Advanced Concepts for Large Data Visualization

Kwan-Liu Ma
Department of Computer Science
University Of California at Davis

SNU, February 28, 2012

Visualization & Interface Design Innovation

SciDAC Institute for Ultrascale Visualization

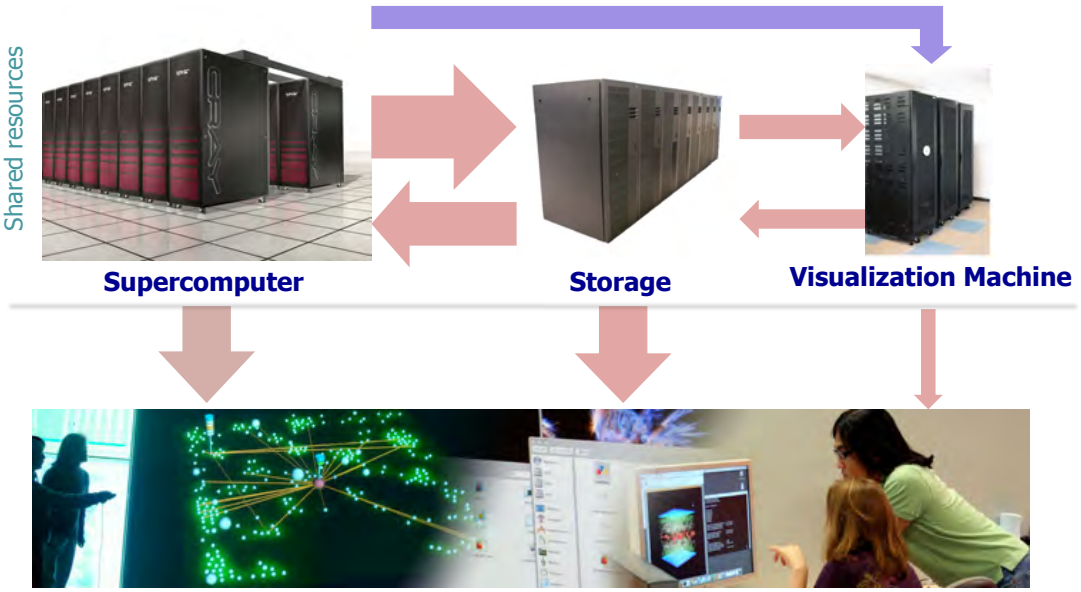
Research Interests

- Scientific Visualization
- Information Visualization
- Visual Analytics
- High Performance Computing
- User Interface Design
- Computer Graphics

Outline

- Large-Scale Scientific Data Visualization
  - In situ visualization
  - Multi-dimensional particle data visualization
- Visibility Directed Volume Visualization (Vis '09)
- Network Visualization and Analysis (InfoVis '08, IEEE TVCG)
- Software Visualization (InforVis '09)
- Dynamics Video Narratives (SIGGRAPH '10)
- Multi-scale Views of 3D Models (SIGGRAPH Asia '11)

Scientific Supercomputing

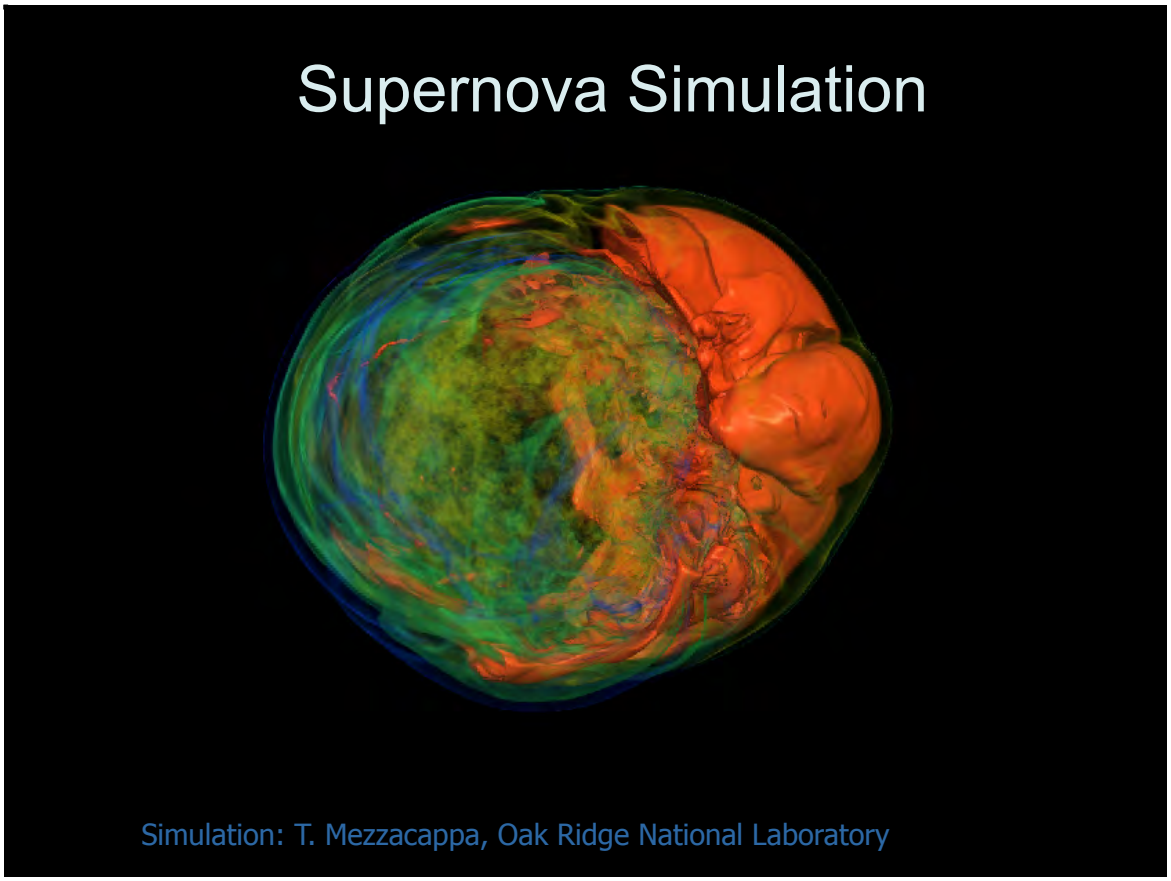


In Situ Visualization

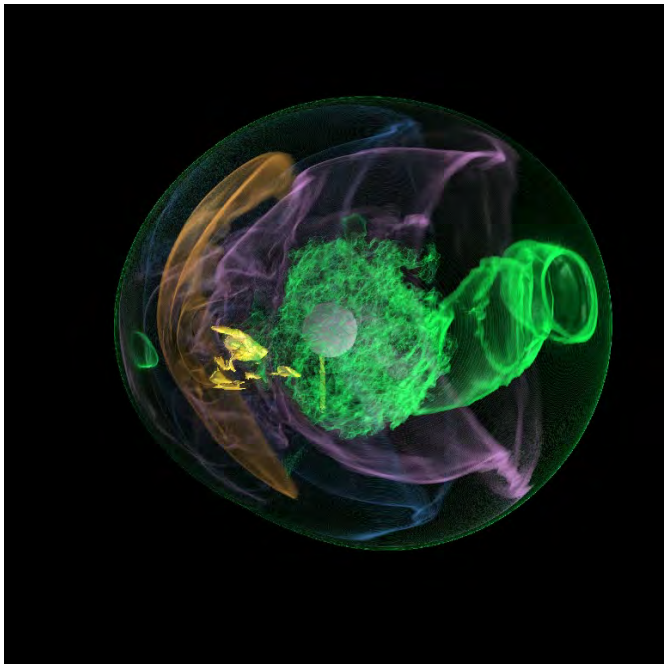
- Reduce/visualize data in situ as the simulation is running
- Process the data before it is written to the disk.
- Two technical approaches:
  - Co-located on a node (techniques that exploit data locality)
  - Concurrent processing (shipping data to dedicated vis/analysis nodes, possibly reducing first)
- In situ visualization enable:
  - Monitoring and validation
  - Data reduction and triage
  - Steering data reduction and analysis
  - Debugging and performance optimization

