



FCLib: A Library for Feature Characterization

**Wendy Koegler, Renata McCoy and
Philip Kegelmeyer**

Sandia National Laboratories

DOEGCF 2004

April 26-27

Santa Fe, NM

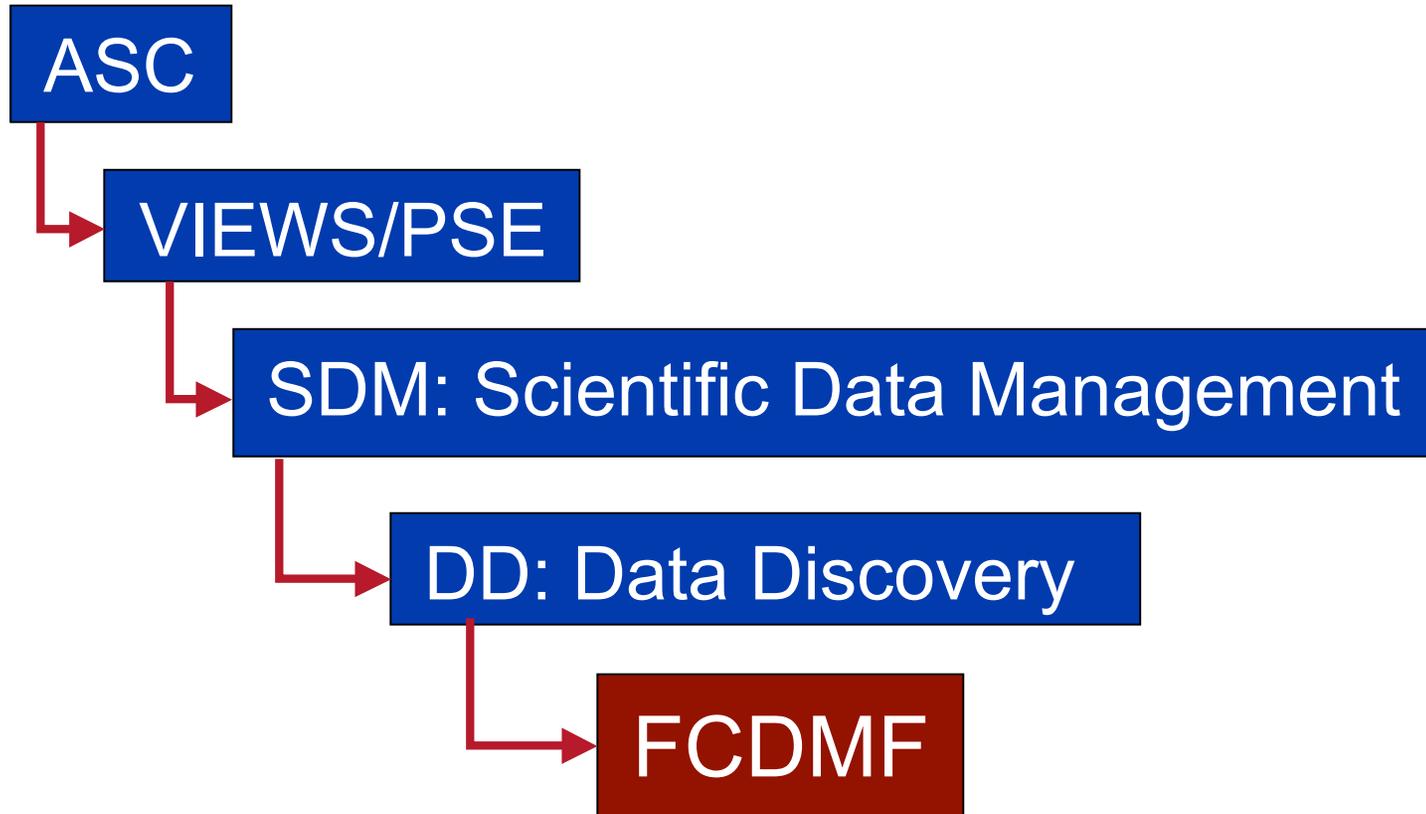


Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company,
for the United States Department of Energy under contract DE-AC04-94AL85000.





