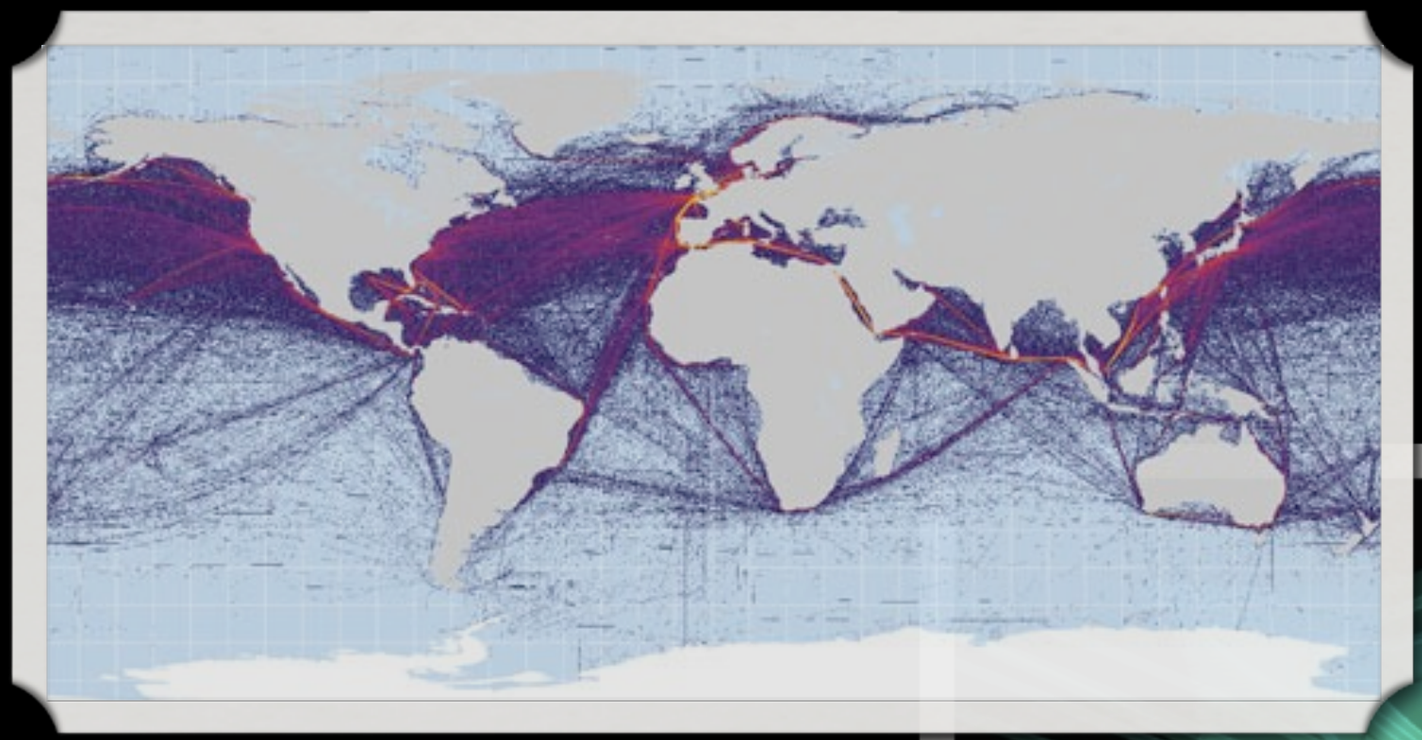
Parallel Visualization With Python

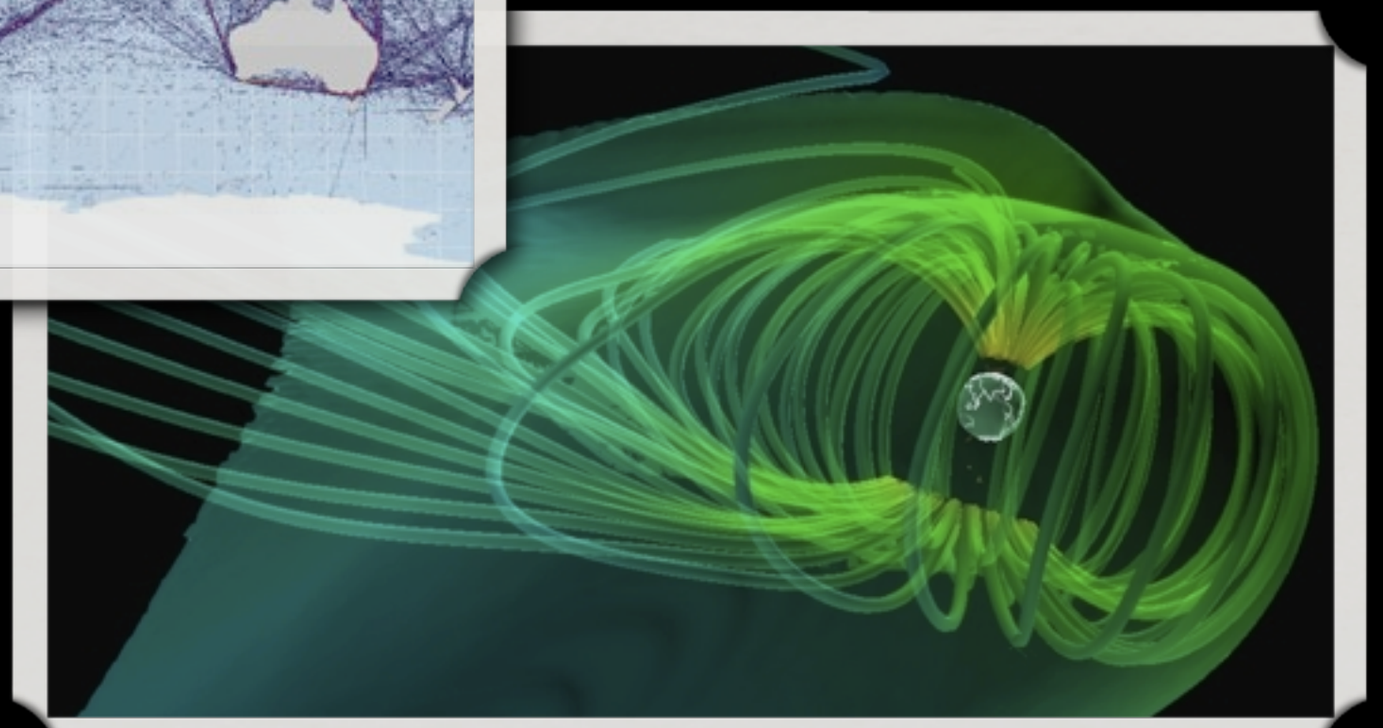
Christine Corbett Moran

EuroPython 2011

Why Visualize?

