VisTrails: Enabling Interactive, Multiple-View Visualizations

Louis Bavoil Steven P. Callahan Juliana Freire Claudio T. Silva Carlos Scheidegger Huy T. Vo University of Utah Patricia Crossno Sandia



Motivation

- Develop a system to enable the creation and maintenance of a large number of complex visualizations
 - Ability to create complex pipelines
 - Ability to track changes to visualizations
 - Need to enable comparisons:
 - Across multiple versions of the same visualization
 - Same visualization on different sets of data
 - Efficient, easy to use, portable, and simple to implement

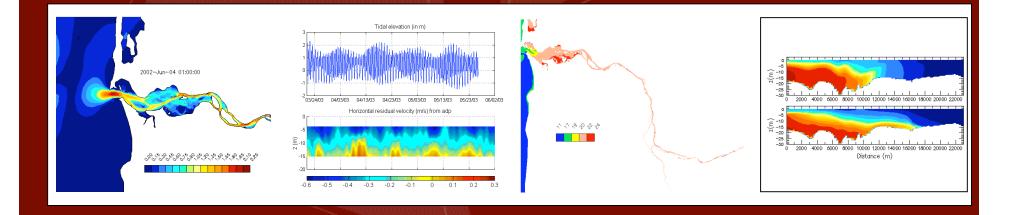
Motivating Example: CORIE

CORIE is an Environmental Observation and Forecasting Systems (EOFS) that combines real-time sensor measurements with advanced computer models to increase reliability of complex, dynamic environmental systems

Thousands of visualizations daily

No management infrastructure

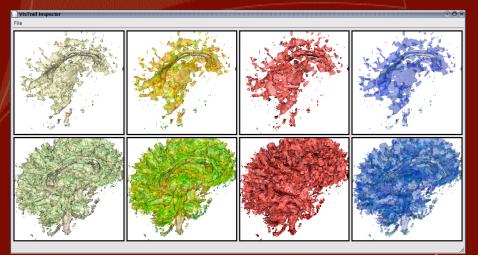
http://www.ccalmr.ogi.edu/CORIE/





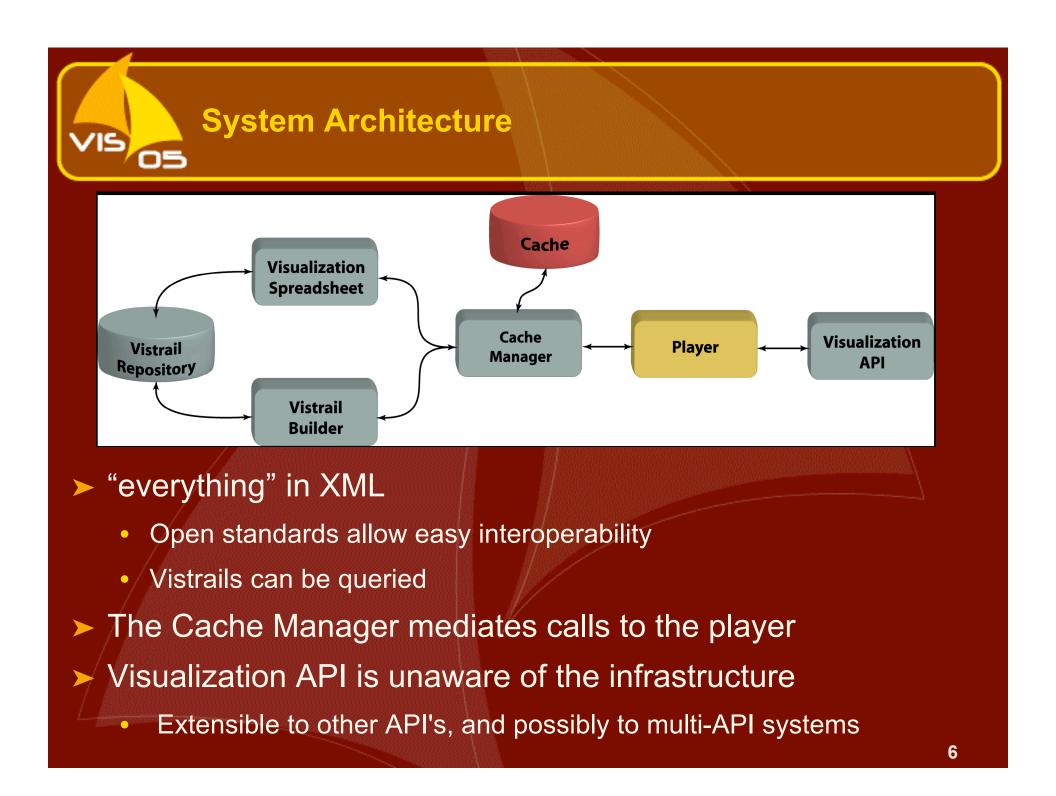
A new system that enables interactive, multiple-view visualizations

