



Cornell University  
Center for Advanced Computing

# Parallel Lab

Aaron Birkland  
Cornell Center for Advanced Computing

Data Analysis on Ranger  
January 2012



## Getting started

- Go to Longhorn portal:  
<http://portal.longhorn.tacc.utexas.edu>
- Sign in using your XSEDE credentials

Not logged in.

**TACC XSEDE Longhorn Visualization Portal**

Home Allocations Jobs Help

**Welcome to the Longhorn Visualization Portal!**  
A simplified interface to the Longhorn visualization resource

**Featuring:**

- + Visualization job submission and monitoring
- + Remote, interactive, web-based visualization
- + Guided visualization using EnVision
- + Current resource usage and allocation view

**System Requirements:**

- + JavaScript & Java applets Enabled
- + Java 1.4.2+
- + Windows: Firefox 3+

Authenticate as:

- XSEDE User Portal User
- TACC User Portal User

Username

Password

Login

Job Submission VNC Visualization Session EnVision Guided Visualization



## Set your VNC password

- Click “here” to set/change your VNC password
  - This is exactly equivalent to using `vncpasswd` on the command line
  - If you forget to do this, and have never set the vnc password before, you will be prompted. Don't worry

**TACC XSEDE Longhorn Visualization Portal** XSEDElab18 logout  
No job running

Home Allocations Jobs Help

### Start a Job

Resource: Longhorn

Session type:  VNC  EnVision guided visualization

Desktop resolution: 1280x1024

Number of nodes: 2

Wayness (processes per node): 4

Note: increasing the number of nodes will only increase performance for parallel applications (e.g. ParaView or VisIt). The wayness parameter is only relevant to parallel applications, not determines how many processes are executed per node. A parallel application is executed.

Start

[Click here to set your VNC password.](#)

### Available Resources

- Longhorn

Longhorn (longhorn.tacc.utexas.edu), TACC's Dell XD Visualization Cluster, contains 2048 compute cores, 14.5 TB aggregate memory and 512 GPUs. Longhorn has a QDR InfiniBand interconnect and has an attached Lustre parallel file system. Longhorn is connected by 10GbE to Ranger's Lustre parallel file system thus making it more convenient to work on datasets generated on Ranger. Longhorn has 256 nodes + 2 login nodes, with 240 nodes containing 48GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 5800 GPUs. Longhorn also has an additional 16 large-memory nodes containing 144GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 5800 GPUs. For more detailed information on Longhorn, please see the Longhorn User Guide.

**Queue information:**  
updated at January 8, 2012, 12:34:41 pm (refresh)

Available

The Longhorn queues are open.  
125 nodes available out of 250 total.

Used

**ACTIVE JOBS**

JOBID	JOBNAME	USERNAME	STATE	COPE	REMAINING	STARTTIME
122446	Hadoop1v	lwoven	Running	40	22:32:19	Sun Jan 8 11:07:01
122483	iq_mss0007	lars	Running	120	05:07:19	Sun Jan 8 11:42:01
122504	iq_mss0008	lars	Running	120	05:36:19	Sun Jan 8 12:11:01
122509	iq_mss0009	lars	Running	120	05:39:19	Sun Jan 8 12:14:01
122510	iq_mss0010	lars	Running	120	05:43:19	Sun Jan 8 12:18:01
122511	h2-1P500	bloring	Running	64	00:53:19	Sun Jan 8 12:28:01
122512	h2-1	bloring	Running	64	00:55:19	Sun Jan 8 12:30:01



## Set your VNC password

- Enter a password in the window that pops up. Click “Set VNC password” button, then close
  - Do NOT use an important (e.g. login) password for this!! Remember, VNC passwords protect a running session only, and might be shared with others

The screenshot shows the TACC XSEDE Longhorn Visualization Portal. The page title is "TACC XSEDE Longhorn Visualization Portal". The navigation menu includes "Home", "Allocations", "Jobs", and "Help". The main content area shows "Processing..." and "Available Resources". Under "Available Resources", there is a section for "Longhorn" with a description of the cluster's capabilities. Below this is a "Queue information:" section with a pie chart showing "Available" and "Used" resources. A "Set VNC Password" dialog box is overlaid on the page, containing fields for "Password:" and "Re-enter:", a "Passwords match!" message, and "Set VNC Password" and "Cancel" buttons. Below the dialog box, there is a table of "ACTIVE JOBS" and "WAITING JOBS".