- insight, communication, comparison, appreciation

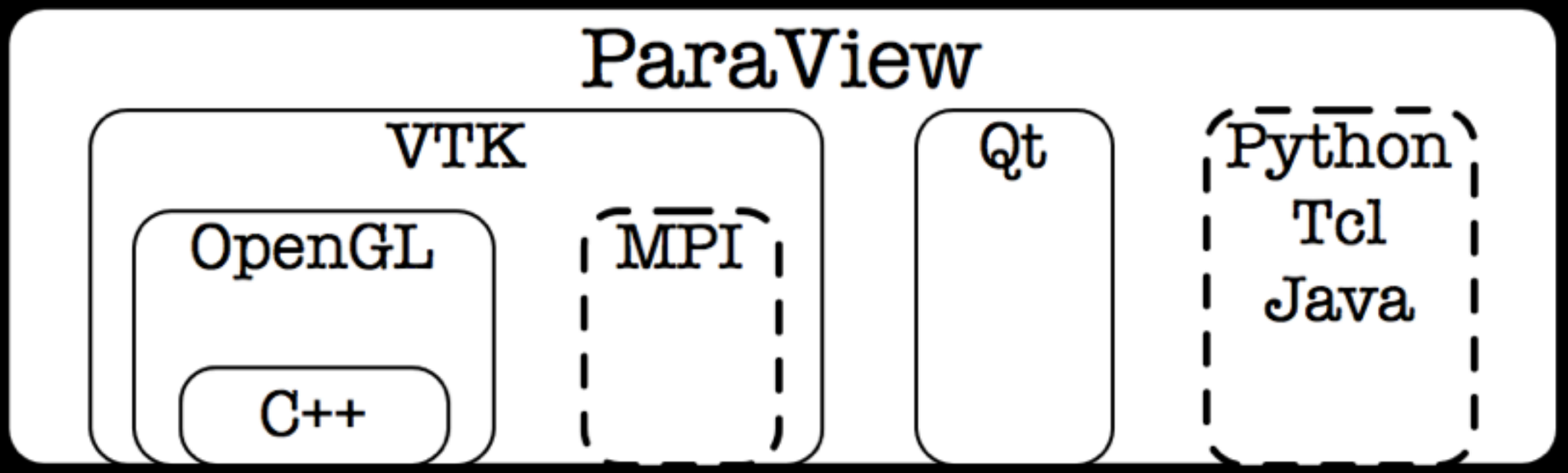




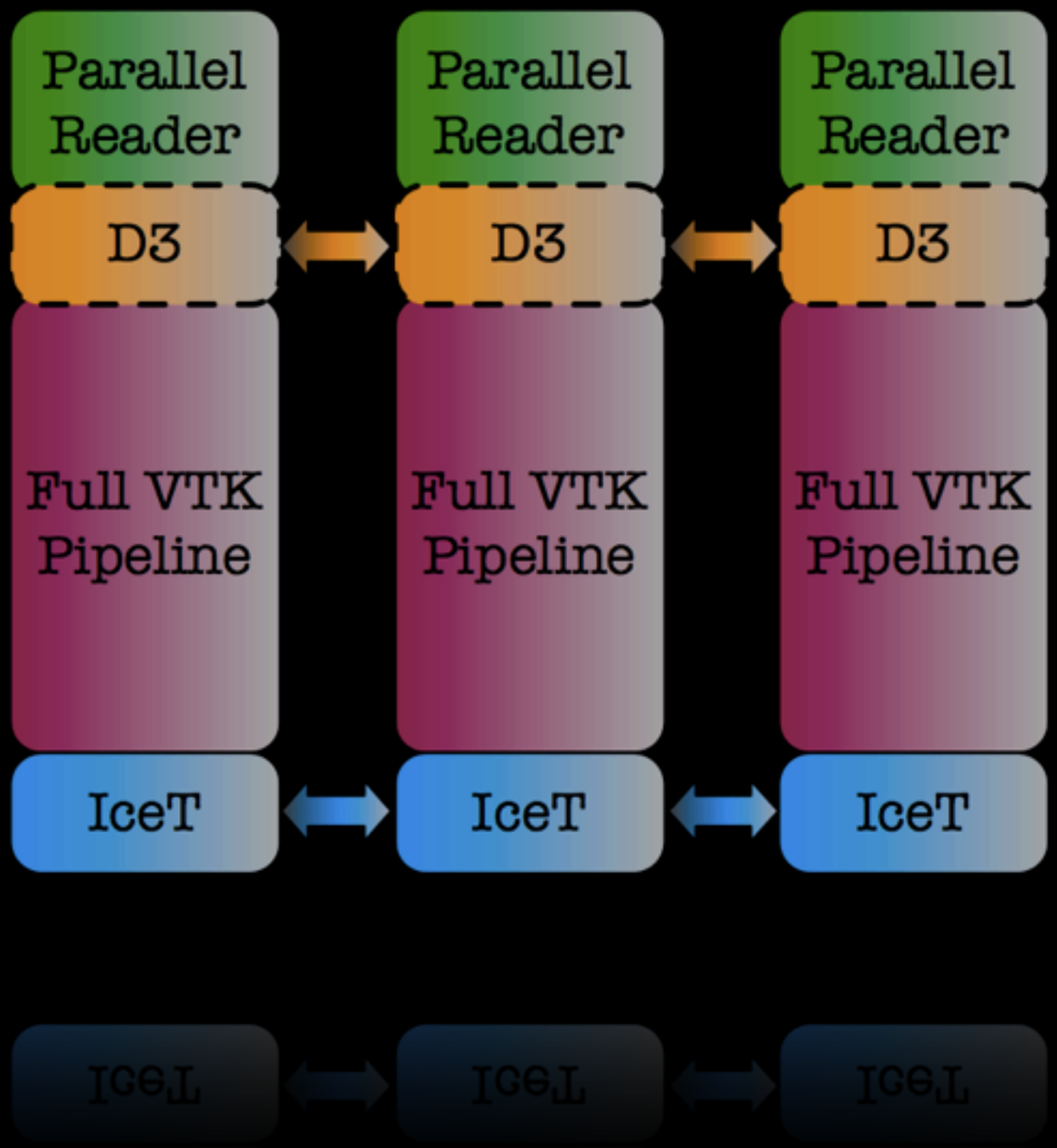
Why Parallel?

- desktop machines have many cores
  - 2,4,6,12...
- datasets are big
  - Ngrams: 2.2TB
  - Wikipedia Traffic stats: 150GB
- interact with data = discover new patterns