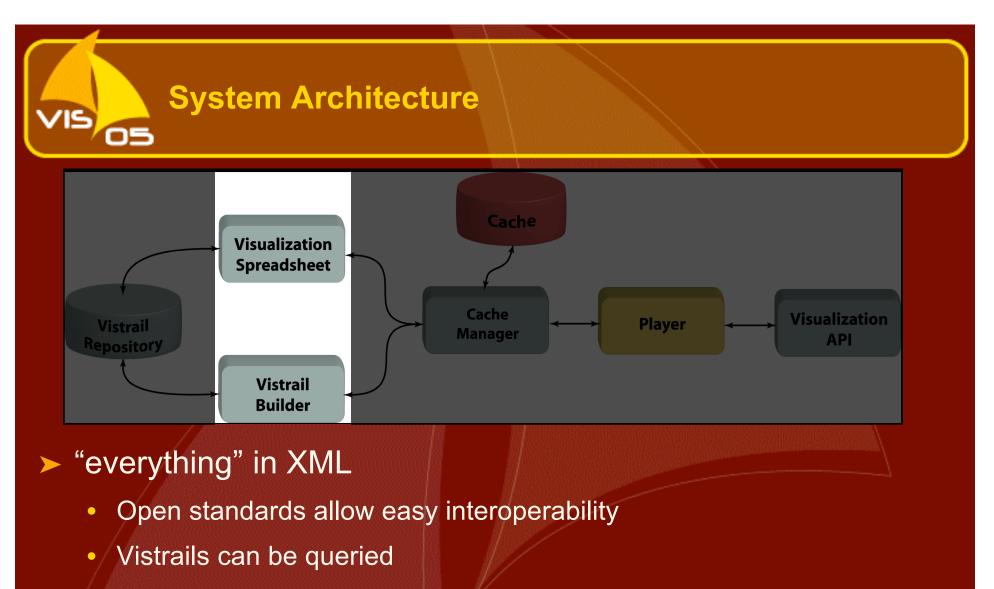
- Simplify the creation and maintenance of a large number of visualizations
- Detailed provenance of visualization results
- Separation between pipeline specification and execution instances
- Optimization of execution through caching