Requirements & Challenges

- Integration of simulation and visualization codes
- Low memory overhead
- Sharing the domain decomposition and data structures
- Low computational cost
- Scalable parallel visualization algorithms
- Not knowing what is interesting/important up front
- Additional requirements for interactive monitoring, steering, and different types of visualization



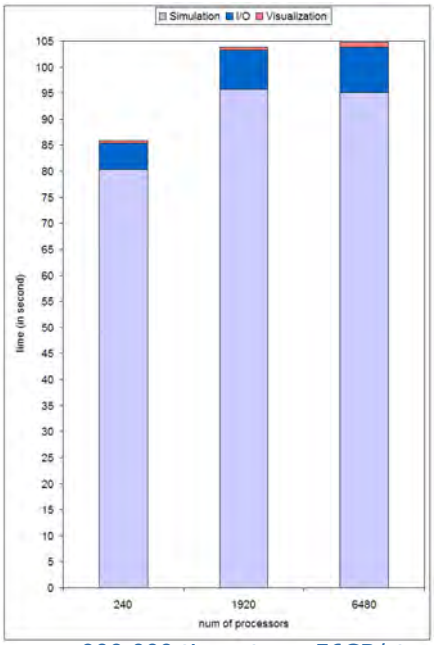
Supernova Simulation



Simulation: T. Mezzacappa, Oak Ridge National Laboratory

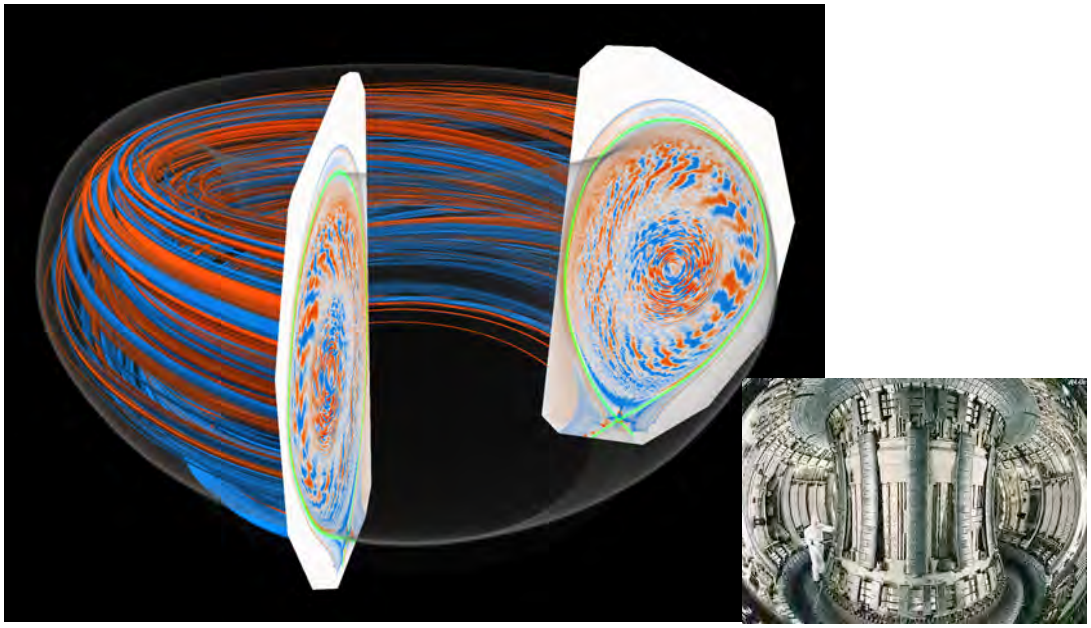
In Situ Visualization of Combustion Simulations

CH2O/HO2 CH3/OH



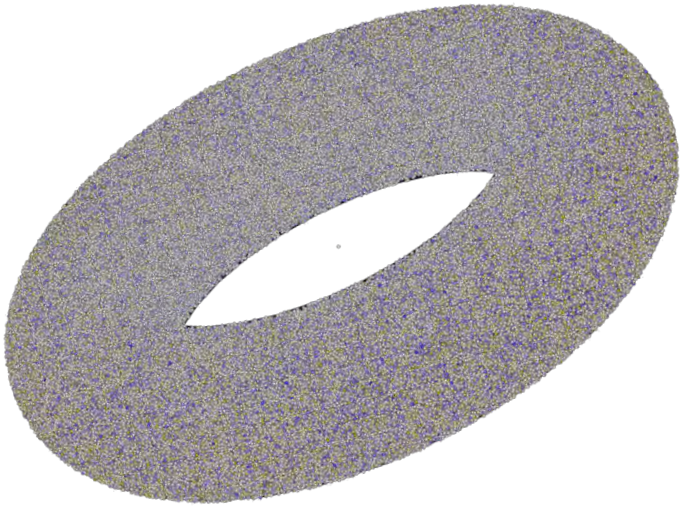
900,000 time steps, 56GB/step

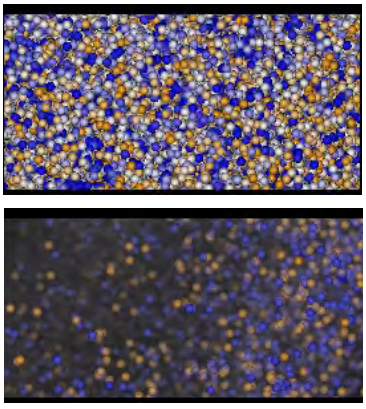
Fusion Simulations



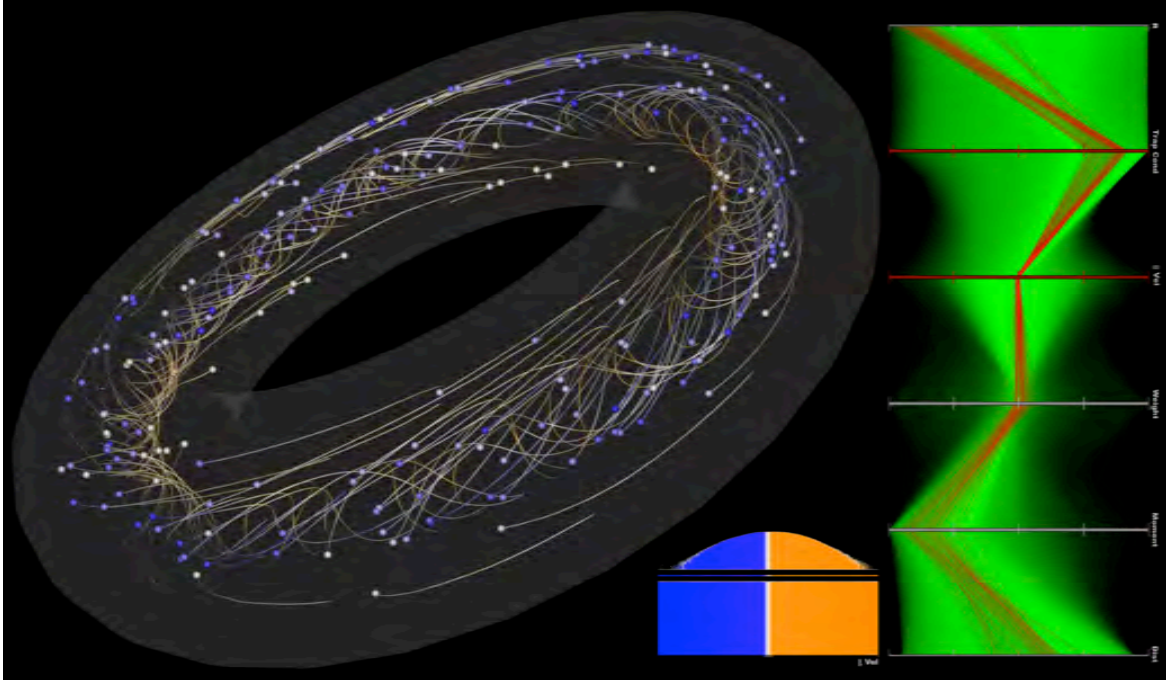
Simulation: Dr. S. Ethier, the Princeton Plasma Physics Laboratory