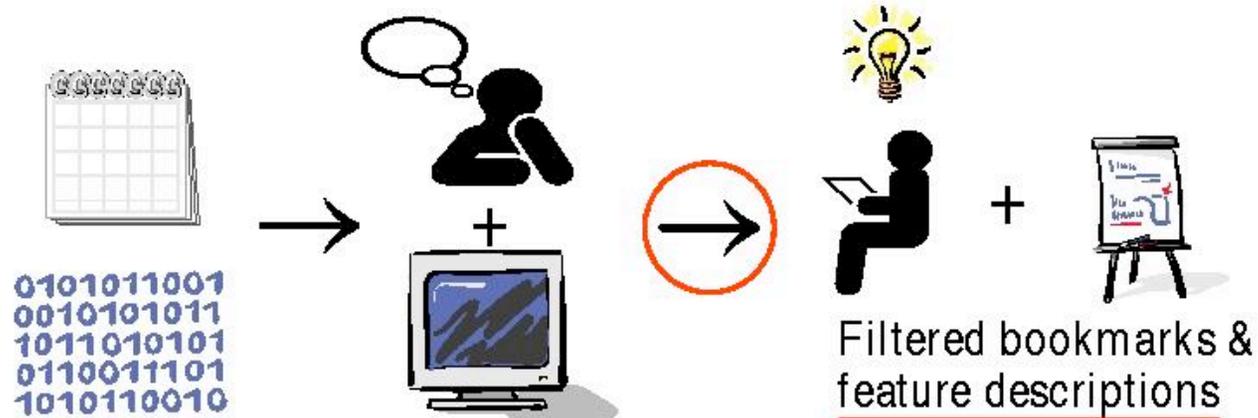
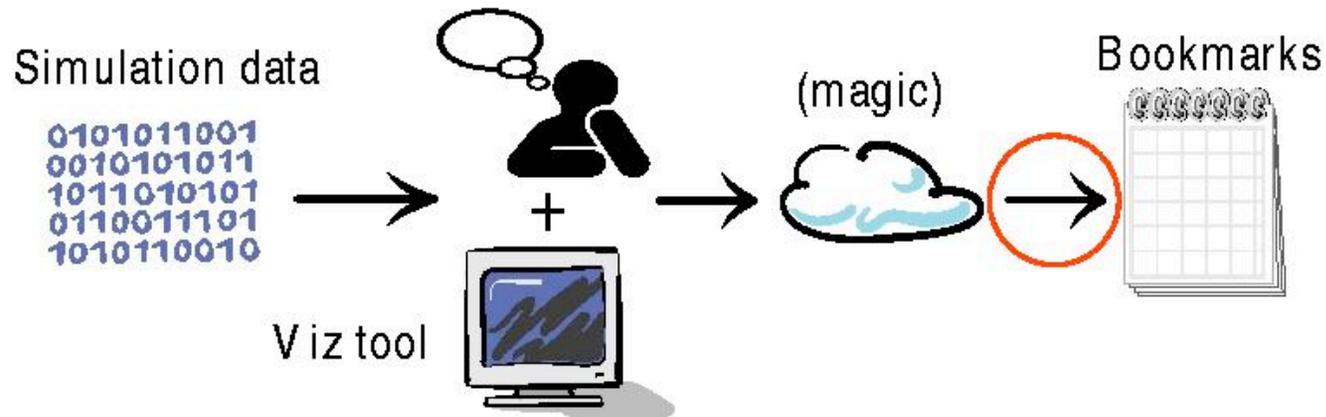
Project FCDMF: Feature Characterization