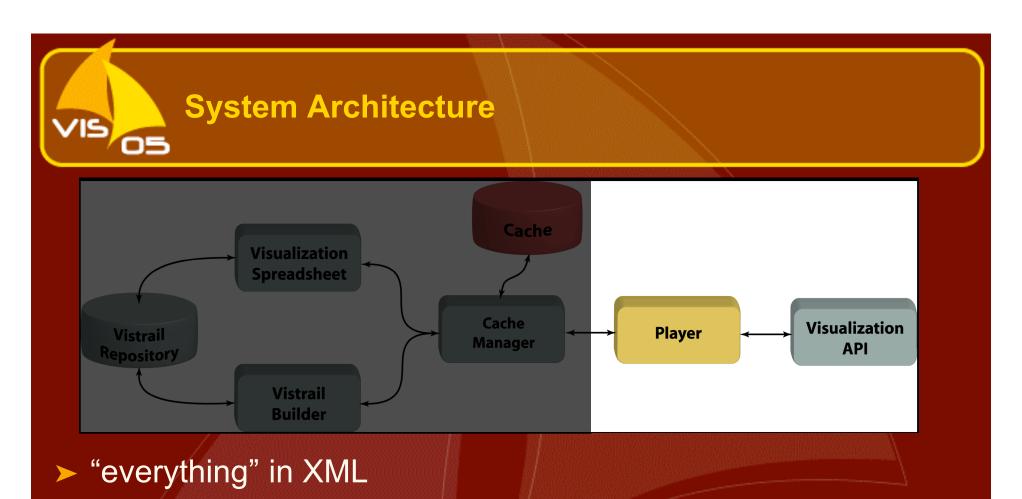
Previous Work

- Many previous dataflow visualization systems
 - Paraview, Opendx, SCIRun, IRIS Explorer, many others
- Kreuseler et al.: tree history for exploratory data mining
 - Readily applicable to exploratory visualization
 - Tree provides provenance of the process
- Brodlie et al.: extension of IRIS Explorer over Grid resources
- Jankun-Kelly and Ma use a spreadsheet-like interface to aid exploration
- Jankun-Kelly et al provide a model for the visualization exploration process
 - See paper for details

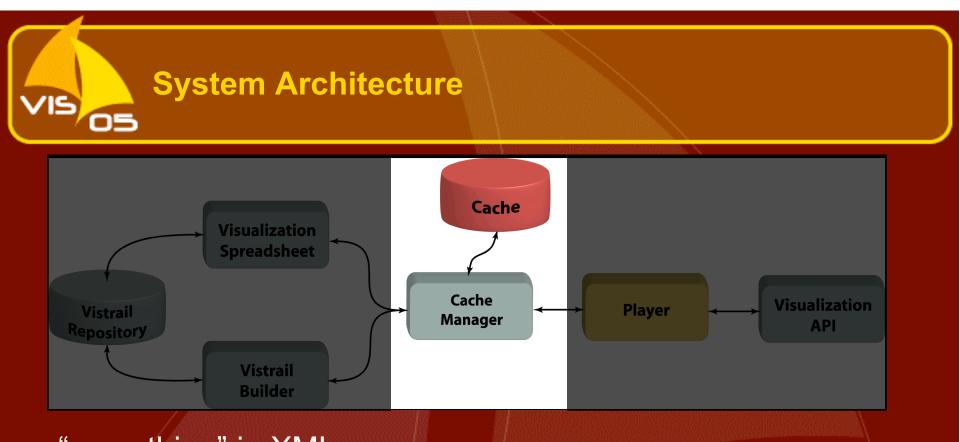




- The Cache Manager mediates calls to the player
- Visualization API is unaware of the infrastructure
 - Extensible to other API's, and possibly to multi-API systems