Multidimensional Particle Data



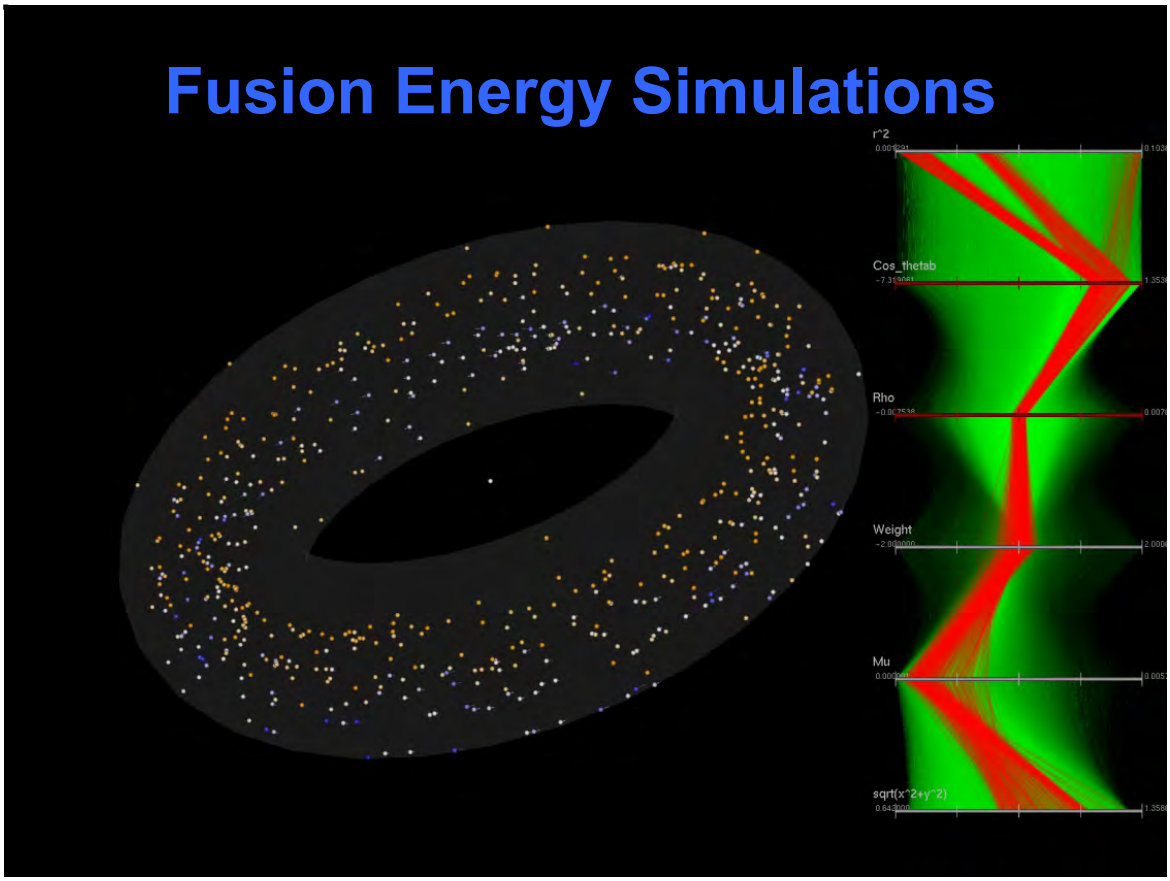


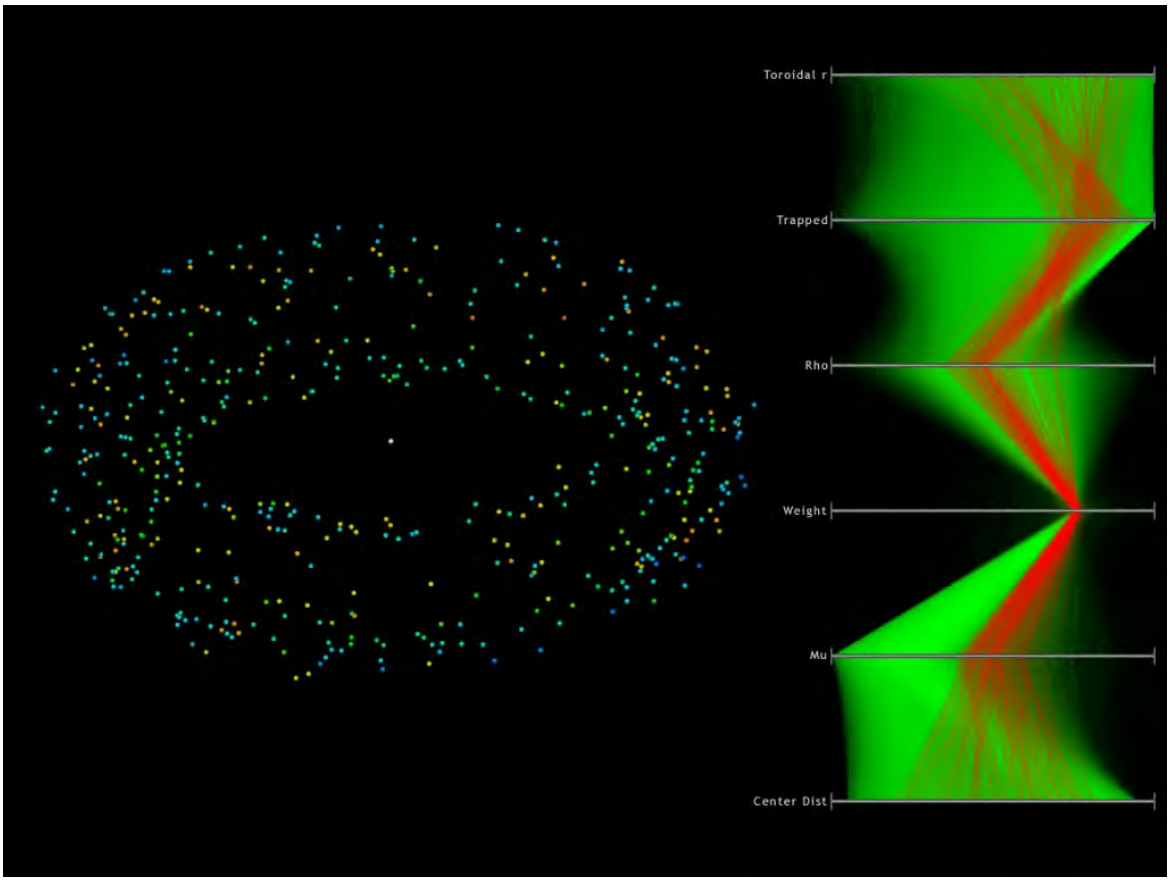
Multidimensional Particles Filtering



Trapped particles that change direction frequently

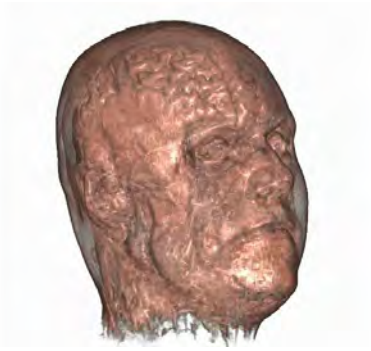
Fusion Energy Simulations





Volume Data Classification Challenges

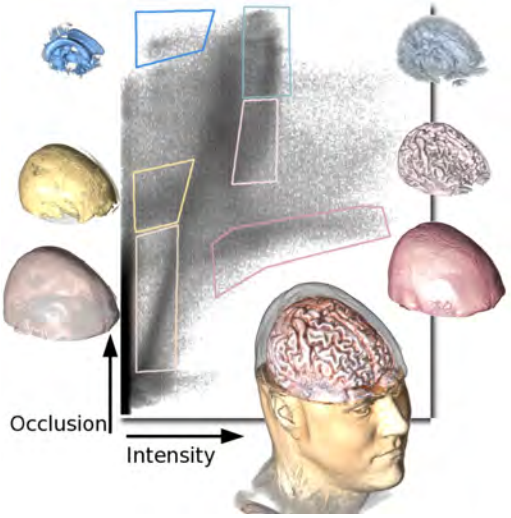
- Each image modality has its own particular strengths and limitations.