<http://www.ca.sandia.gov/ASCI/sdm/> → Data Discovery → FCDMF



Data Discovery: Use Scenario





Outline

- **Background**
- **Library overview**
- **Custom applications**

- **Side themes**
 - **Visualization vs(?) data analysis**
 - **Data models**

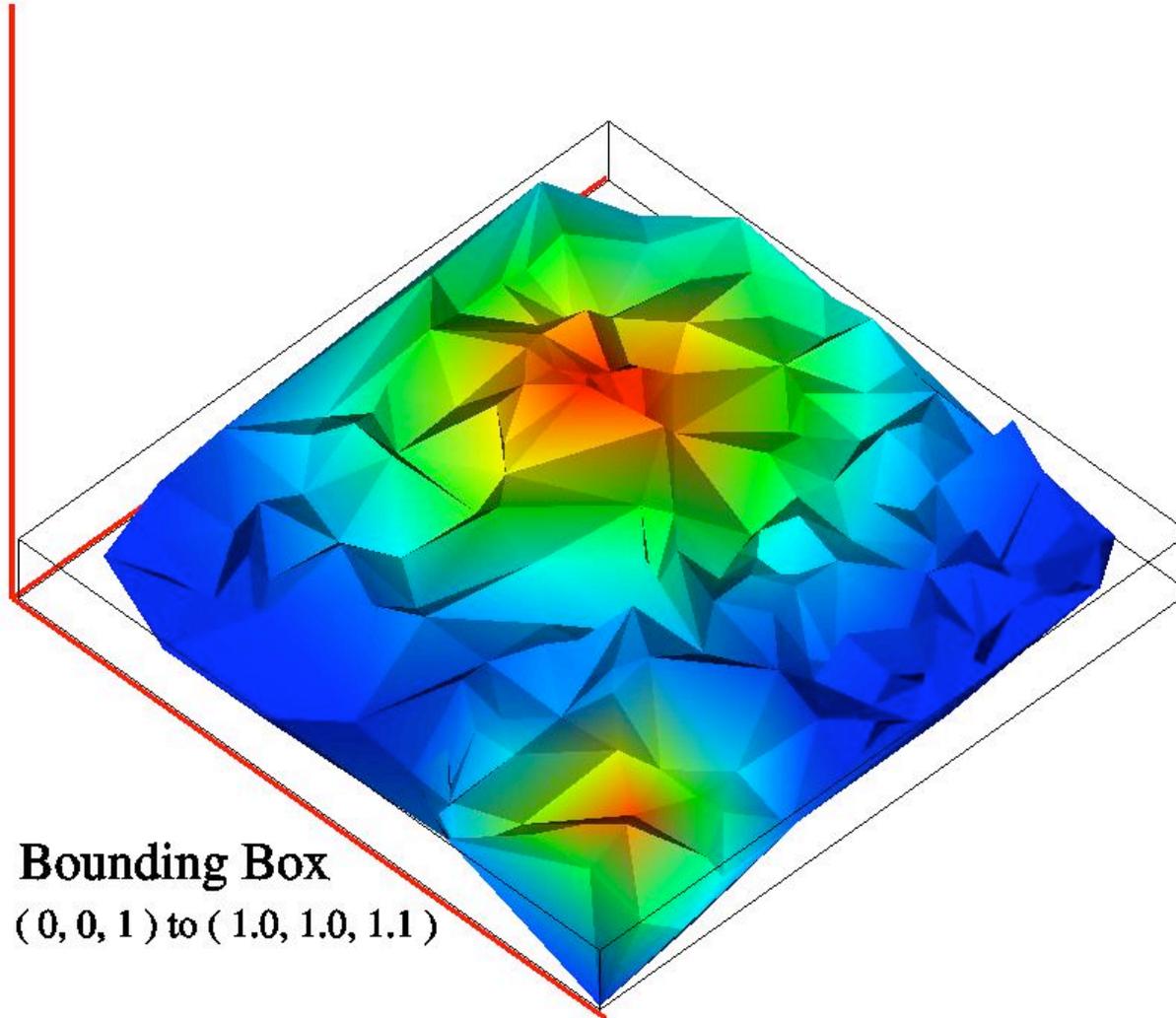


What are Features?

- **Attribute**: a low level property of the data
 - Physics variables (temperature, velocity, stress)
 - Geometry variables (local curvature, connectivity)
- **Feature**: a high level structure or phenomena
 - Vortex tube
 - Flame front
 - Crumple zone, tear
 - Representation = spatial-temporal region of dataset
- **Characterization**: analysis, reduction, representation, etc.