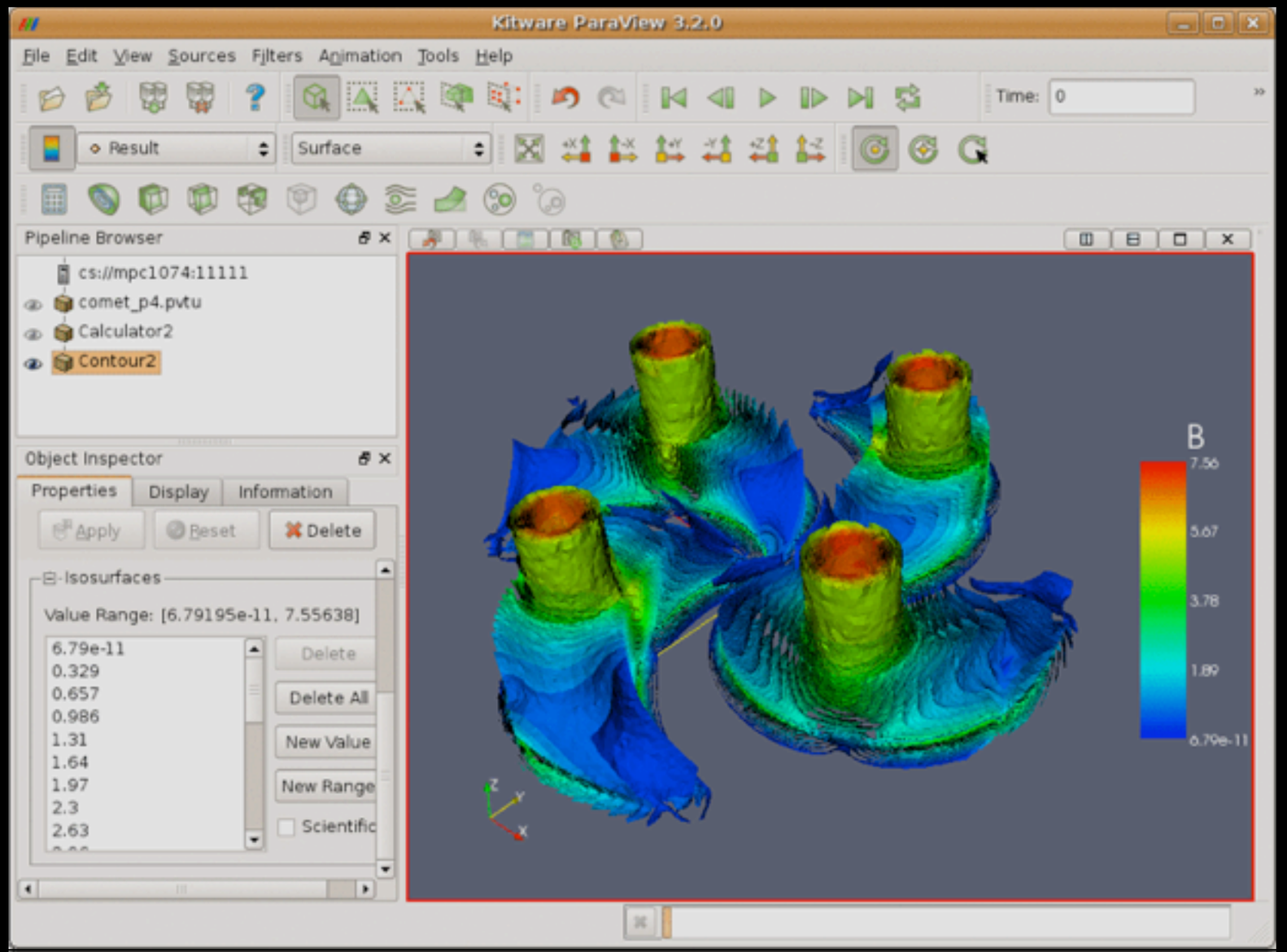
Why ParaView?



Why ParaView?



Why Python?


Kitware ParaView 3.2.0
File Edit View Sources Filters Animation Tools Help
Time: 0
Result
Surface
Pipeline Browser
cs://mpc1074:11111
comet_p4.pvtu
Calculator2
Contour2
Object Inspector
Properties Display Information
Apply Reset Delete
Isosurfaces
Value Range: [6.79195e-11, 7.55638]
6.79e-11
0.329
0.657
0.986
1.31
1.64
1.97
2.3
2.63
Delete
Delete All
New Value
New Range
Scientific
B
7.56
5.67
3.78
1.89
6.79e-11

Why Python?

```
int vtkTipsyReader::RequestInformation(
  vtkInformation* vtkNotUsed(request),