- A single scalar value cannot always uniquely define a feature of interest



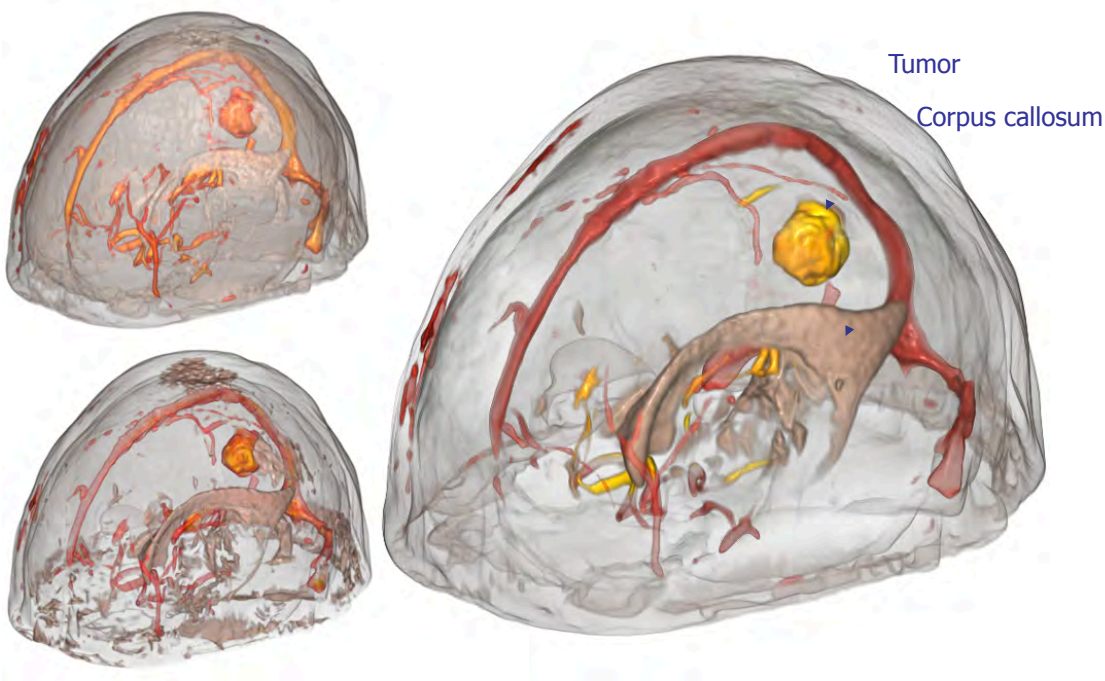
- Different features may have similar intensity values
- A feature may change its properties over time

Occlusion Spectrum

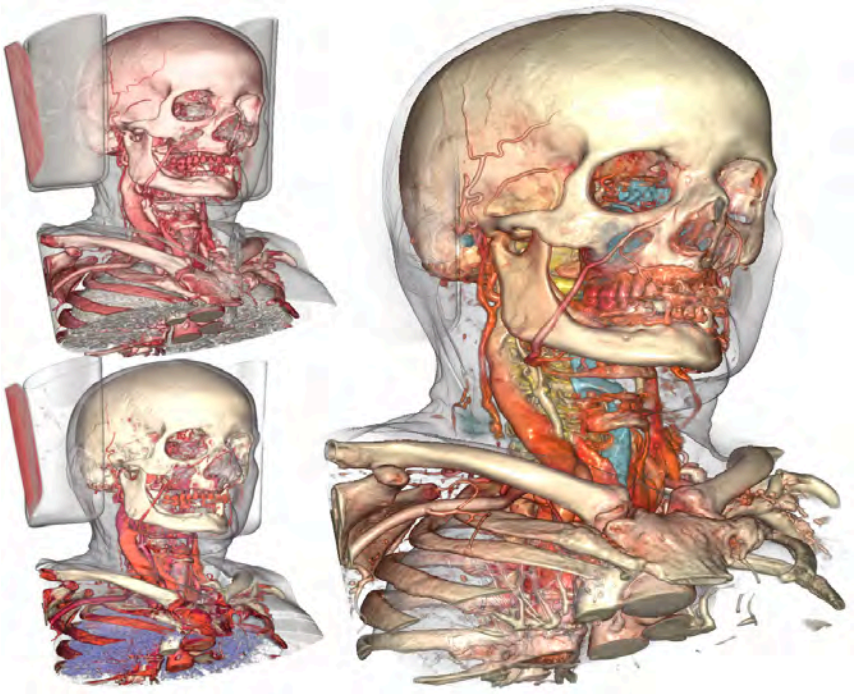
- Classifying based on the ambient occlusion of voxels
- Occlusion patterns often correspond to the spatial structures of features of interest in the data