Processing...

**Available Resources**

- Longhorn

Longhorn (longhorn.tacc.utexas.edu), TACC's Dell XD Visualization Cluster, contains 2048 compute cores, 14.5 TB aggregate memory and 512 GPUs. Longhorn has a QE interconnect and has an attached Lustre parallel file system. Longhorn is connected by 10GigE to Ranger's Lustre parallel file system thus making it more convenient to work generated on Ranger. Longhorn has 256 nodes + 2 login nodes, with 240 nodes containing 48GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 580 also has an additional 16 large-memory nodes containing 144GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 5800 GPUs. For more detailed info please see the [Longhorn User Guide](#).

**Queue information:**  
updated at January 8, 2012, 4:32:48 pm (refresh)

Available  Used The Longhorn queues are of 242 nodes available out of 2

**ACTIVE JOBS**-----  
JOBID JOBNAME USERNAME STATE CORE MEM  
-----  
122446 Hadoop1w lwoven Running 40 18:34:12 Sun Jan 8 11:07:01  
122535 Hadoop-tes roller Running 24 04:40:12 Sun Jan 8 15:13:01  
2 active jobs : 8 of 248 hosts ( 3.23 %)

**WAITING JOBS**-----  
JOBID JOBNAME USERNAME STATE CORE WCLIMIT QUEUETIME  
-----

**WAITING JOBS WITH JOB DEPENDENCIES**---



## Start a VNC session

- Make sure VNC session type is selected
- Specify number of nodes = 2
- Specify wayness (processes per node) = 4
- Click “Start”

**TACC XSEDE Longhorn Visualization Portal** XSEDEVapb18 logout  
No job running

Home | Applications | Jobs | Help

### Start a Job

Resource: Longhorn

Session type:  VNC  EnVision guided visualization

Desktop resolution: 1280x1024

Number of nodes: 2

Wayness (processes per node): 4

*Note: increasing the number of nodes will only increase performance for parallel applications (e.g. ParaView or VisIt). The wayness parameter is only relevant to parallel applications, and determines how many processes are spawned per node when the parallel application is executed.*

[Click here to set your VNC password.](#)

### Available Resources

- Longhorn

Longhorn (longhorn.tacc.utexas.edu), TACC's Dell XD Visualization Cluster, contains 2048 compute cores, 14.5 TB aggregate memory and 512 GPUs. Longhorn has a QDR InfiniBand interconnect and has an attached Lustre parallel file system. Longhorn is connected by 10GigE to Ranger's Lustre parallel file system thus making it more convenient to work on datasets generated on Ranger. Longhorn has 256 nodes + 2 login nodes, with 240 nodes containing 48GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 5800 GPUs. Longhorn also has an additional 16 large-memory nodes containing 144GB of RAM, 8 Intel Nehalem cores (@ 2.5 GHz), and 2 NVIDIA Quadro FX 5800 GPUs. For more detailed information on Longhorn, please see the Longhorn User Guide.

**Queue information:**  
updated at January 8, 2012, 12:34:41 pm (refresh)

Available  Used

The Longhorn queues are open.  
125 nodes available out of 250 total.

**ACTIVE JOBS**

JOBID	JOBNAME	USERNAME	STATE	CORE	REMAINING	STARTTIME
122446	Hadoop14	lvowen	Running	40	22:32:19	Sun Jan 8 11:07:01
122483	iq_mss0007	lacs	Running	128	05:07:19	Sun Jan 8 11:42:01
122504	iq_mss0008	lacs	Running	128	05:36:19	Sun Jan 8 12:11:01
122509	iq_mss0009	lacs	Running	128	05:39:19	Sun Jan 8 12:14:01
122510	iq_mss0010	lacs	Running	128	05:43:19	Sun Jan 8 12:18:01
122511	h2-19500	bloring	Running	64	00:53:19	Sun Jan 8 12:28:01
122512	h2-1	bloring	Running	64	00:55:19	Sun Jan 8 12:30:01