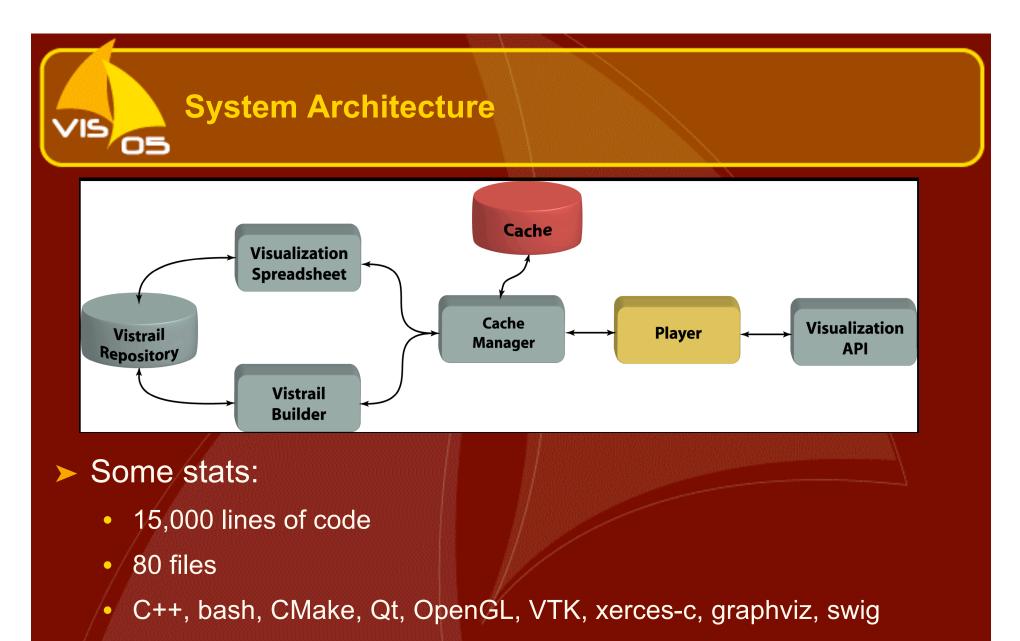


- Open standards allow easy interoperability
- Vistrails can be queried
- The Cache Manager mediates calls to the player
- Visualization API is unaware of the infrastructure
 - Extensible to other API's, and possibly to multi-API systems



"everything" in XML

- Open standards allow easy interoperability
- Vistrails can be queried
- The Cache Manager mediates calls to the player
- Visualization API is unaware of the infrastructure
 - Extensible to other API's, and possibly to multi-API systems



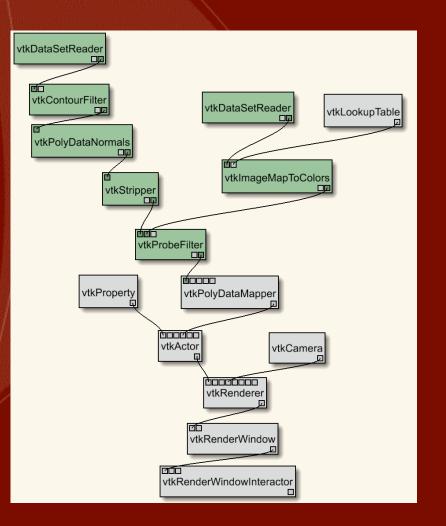
• Windows, Linux, Mac OS X

Stay tuned for updates: http://www.sci.utah.edu/~vgc

The Vistrail model

- Vistrail: sequence of operations used to generate a visualization
- Parameter settings are distinguished from the dataflow

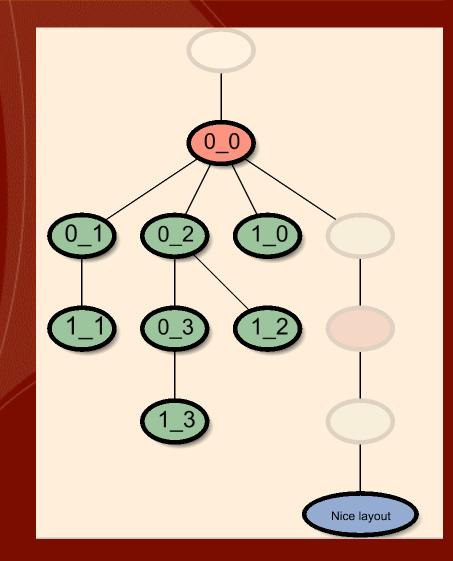
Modules represent filters in the dataflow network, and connections determine the dependencies