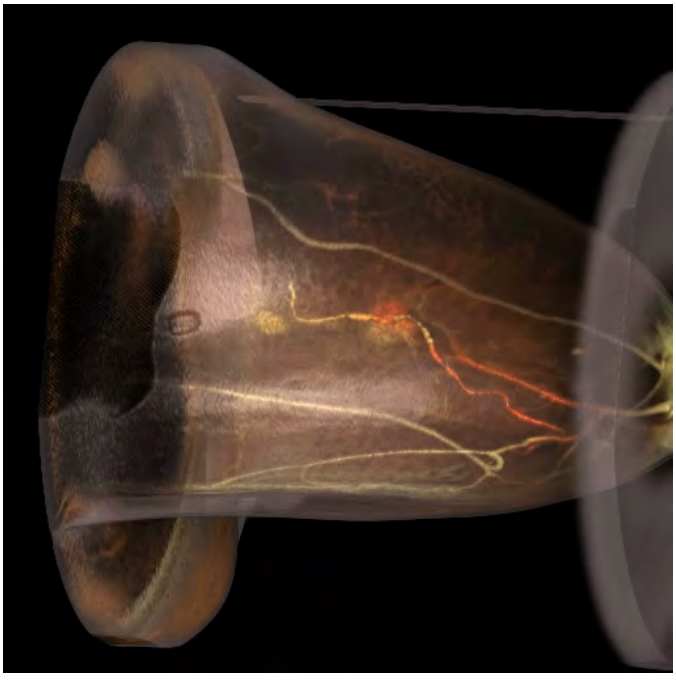
Occlusion Spectrum



Occlusion Spectrum



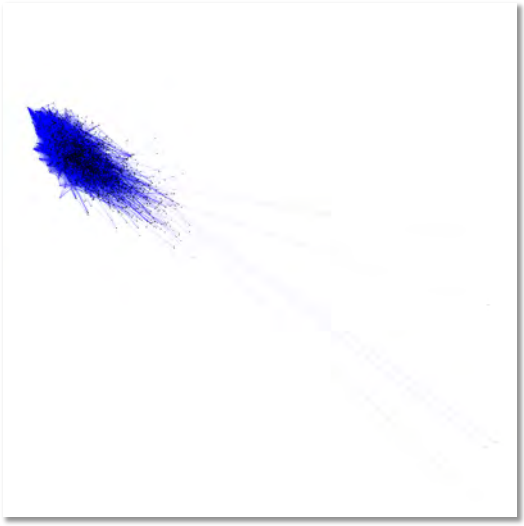
Occlusion Spectrum



Data provided by J. Boone, Medical School, UC Davis

The Graph Layout Problem

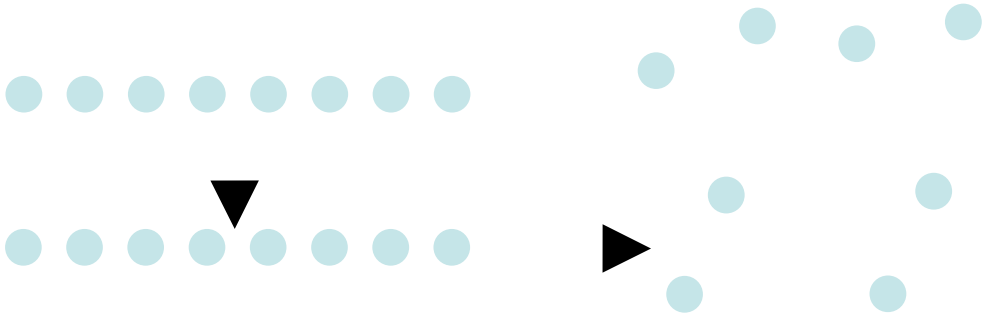
- The cost of displaying a graph
- The hairball problem of large graph layouts
  - Large, dense graphs become a mess
  - Inefficient use of space
  - Details cluttered
- Solutions
  - Filtering
  - Clustering
  - Abstraction
  - Focus+context

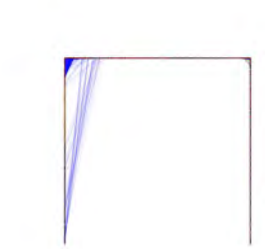


California data 6,107 nodes 15,160 edges
High dimensional embedding method

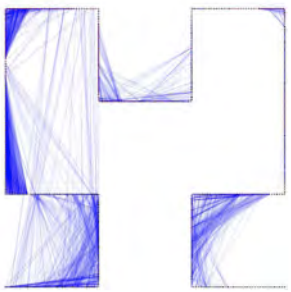
Space Filling Curve Based Layout

- Hierarchically cluster the nodes (if no clustering given)
- Traverse the hierarchy to order the nodes
- Place the nodes in that order along a space filling curve