## Start a VNC session

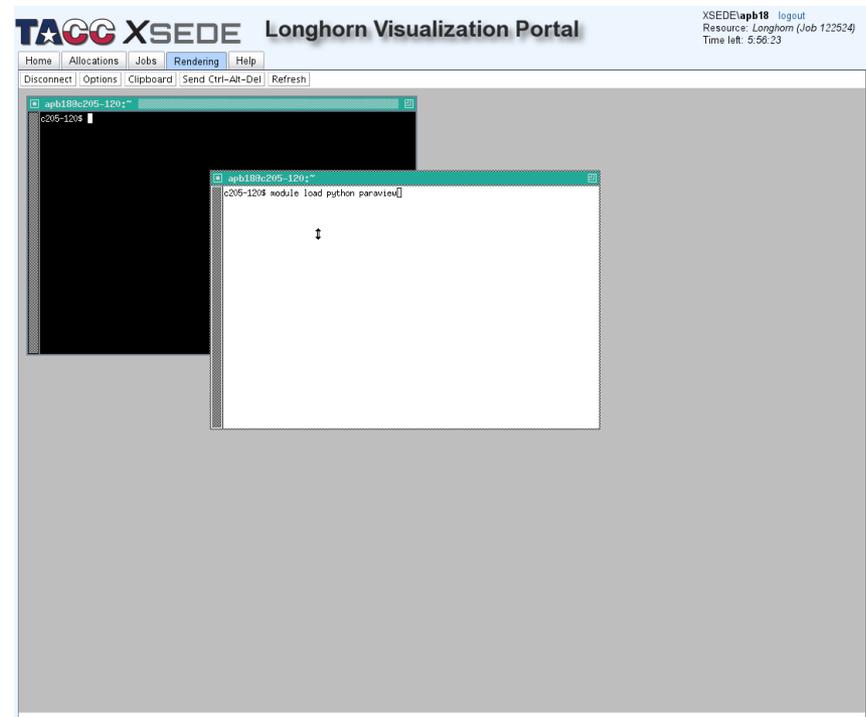
- A VNC session is now running. By default, you are brought to a “rendering” tab which has a built-in VNC viewer. You can use this if you wish, or just use your own VNC client.
- Type in your VNC password to access the session.

A screenshot of the XSEDE Longhorn Visualization Portal. The page title is "XSEDE Longhorn Visualization Portal". In the top right corner, it shows "XSEDElab18 logout", "Resource: Longhorn (Job 122524)", and "Time left: 5:59:45". The main navigation bar includes "Home", "Allocations", "Jobs", "Rendering", and "Help". Below this is a secondary navigation bar with "Disconnect", "Options", "Clipboard", "Send Ctrl-Alt-Del", and "Refresh". The central content area is titled "VNC Authentication" and contains a "Password:" label, a text input field, and an "OK" button.



## Run ParaView

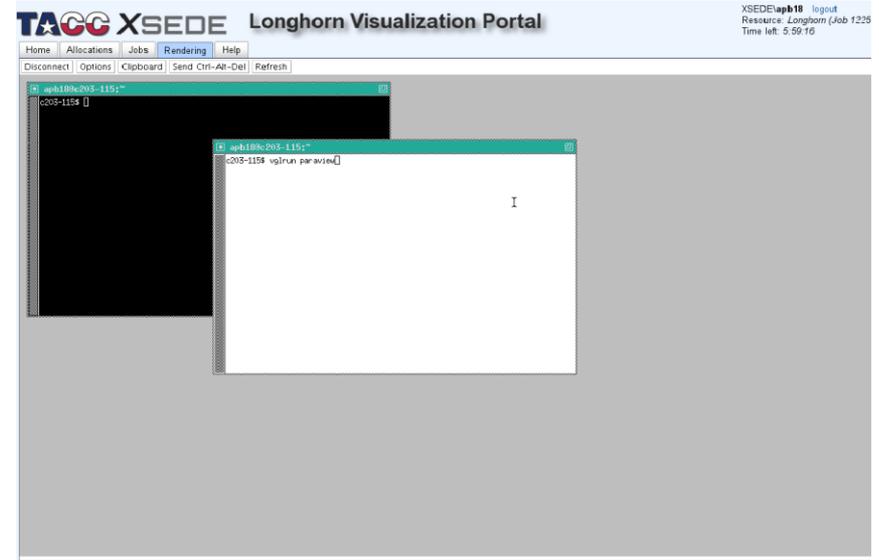
- You will be brought to the VNC desktop.
- In the white xterm, type  
`module load python  
paraview`
- You will receive no visual feedback or return value
- This makes paraview available on the commandline.