  vtkInformationVector** vtkNotUsed(inputVector),
  vtkInformationVector* outputVector)
```

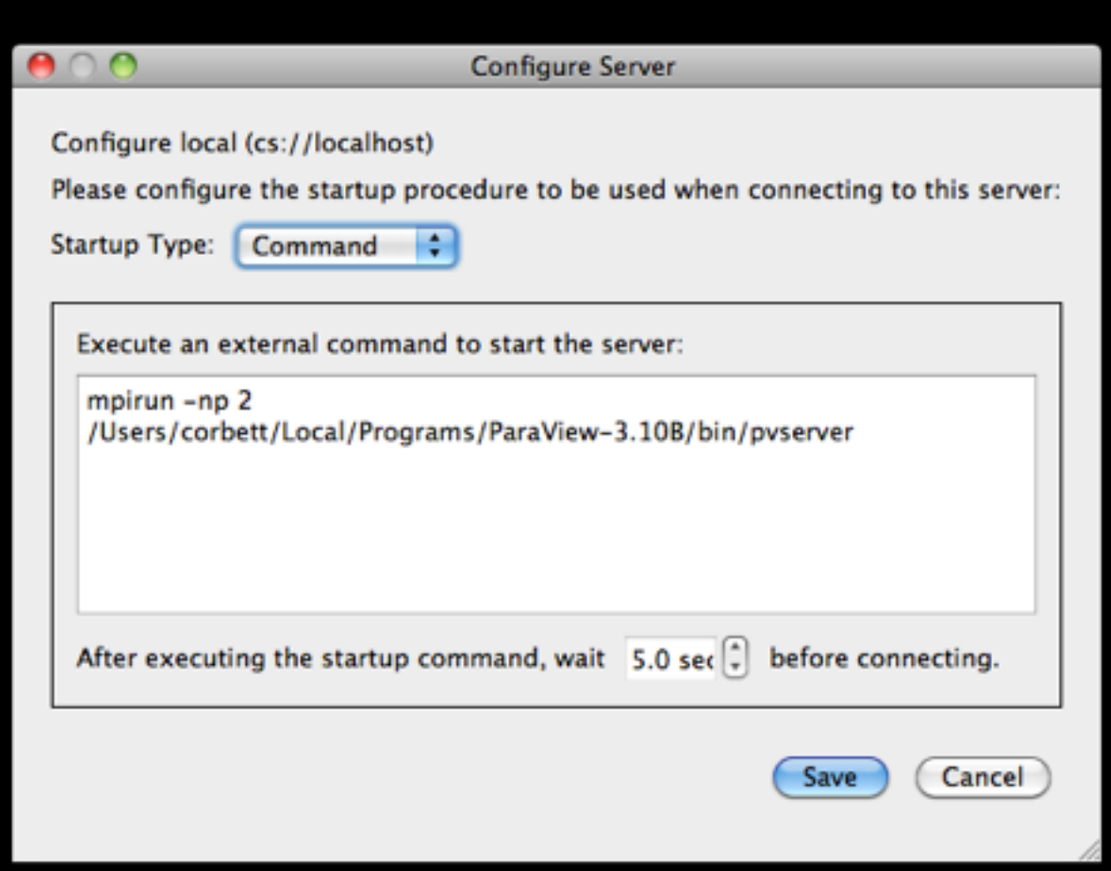
```
<ServerManagerConfiguration>
  <ProxyGroup name="sources">
    <SourceProxy name="TipsyReader"
                 class="vtkTipsyReader">

      <StringVectorProperty
        name="MarkFileName"
        command="SetMarkFileName"
        number_of_elements="1">
        <FileListDomain name="files"/>
        <Documentation>
          If set, only particles specified in the marked file are read in.
        </Documentation>
      </StringVectorProperty>
```