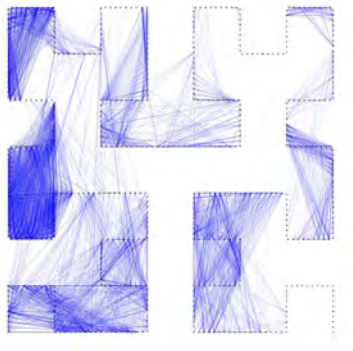




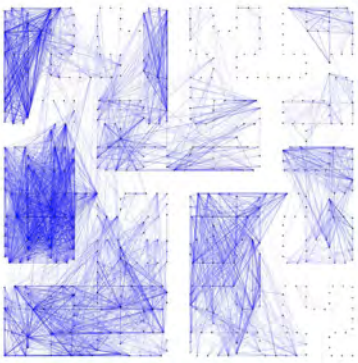
Order 1



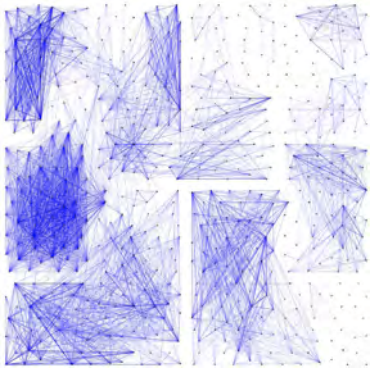
Order 2



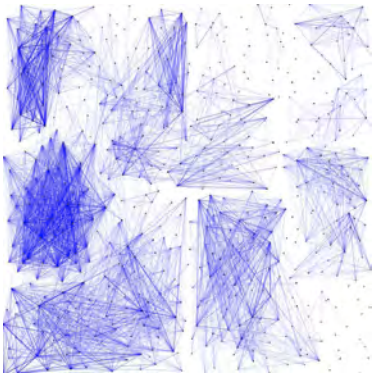
Order 3



Order 4



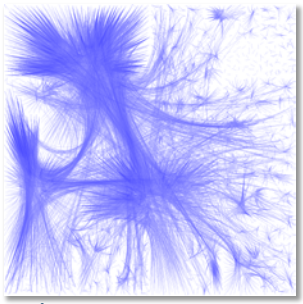
Order 5



Order 11

Fast Graph Layout

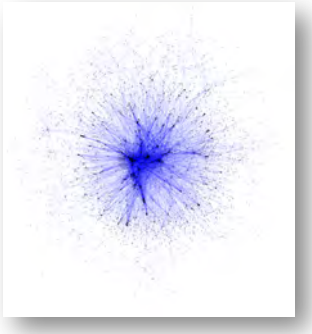
Our Methods: ~1s



Hibert curve

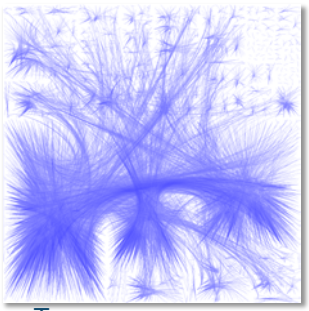
Gosper curve

Force-directed method: 10,737s

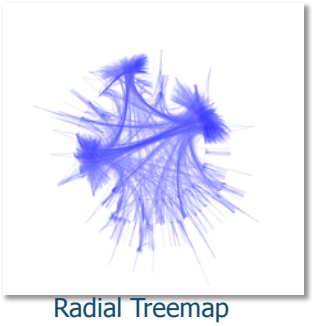


6,107 nodes 15,160 edges

Our Methods: ~1s

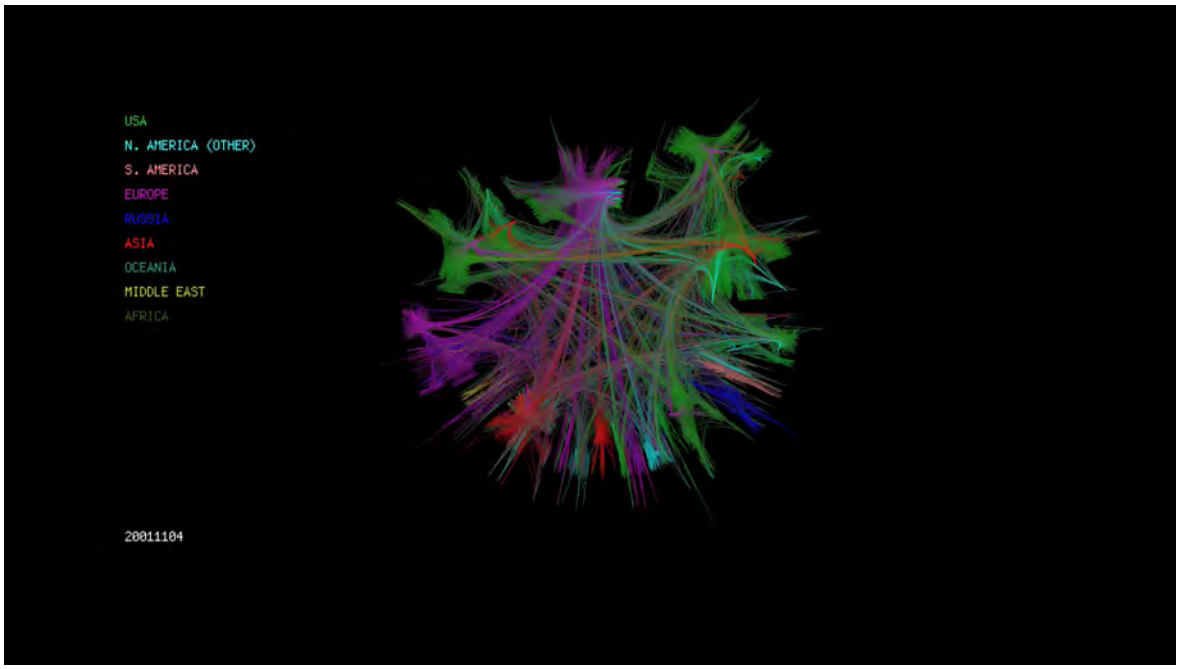


Treemap



Radial Treemap

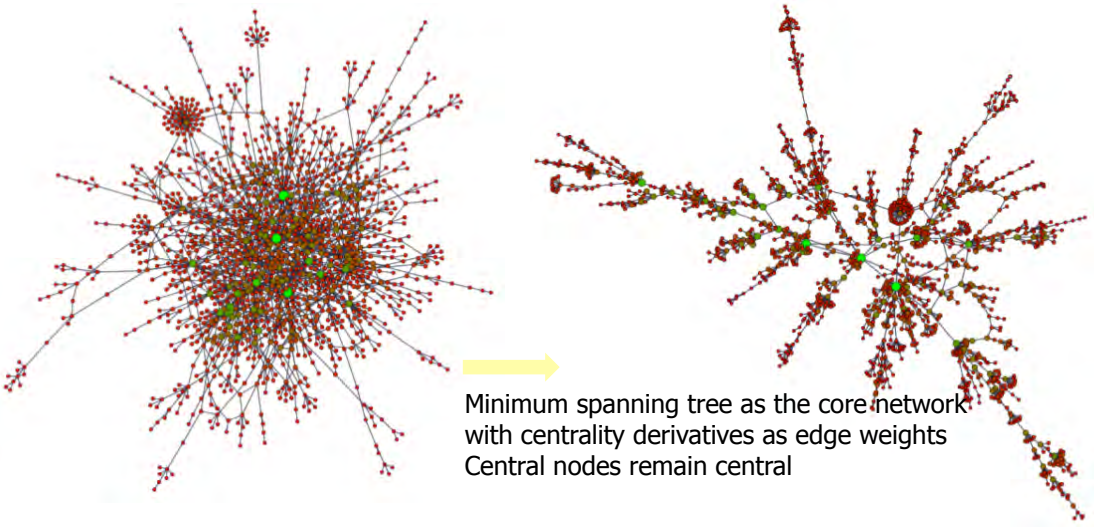
Visualizing Internet Connectivity



Centrality Sensitivity

- Centralities (degree, between-ness, closeness, eigenvector, Markov, …) indicate how *important* a node is in a network.
- Studying the ***sensitivity*** and *stability* of a network in terms of different metrics for centrality allow us to
  - Filter the network
  - Obtain an overview of the network
  - Search and explore in the network based on relative importance
- Compute sensitivity as the derivative of the centrality function, approximate derivatives of centrality using finite difference, and validate by computing the mean square error of the linear fit between the approximated and analytical values

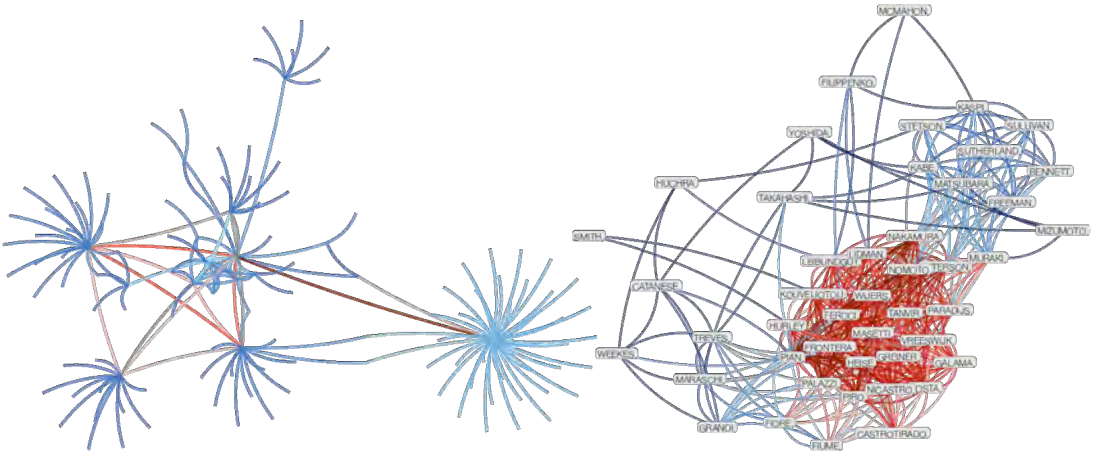
Simplification



Minimum spanning tree as the core network
with centrality derivatives as edge weights
Central nodes remain central

Network of protein-protein interaction (~1500 nodes)

Overview of Sensitivity

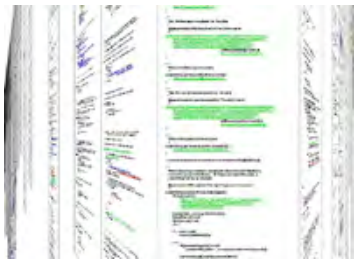


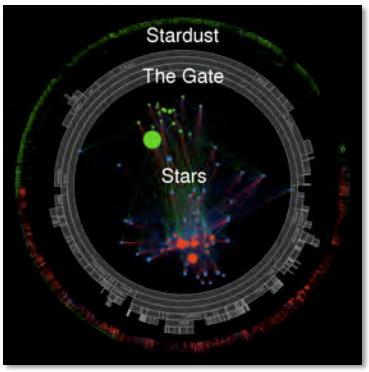
Friendster social network
Links exhibit negative sensitivity (red) between cluster centers

Astrophysics co-author network
One competitive network (red) and one collaborative network (blue)