## Run ParaView

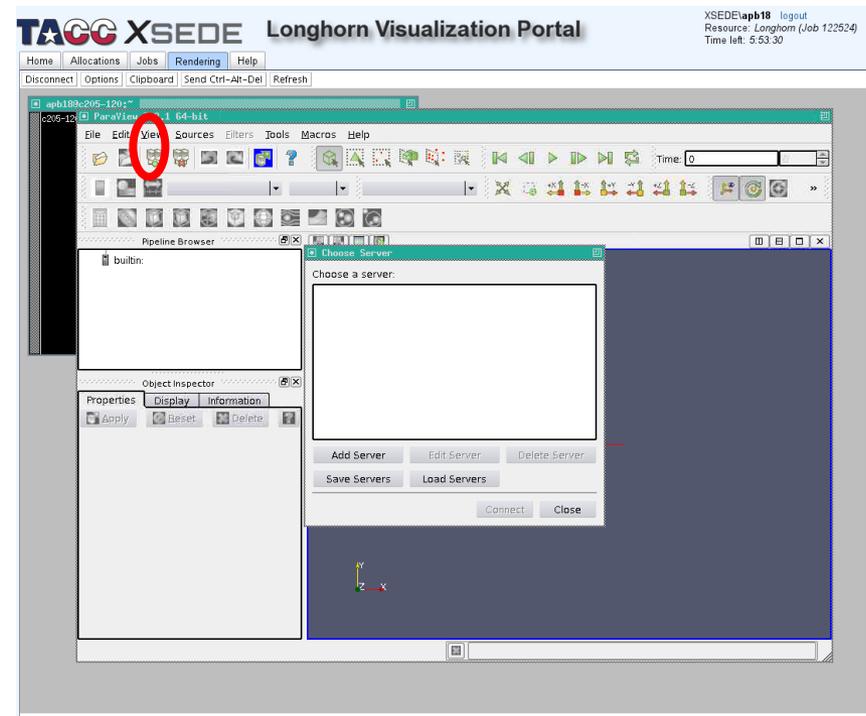
- Type `vglrun paraview`
- This will start paraview and allow it to use the system GPU for rendering
- By default, it does not start up in parallel mode. That has to be enabled next.





## Run ParaView

- Click the “connect” button (or choose file->connect)
- This will bring up a “choose a server” dialog.
- Select “add server”
- Paraview can launch backend parallel processes for us and connect to them, but we need to configure it to do so.