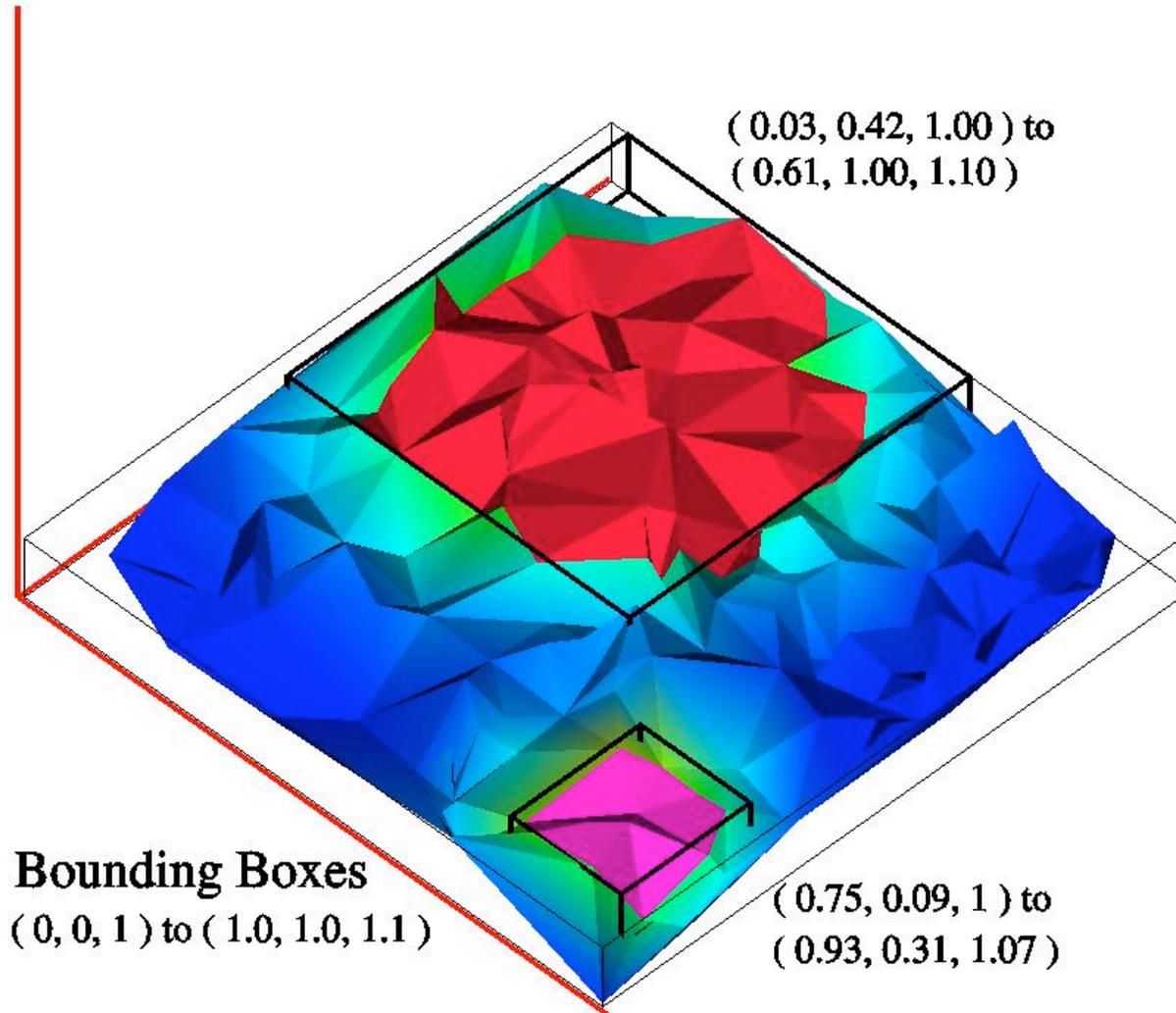


Example: Attribute = Temperature

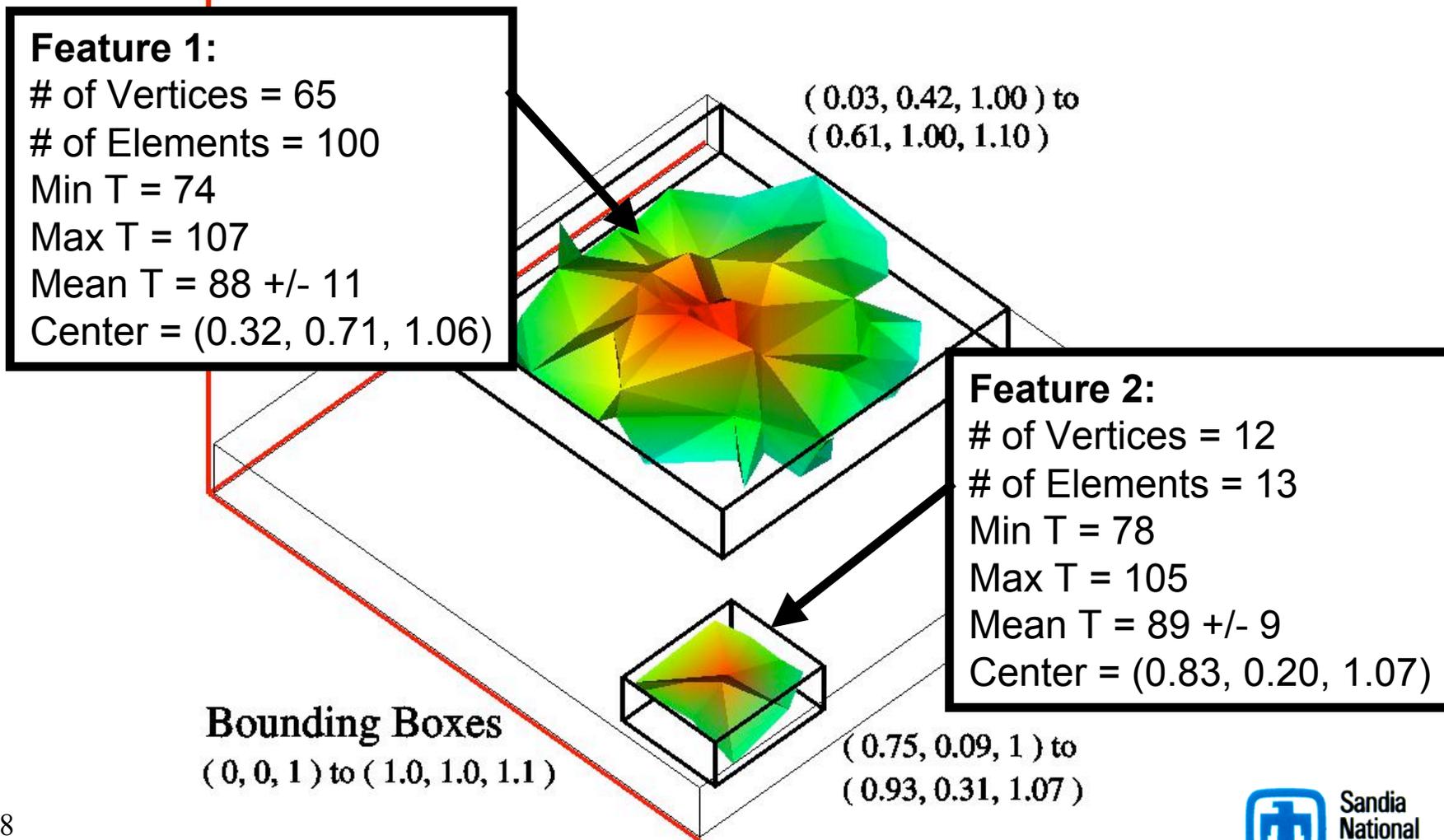




Example: Features = High Temperature



Example: Characterizations





FCDMF Goal

Goal:

- Provide **general resources** for the creation and use of feature characterizations.
- **Automate** characterization steps, but not replace human

Challenges:

- Multiple levels of users.
- Features and characterizations tend to be very specific to domain area and data representation.
- Algorithms an open area of research.

Advantages

- Algorithms and computations are not required to scale.



Feature Characterization Library (FCLib)

- **Prototype C library for constructing characterization routines and applications**
 - **Feature aware**
 - **Simple API for two levels of users**
 - **Building blocks for characterization developers**
 - **“Check-box” characterizations**
 - **Unstructured meshes**
 - **Serial**
 - **IO = SAF**



FCLib Components

- **Data manipulation**
- **Feature processing**
- **Characterization building blocks**
- **Generic characterizations**
- **Domain specific characterizations**
 - When build one of these, try to milk for generic characterizations and building blocks
- **<http://calico.ca.sandia.gov/fclib-doc/>**