Why Python?

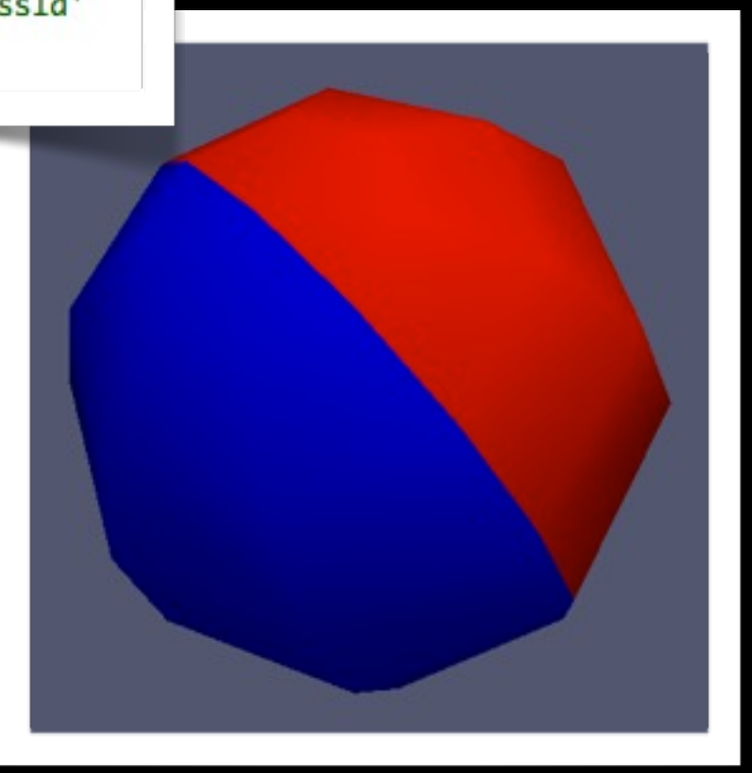
```
# Reading in from file
points = CSVReader(FileName=[ROOT_DIR+'/Tutorials/Data/simple.csv'])
points.UpdatePipeline()
# Making this spreadsheet into coordinates we can visualize
table = TableToPoints()
table.XColumn = 'x'
table.YColumn = 'y'
table.ZColumn = 'z'
table.UpdatePipeline()
# Instead of points lets visualize things as a sphere
sphereglyph = Glyph(GlyphType='Sphere')
dr=Show()
dr.ColorArrayName='random'
Render()
ResetCamera()
#Now Let's Save the Image into a Png
WriteImage(ROOT_DIR+'/Tutorials/Screenshots/CSV.png')
```

What Progress?