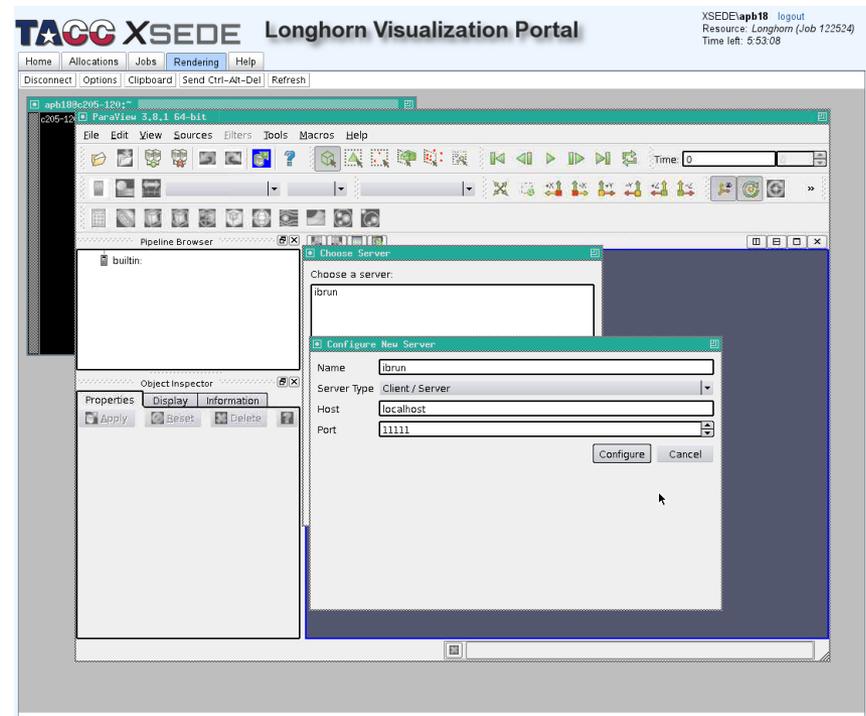




## Run ParaView

- Once you click “add”, a “configure new server” dialog will pop up.
- Give it this configuration a name (e.g. “ibrun”) in the “name” field
- Leave the other fields alone
  - Server Type: Client/Server
  - Host: Localhost
  - Port: 11111
- Click “configure”