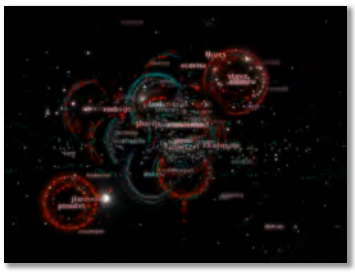
Software Visualization

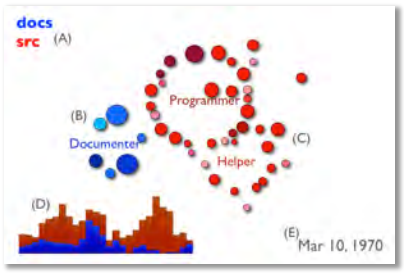
- Source code
- Monitoring program execution
- Debugging
- Analyzing runtime performance
- **Software evolution**
- **Developers** social network





Code_Swarm

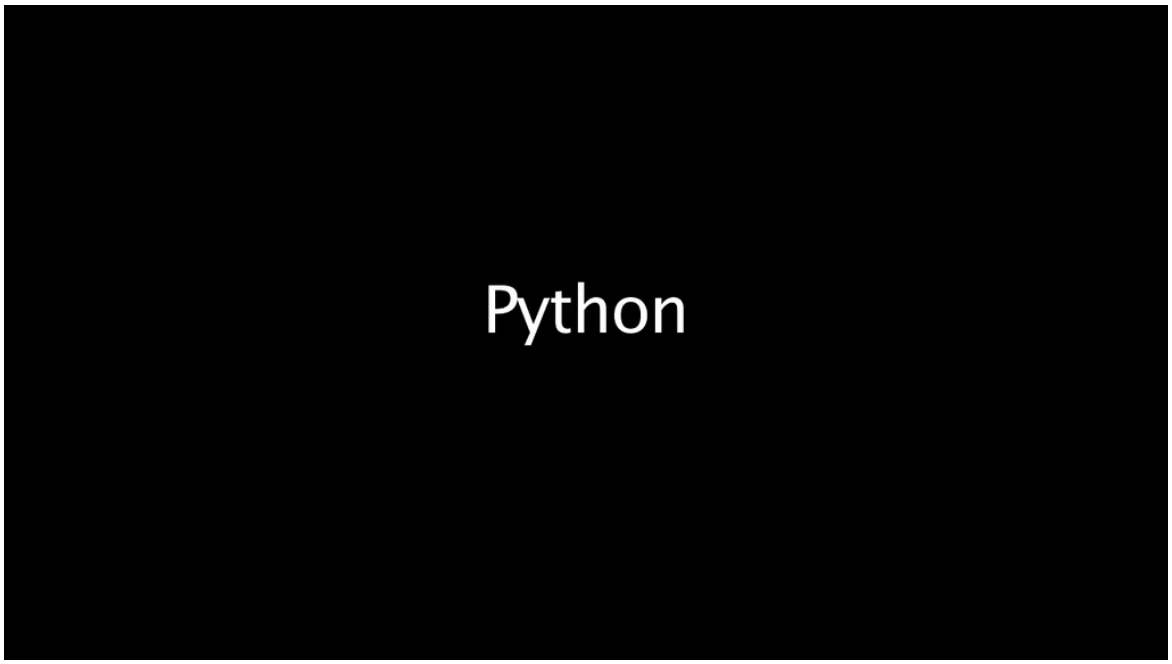




- Organic visualization to avoid a rigid layout
- An animation showing the history of commits in a project
- Both developers and files are shown as moving elements

- When a developer commits a file, the file lights up and flies towards that developer
- Each file is colored according to its purpose
- An inactive file/developer will fade away
- A histogram keeps a reminder of what has come before
- In this space the centrality of authors grasps attention!