Simple Data Interface

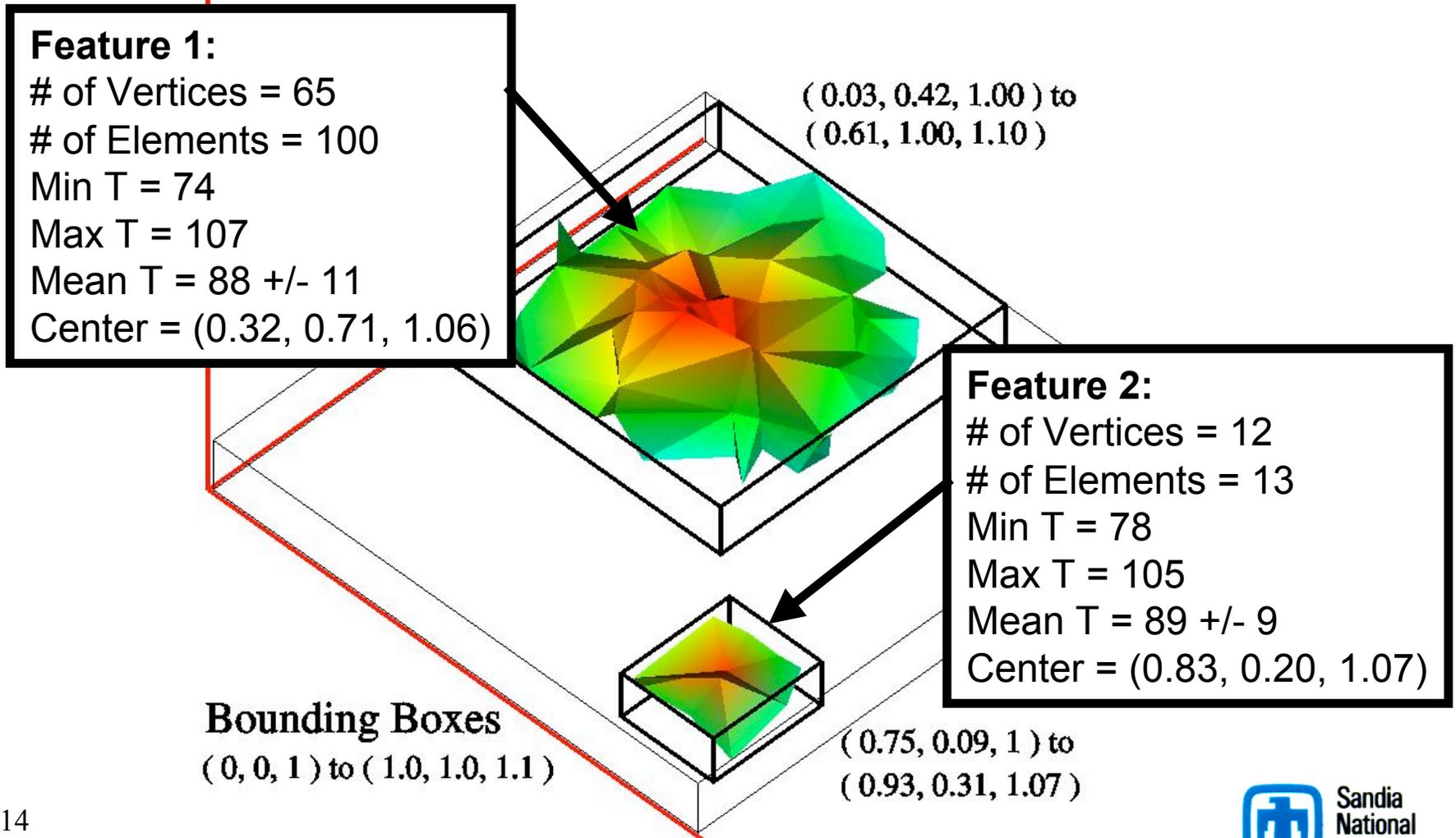
- **Five basic object types**
 - Dataset
 - Sequences (e.g. time series)
 - Meshes
 - Variables (sequence variable = array of vars)
 - Subsets = set of sub-element IDs (e.g. vertices)
- **Opaque handles**
- **Metadata vs. big data calls**
- **Most users will not have to manipulate big data arrays**



Building Blocks & Simple Characterizations

- **Mesh-based**
 - Bounding box, Centroid
 - Edge lengths, Face areas, Element volumes
 - Vertex normals, Surface normals
 - Topological neighbors
 - Geometric neighbors
- **Variable-based**
 - Weighted centroid
 - Variable math
 - Threshold
 - Statistics (min/max/mean/std.dev.)

Example: Variable Statistics of Features





Example: Variable Statistics of Features

```
fc_openDataset("tri-surface.saf", &dataset);
fc_openMesh(dataset, "tri mesh", &mesh);
fc_openVariable(mesh, "temperature", &var);

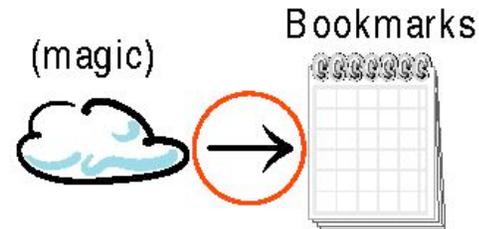
fc_threshold(var, ">=", 100.0, &subset);
fc_segment(subset, &numFeature, &features);
for (i = 0; i < numFeature; i++) {
    fc_getSubsetMinMax(var, feature, &min, &minID, &max, &maxID);
    fc_getSubsetMeanStdev(var, feature, &mean, &stdev);
    /* print results */
}

/* close things */
```



