows indicate O(query size)



Visualization

Grids and bound Attributes

X : (x_coord)
Y : (y_coord)
Z : (z_coord)
T : (timestep)
X x Y : (h, b)
T x X x Y : (elevation, radiation, wind)
T x X x Y x Z : (salt, temp, u, v, w)

timestep < 20

Visualization Primitives

screen x	x_coord
screen y	y_coord
screen z	
Animate	timestep
Color	max(salt)
Glyphs	u*10, v*10

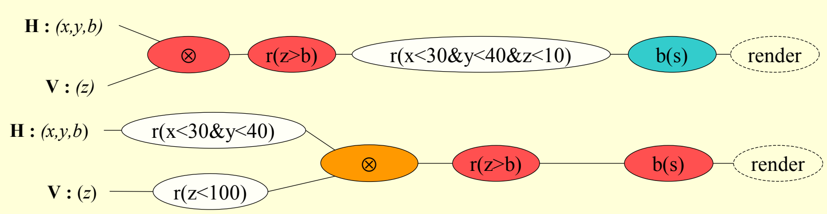
Map gridfield attributes onto visualization primitives: 3 spatial dimensions, color, etc.

Aggregations, restrictions and computations can be expressed.

Optimization

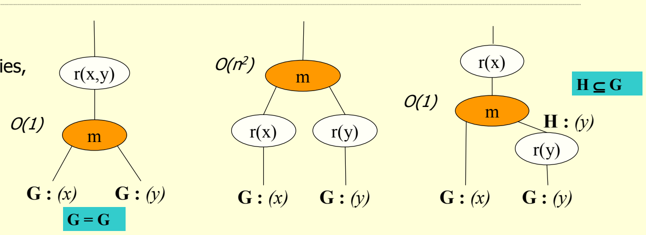
Push restrict operator through bind (converting it to a form of semi-join).

We can push the restrict operator through the cross product in a familiar manner.



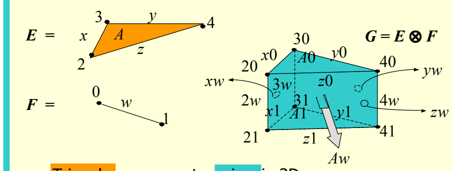
Pushing restricts can destroy grid equivalencies, leading to costly plans.

The merge operator is O(1) if one grid is a subgrid of the other.



Operators

Cross Product



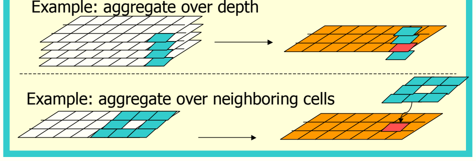
Aggregate

1) For each cell in the target grid, assign cells from the source grid using an *assignment function.*

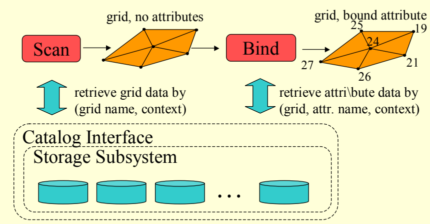
2) Apply a typical *aggregation function.*

Example: aggregate over depth

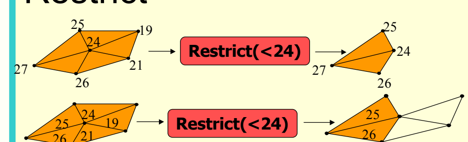
Example: aggregate over neighboring cells



Scan/Bind

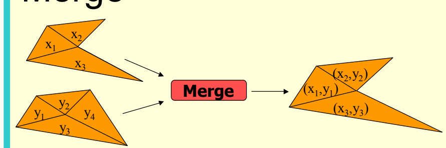


Restrict



Preserves well-formedness property: cells referenced in the incidence relationship must exist!

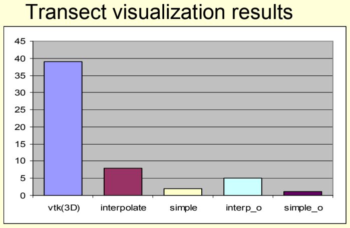
Merge



Topological intersection of two grids.

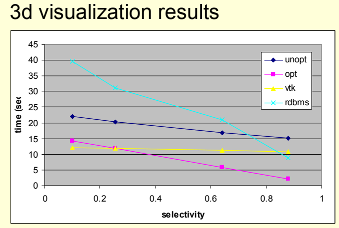
Performance

Transect visualization results



Results for a vertical transect visualization (Figure 3). Exposing algebraic equivalences (bars marked interpolate, simple, interp_o, and simple_o) allows significant savings over over direct computation using a visualization library (bar labeled vtk3D). See Howe, Maier VLDB 2004 for details.

3d visualization results



Our optimized plan (opt) for computing a 3D visualization (similar to Figure 5) offers better performance than both a visualization library (vtk), a relational database implementation (rdbms), and an unoptimized plan (unopt). See Howe, Maier 2004 for details.

Bill Howe, Portland State University, bill@cs.pdx.edu