- Creating simple visualizations

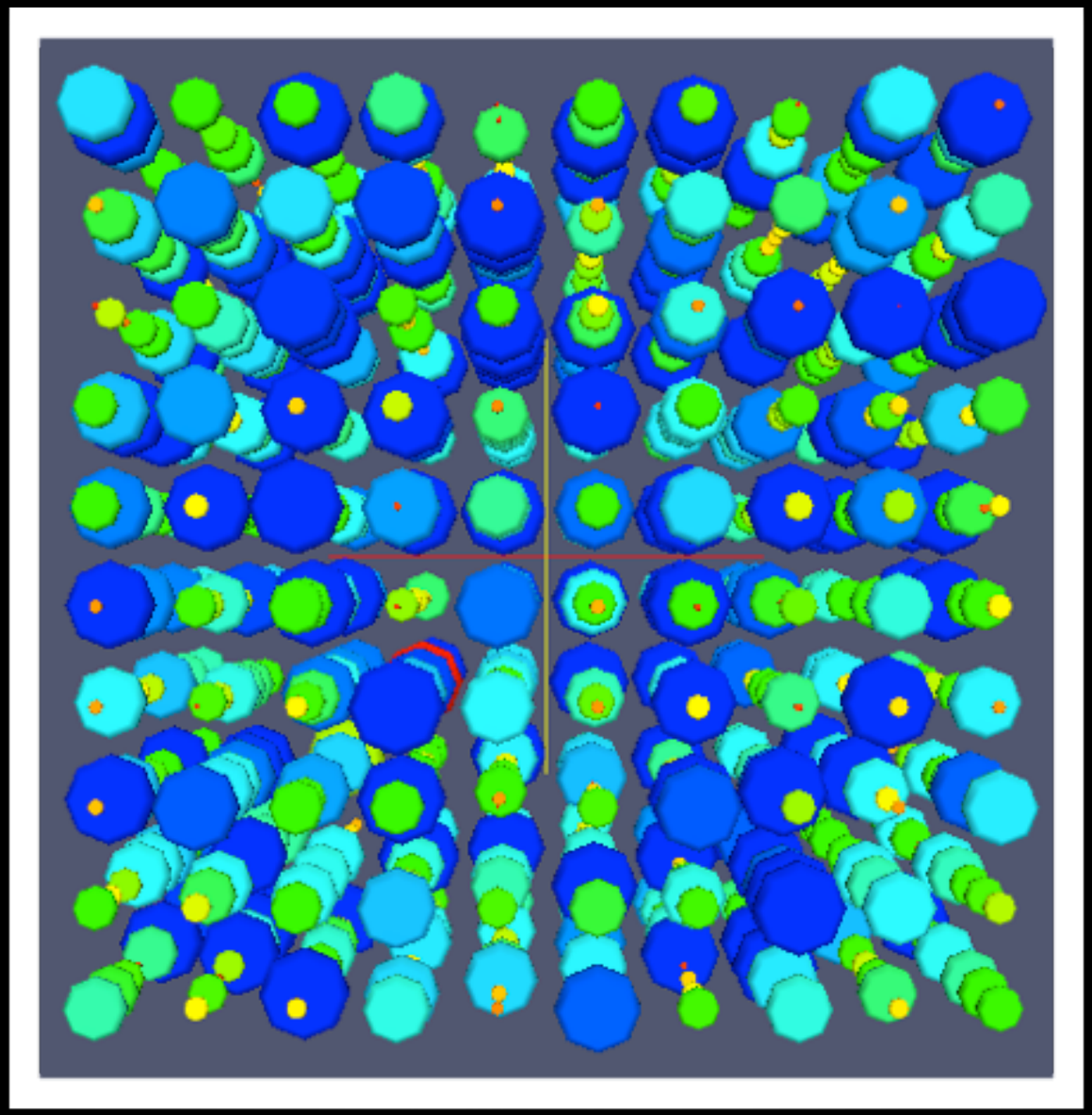



```
sph=Sphere()
dr1=Show()
Render()
pid=ProcessIdScalars()
dr2=Show()
dr2.ColorArrayName='ProcessId'
Render()
```



What Progress?