Custom Characterization App I

- **Customer: View Selection**
- **Find Regions of High Saliency**
 - Report bounding boxes
 - But first “smooth” the variable



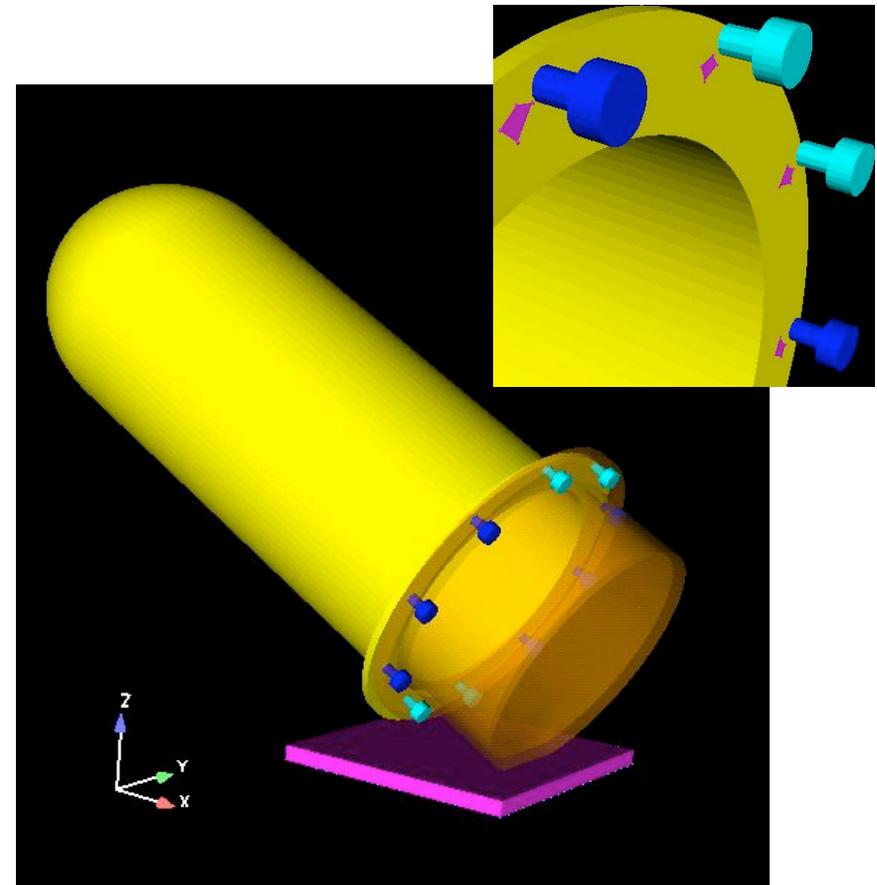
```
>threshBoundingBox -r 10 data.saf temperature ">" 100.0
```

```
0 7.2727 0.6667 0.0054 9.0909 2.0000 0.3050
0 0.9090 4.0000 0.0054 5.4545 7.3333 4.7532
1 7.2727 0.6667 0.0012 9.0909 2.6667 0.7142
1 0.0000 3.3333 0.0009 5.4545 8.0000 4.7532
2 0.0000 0.6667 0.0002 9.0909 8.0000 4.7532
3 0.0000 0.0000 0.0000 10.0000 8.0000 4.7532
4 ...
```



Custom Characterization App II

- Customer: Jay Dike, Analyst
- Evaluate spot weld integrity from Sierra Presto runs
 - How many fail & when
 - Failure parameter **new!**
- Fully automated
- Output graphs and data files
- Integrated with SimTracker





Future Plans

- **Continue soliciting users for requests**
- **On-the-fly characterization, tie-in with Sierra**
- **Integration with viz packages and/or GUIs**
- **More time-based characterizations, events**
- **“Smart” dataset comparisons using characterized features.**