Visualization Provenance

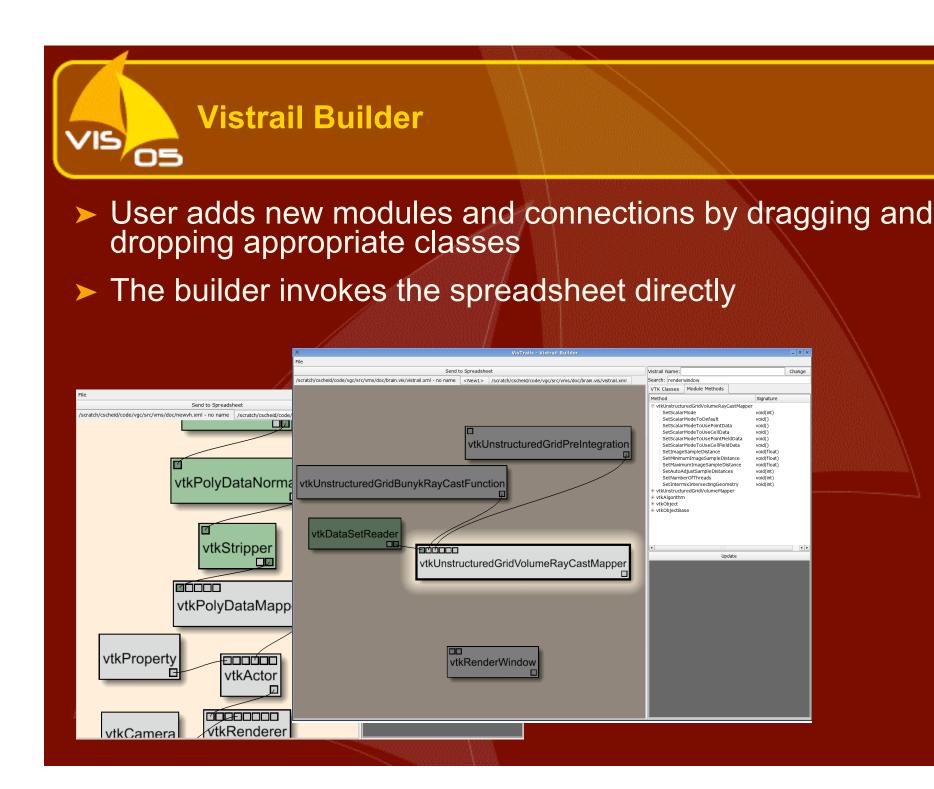
 VisTrails stores visualizations in a Vistrail Collection

- version tree given from metadata
- Exporting a visualization result ensures reproducibility





- User adds new modules and connections by dragging and dropping appropriate classes
- > The builder invokes the spreadsheet directly



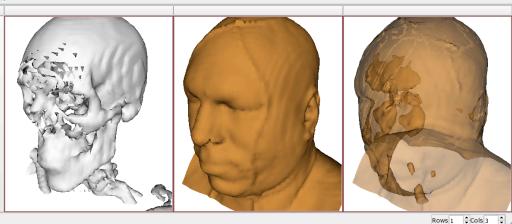


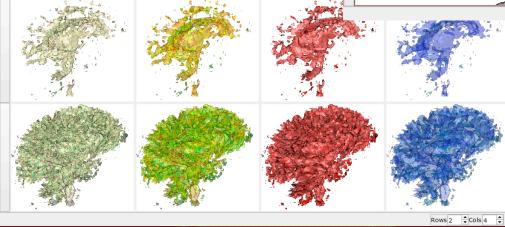
- User adds new modules and connections by dragging and dropping appropriate classes
- > The builder invokes the spreadsheet directly

×	VisTrails - Vistrail Builder	_ 6 ×
Fie	I to Spreadsheet Vistrail Name: d	hange
/scratch/cscheid/code/vgc/src/vms/doc/brain.vis/vistrail.xml - no name	<new1> /scratch/cscheid/code/vgc/src/vms/doc/brain.vis/vistrail.xml Search: renderwindow</new1>	
Fie	UTX Flacese Module Methode	d
Send to Spreadsheet /scratch/cscheid/code/vgc/src/vms/doc/newvh.xml - no name /scratch/cscheid/code/	Vistrail Name: skin	Change
	Search: dataset	
	VTK Classes Module Methods	
ビ vtkPolyDataNorma vtkUnstructuredGridBunykRayCa	Name	Туре 🔺
	😑 vtkObjectBase	Abstract
	⊖ vtkObject	Object
VtkDataSetReader	vtkAlgorithm	Object
vtkStripper	vtkAbstractMapper	Abstract
vtkUnst	n 😑 vtkAbstractMapper3D	Abstract
	😑 vtkMapper	Abstract
vtkPolyDataMapp	vtkDataSetMapper	Object
	vtkDataObjectAlgorithm	Filter
	vtkDataSetToDataObjectFilter	Filter
vtkProperty	vtkDataReader	Filter
vtkActor	vtkDataSetReader	Filter
	vtkDataSetAlgorithm	Filter
	vtkDataObjectToDataSetFilter	Filter
vtkCamera	vtkInterpolateDataSetAttributes	Filter
	vtkGenericDataSetAlgorithm	Abstract
	vtkHierarchicalDataSetAlgorithm	Abstract