- Reading in data from CSV and binary files



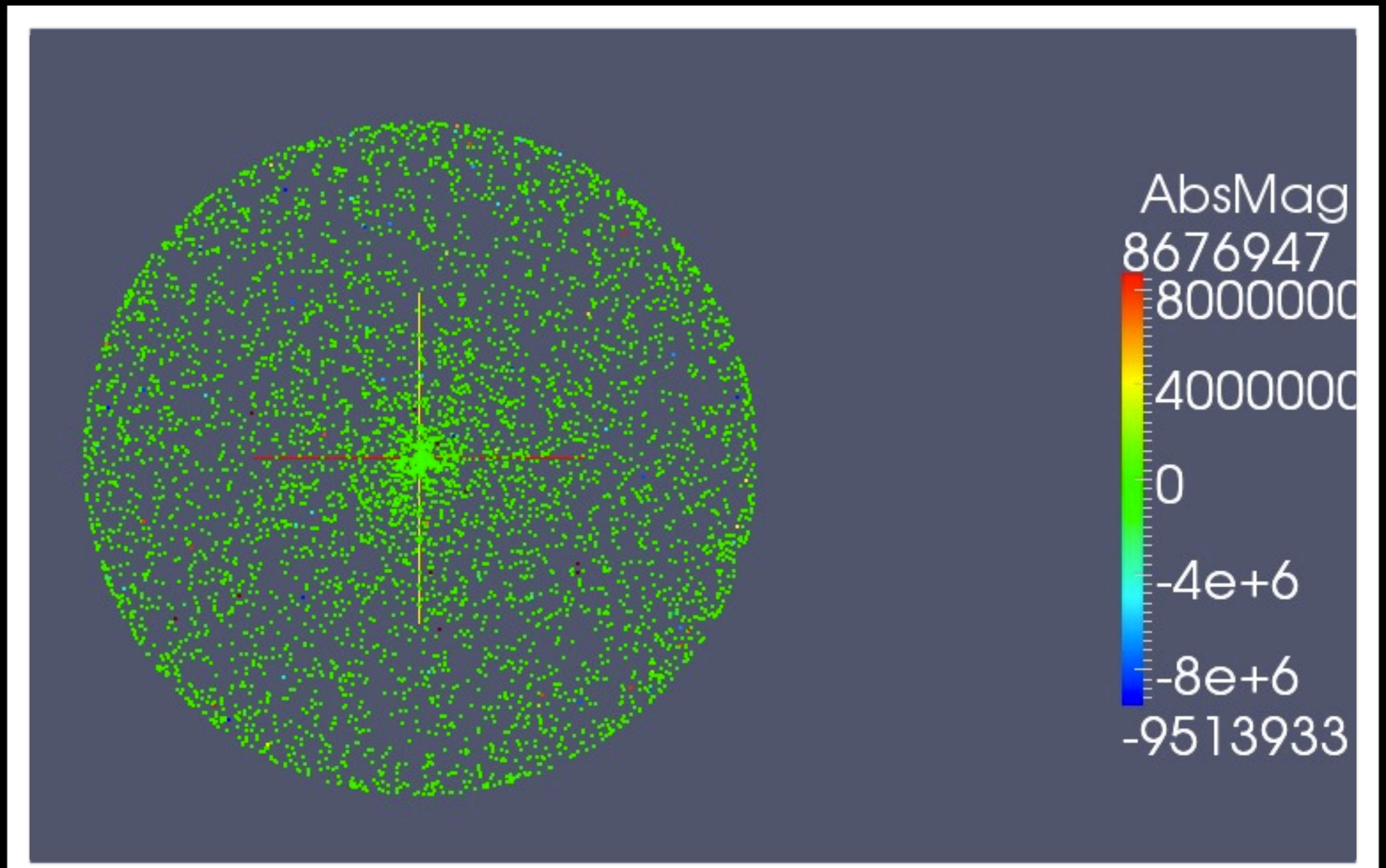
What Progress?

- Reading in data from CSV and binary files

```
# Reading in the data
random_raw = ImageReader(FilePrefix='random.raw',
   ScalarArrayName = 'random number',
   DataExtent = [0, 9, 0, 9, 0, 9],
   DataByteOrder = 'LittleEndian',
   DataScalarType = 'double')
 random_raw.UpdatePipeline()
# Structured data such as this doesn't automatically get a display in the
# GUI, we have to create a glyph.
glyph = Glyph( GlyphType="2D Glyph",
   Scalars = ['POINTS', 'random number'],
   ScaleMode = 'scalar'
   )
glyph.GlyphType.GlyphType = 'Vertex'
dr=Show()
dr.ColorArrayName='random number'
Render()
```

What Progress?