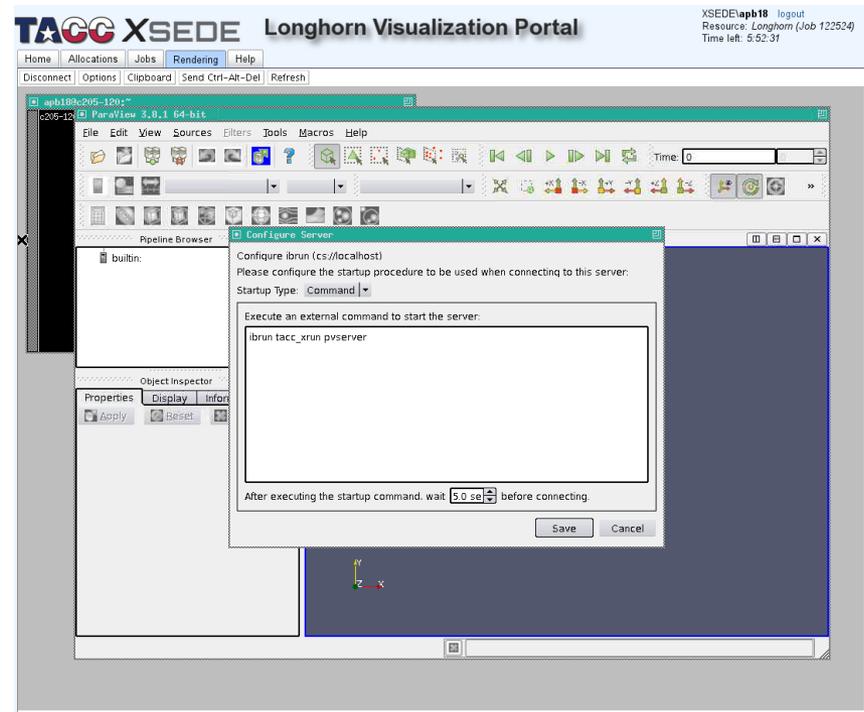
1/20/2012





## Run ParaView

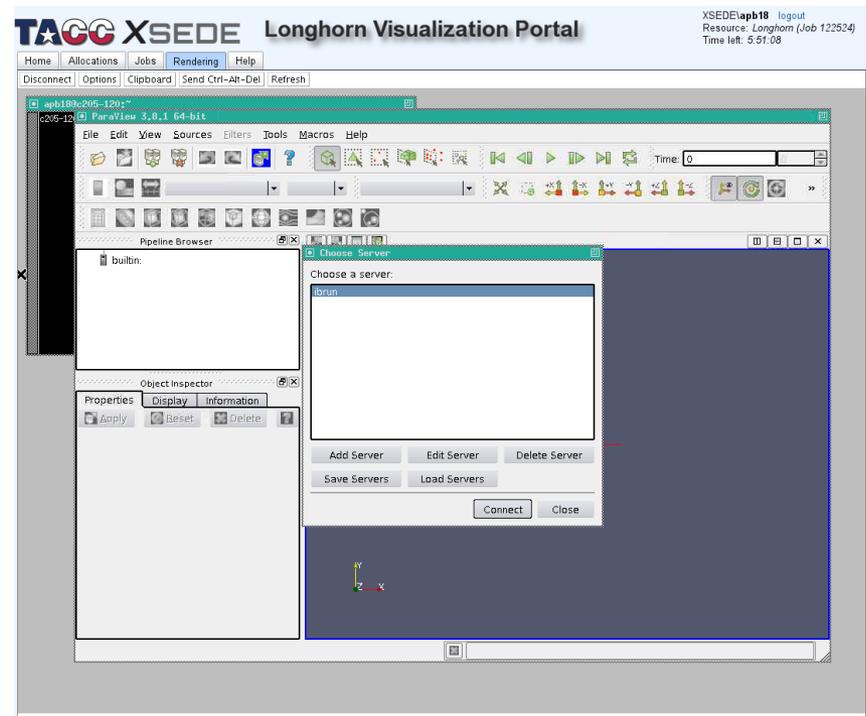
- A “configure server” dialog will open
- Make sure Startup Type: Command is selected
- In the large text box, type in “env NO\_HOSTSORT=1  
ibrun tacc\_xrun  
pvserver”
  - This is the command that paraview will use to launch parallel backends
- Click “save”





## Run ParaView

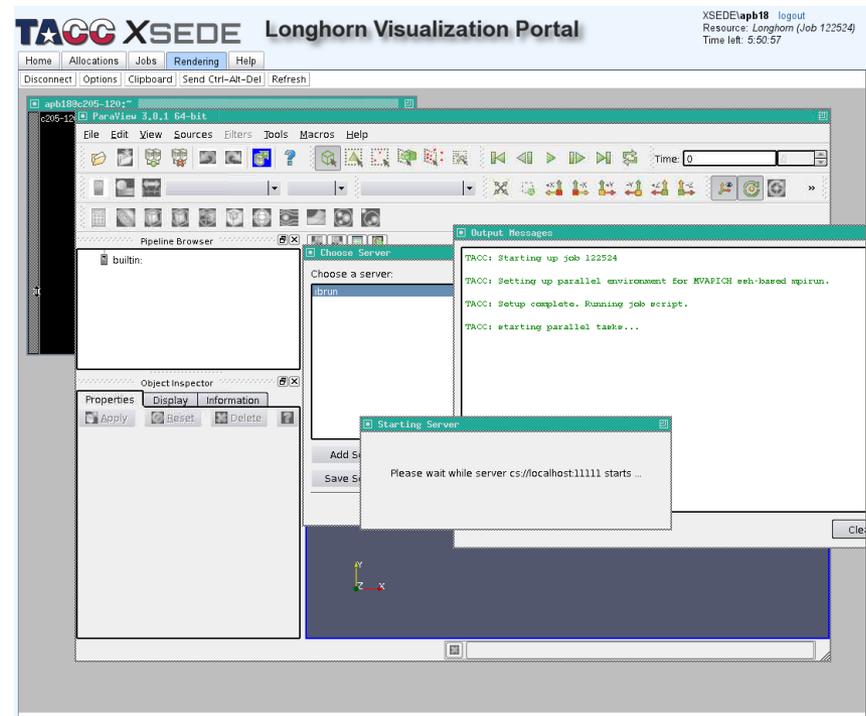
- Select “Save Servers” to have paraview remember these settings
  - Otherwise, you’ll need to type this information every time you restart paraview!
- It will prompt you for a filename. Just type in something like “servers” and click OK
- Click “connect” to have paraview launch the parallel backends.