- User can compare a large number of visualizations in the spreadsheet
- Views can be linked





The Cache Manager	
Reader Isosurface PolyNormals Stripper PolyMapper Actor Renderer Interactor	
 The Cache Manager determine Each module is broken into a s Each subnetwork receives a u modules, connectivity and para 	series of subnetworks nique ID, comprising its

The Cache Manage	
Reader Isosurface PolyNormals PolyNormals PolyMapper PolyEarce PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper	
The Cache Manager deterr	nines pipeline sharing

- Each module is broken into a series of subnetworks
- Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters

The Cache Manage	r
Reader Reader Isosurface Isosurface PolyNormals PolyNormals Stripper Cache Lookup PolyMapper PolyMapper PolyMapper PolyMapper Menderer Renderer Interactor Interactor	
The Cache Manager deterion	mines pipeline sharing

- Each module is broken into a series of subnetworks
- Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters
 - Results are linked to the ID, and only computed if missing in the cache

The Cache Manager	
Reader Reader Isosurface PolyNormals PolyNormals PolyNormals PolyMapper PolyMapper	

- The Cache Manager determines pipeline sharing
- Each module is broken into a series of subnetworks
- Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters
 - Results are linked to the ID, and only computed if missing in the cache

The Cache Manager	
Reader Reader Isosurface PolyNormals Stripper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper Interactor	Cache Lookup Isosurface

- The Cache Manager determines pipeline sharing
- Each module is broken into a series of subnetworks
- Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters
 - Results are linked to the ID, and only computed if missing in the cache

The Cache Manager	
Reader Reader Isosurface PolyNormals PolyNormals Stripper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper PolyMapper Interactor	Cache Lookup Reader Jossurface Isosurface PolyNormals PolyMapper Actor Renderer Interactor
 The Cache Manager determi Each module is broken into a 	

Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters

The Cache Manager		
Reader Reader Isosurface PolyNormals PolyNormals Stripper PolyMapper PolyMapp	Cache Lookup Reader Isosurface	Reader Isosurface PolyNormals PolyMapper PolyMapper PolyMapper Actor Renderer Interactor Interactor

- > Each module is broken into a series of subnetworks
- Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters

	The Cache Manager		
Reader V Isosurface V PolyNormals V Stripper V PolyMapper V Actor V Renderer V Interactor	Reader Isosurface RolyNormals Cache Lookup TolyMapper Actor Renderer	Cache Lookup Reader Josurface	Reader Isosurface PolyNormals PolyMapper Actor Renderer Interactor Interactor
	ache Manager determi module is broken into a	1 States -	J

Each subnetwork receives a unique ID, comprising its modules, connectivity and parameters

Discussion and conclusions

- VisTrails is a system that allows interactive multiple-view visualizations
- Leverages formal specification of pipelines to increase efficiency
- Allows fast exploration of parameter space with the Visualization Spreadsheet
- Provides detailed provenance of visualization results

VisTrails: Demo

(Check out <u>http://www.sci.utah.edu/~vgc</u> for updates and code)

We'll be at the VTK BOF tonight



Future Work

- Changeset orientation really defines an algebra of pipelines
 - Checking commutativity: move actions around
- Parallelism

- Execution in a grid environment
- Graph layout of time-varying graphs
- Deployment
 - CORIE Vis'03 paper

Acknowledgments

Antonio Baptista

 NSF grants IIS-0513692, CCF-0401498, EIA-0323604, CNS-0541560, and OISE-0405402

► DOE VIEWS and MICS

Sandia National Laboratories is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under Contract DE-AC04-94AL85000

Emanuele Santos, John Schreiner, Wayne Tyler

> AT&T, Kitware, Trolltech, SWIG project

Bruno Notrosso, Gordon Kindlmann