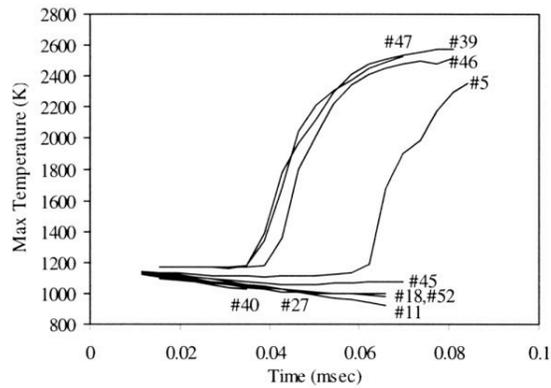


Data Analysis Vs.(?) Visualization

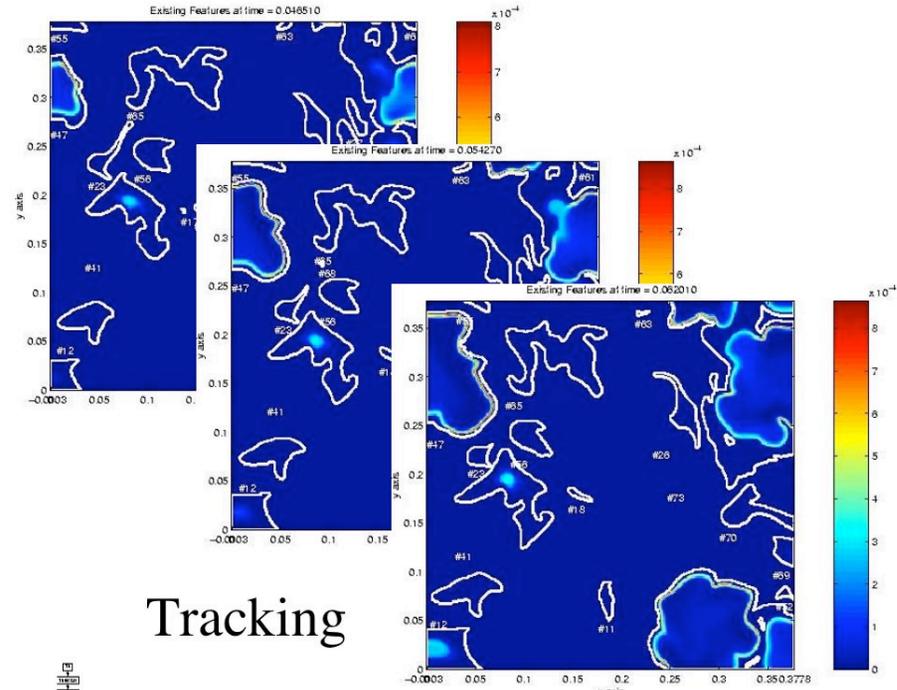
- **Many things in common**
 - Goal of extracting information from data
 - Many similar intermediate data & building blocks
- **Some differences I've encountered**
 - Final information encoding
 - Quantification & Reduction
 - Features: All variables in one spot vs. one variable everywhere.
 - Scalability requirements
 - Structures for topology, neighbor info



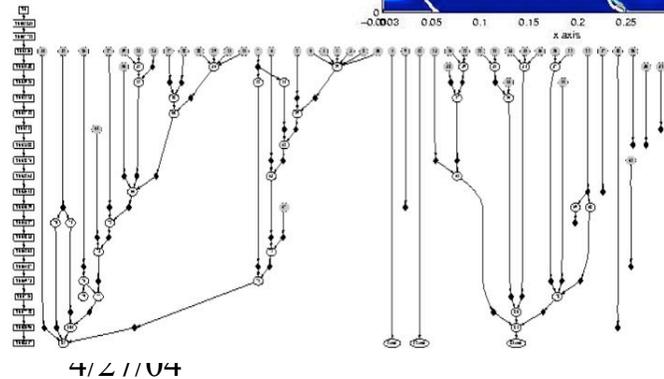
Conclusion



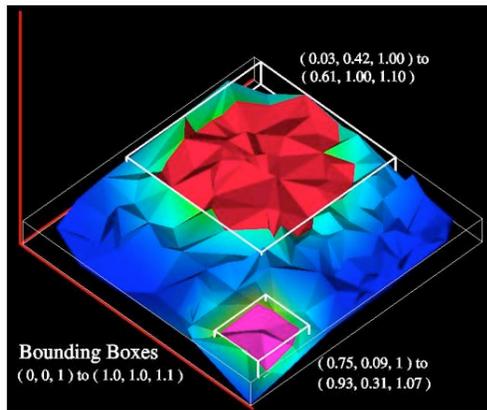
Analysis



Tracking



and more...



Representations