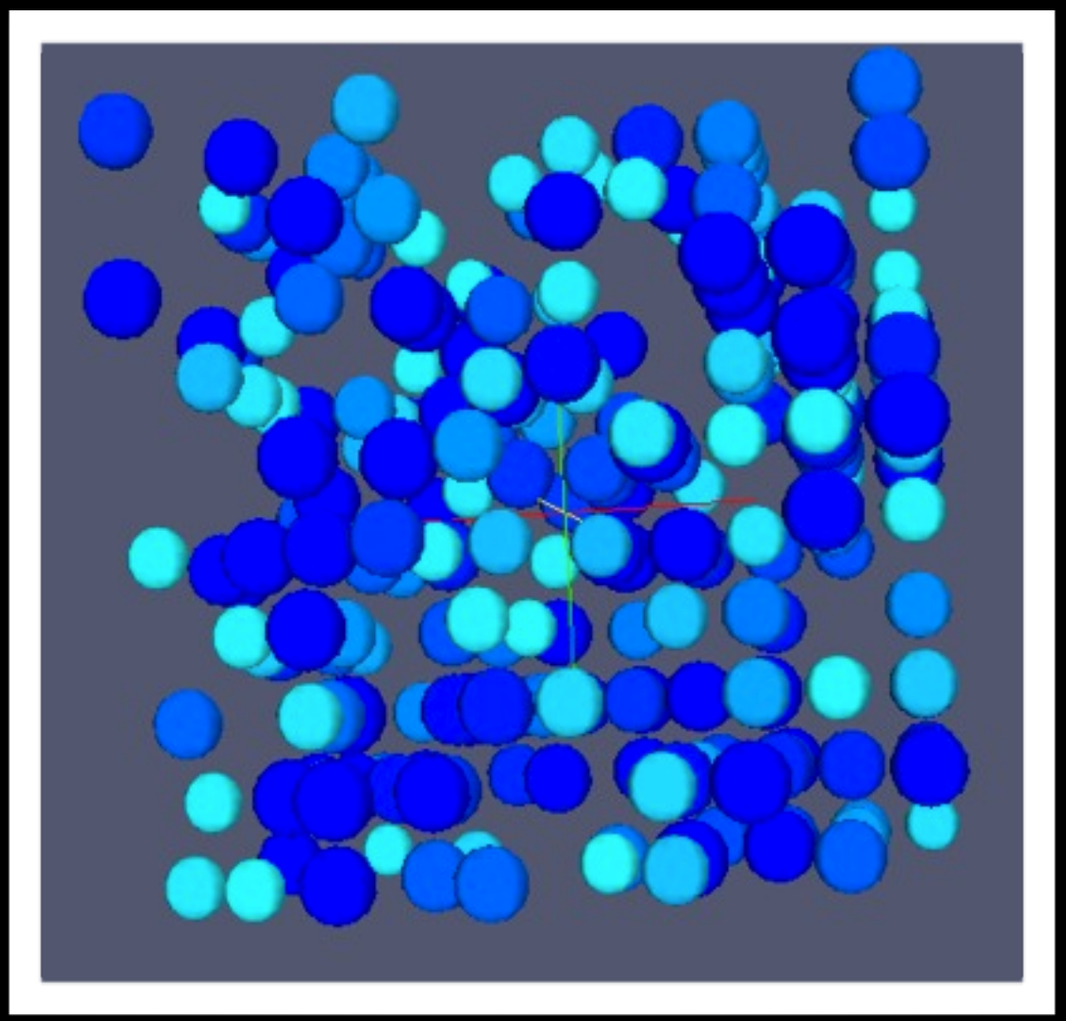
- Playing with data from the web



What Progress?

- Using filters to do analysis

```
thresh = Threshold(Scalars = ['POINTS', 'random number'],
  ThresholdRange = [.75, 1.0])
dr=Show()
dr.ColorArrayName='random number'
Render()
```

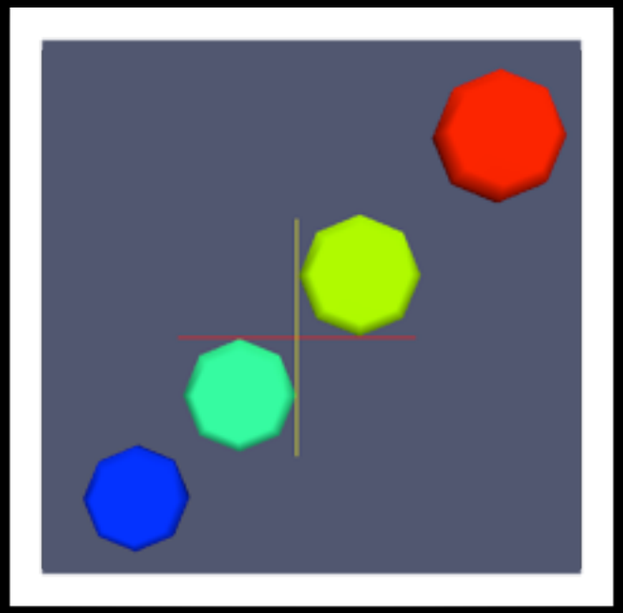


What Progress?

- Writing your own custom readers, writers and data filters

```
from paraview import vtk
# where we'll put our data
output=self.GetOutput()

# Our data: points and arrays
# Points
points=vtk.vtkPoints()
for point in [(0,0,0),(1,1,1),(2,2,2),(3,3,3)]:
  points.InsertNextPoint(point[0],point[1],point[2])

# One Scalar Array
temp=vtk.vtkDoubleArray()
temp.SetName('Temperature')
temp.SetNumberOfComponents(1)

for val in [0.0,1.1,2.2,3.3]:
  temp.InsertNextValue(val)

# Placing these in the output dataset
output.SetPoints(points)
output.GetPointData().AddArray(temp)
# There we are!
```

```
progSource =\
  ProgrammableSource(Script=\
  open(ROOT_DIR+'/Tutorials/Projects/CustomReader/SimplePoints.py').read())
progSource.UpdatePipeline()
glyph = Glyph(GlyphType='Sphere',Scalars = ['POINTS', 'Temperature'],
caleMode = 'off')
dr=Show()
dr.ColorArrayName='Temperature'
dr.LookupTable=GetLookupTableForArray('Temperature', 1 )
Render()
ResetCamera()
WriteImage(ROOT_DIR+'/Tutorials/Screenshots/CustomReader.png')
```



What Projects?

- What's possible
  - Visualize and manipulate TBs of data in one go, in real time.
- What's next
  - What data do you want to visualize/ analyze? Q&A how to make that happen.

Pre-reqs

- Download and run ParaView binary
  - http://www.paraview.org/paraview/resources/software.html
- Problems? Ping @corbett_inc
- Tutorial files on the net, available via USB