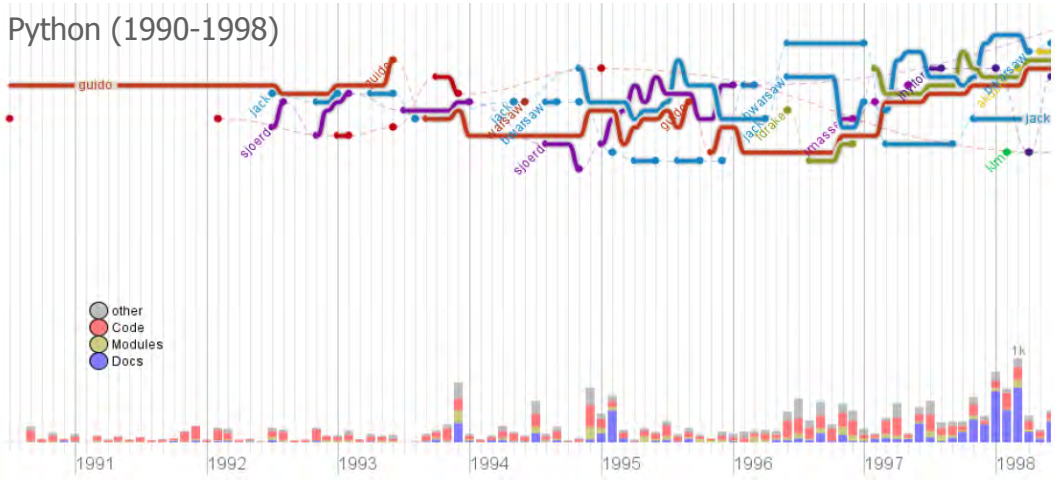
Python – Eclipse – Little Big Planet

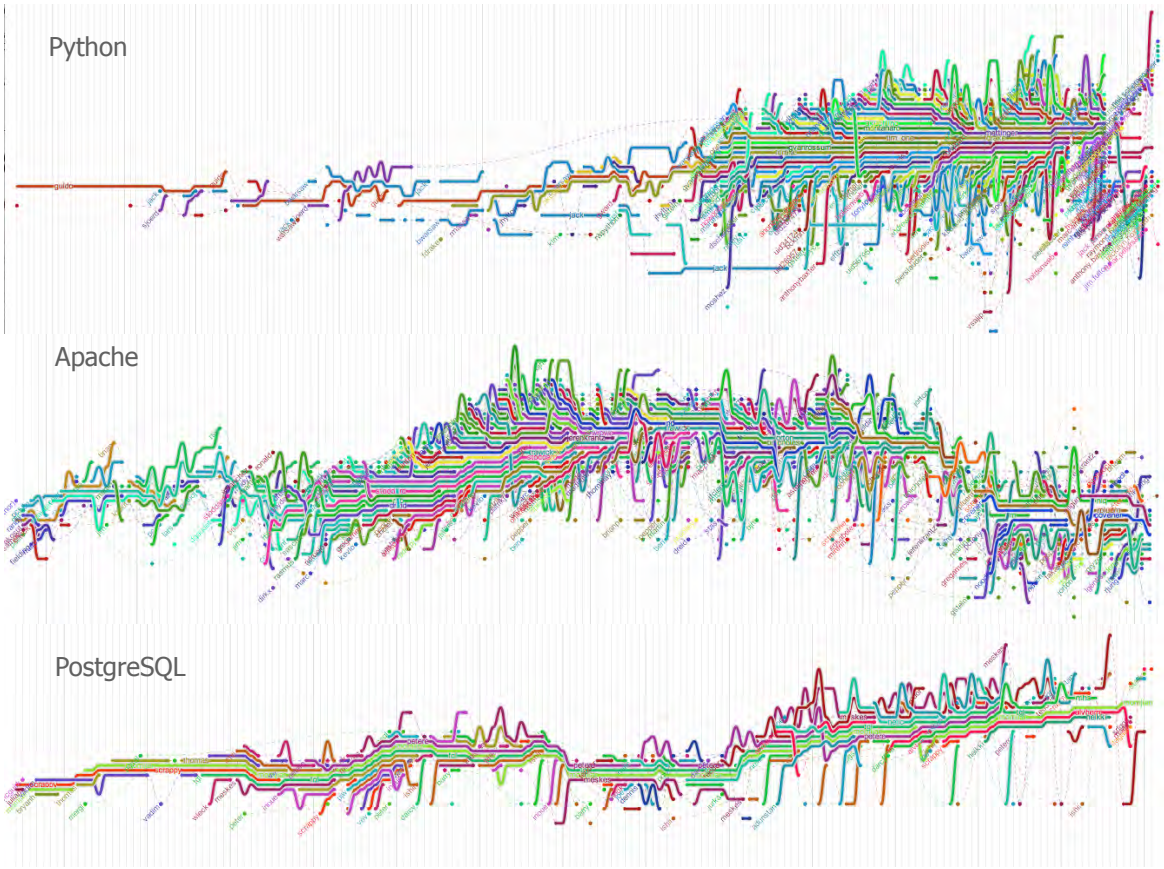


Litter Big Planet

Evolution Storyline Visualization

- Visualizing details of interaction among developers
- Inspired by XKCD's movie narrative charts and metro maps
- Complementing Code_Swarm





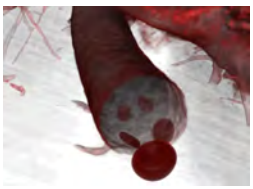
Dynamic Video Narratives

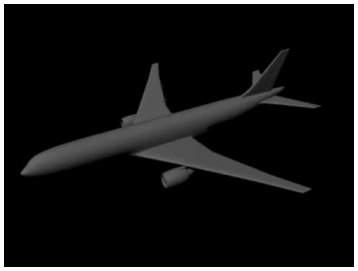
- Narratives as seamless dynamic compositions of a linear collection of mosaics
- A mosaic as a panoramic summary of a short video sequence occurring over a common background

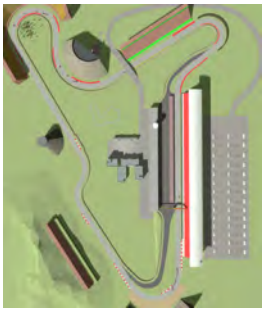
Dynamic Video Narratives

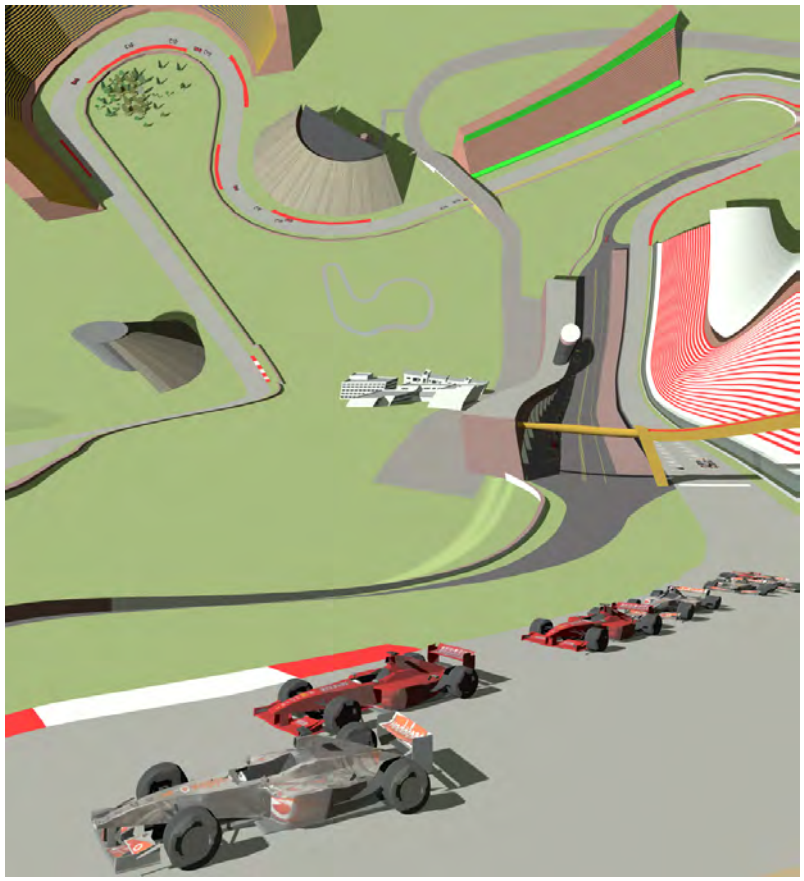
Dynamic Video Narratives

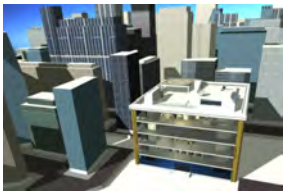
Multi-Scale Views of 3D Models

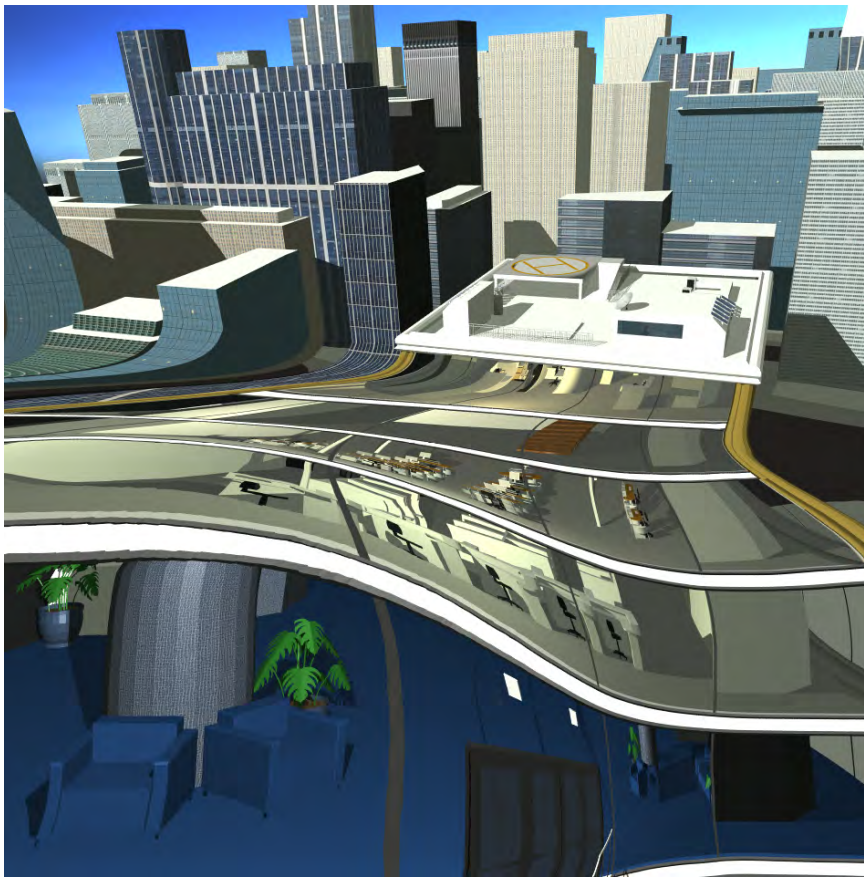












VIDi
Visualization & Interface Design Innovation

Kwan-Liu Ma

ma@cs.ucdavis.edu

http://www.cs.ucdavis.edu/~ma