## Run ParaView

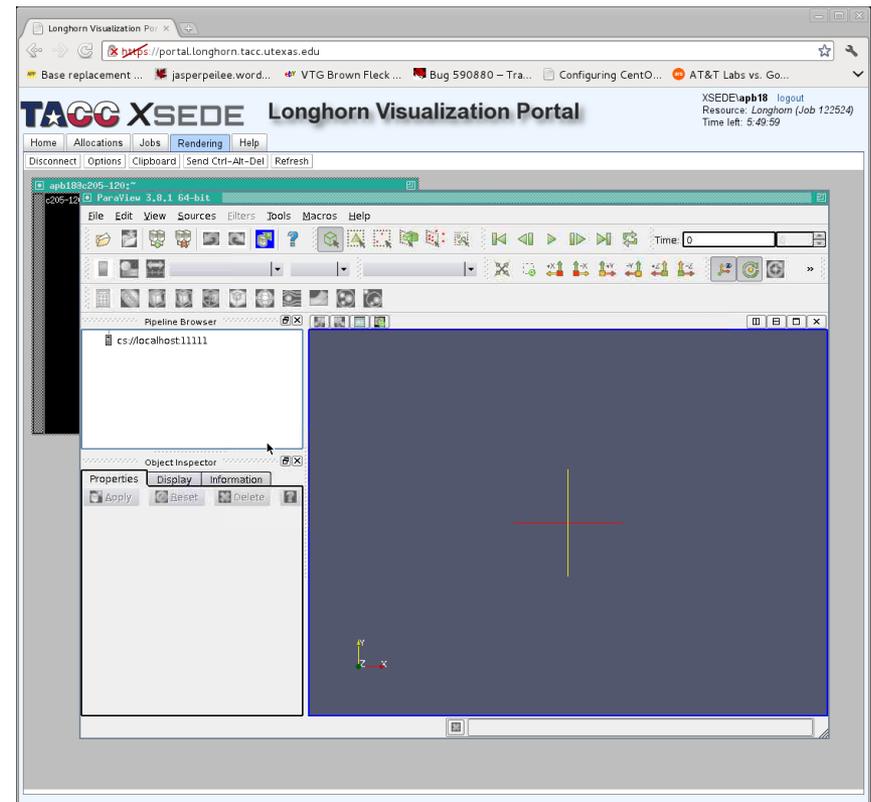
- You will see several windows pop up that report status.
- Once all backends have been launched, you can close the status dialog window if it's in the way.
- The number and location of backends is automatically determined by your initial settings from submitting the VNC job.





## Parallel

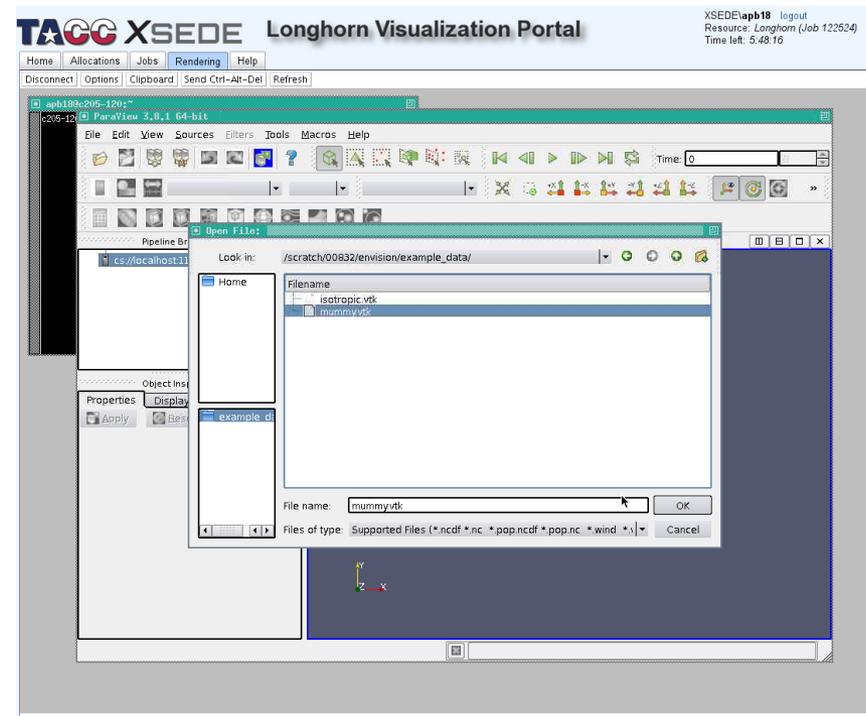
- ParaView is now ready to use. It looks no different from “normal”
- The only visible difference: the pipeline browser now starts with “cs://localhost:11111”
- You can proceed as usual, ParaView will automatically use all available resources in parallel.





## Parallel

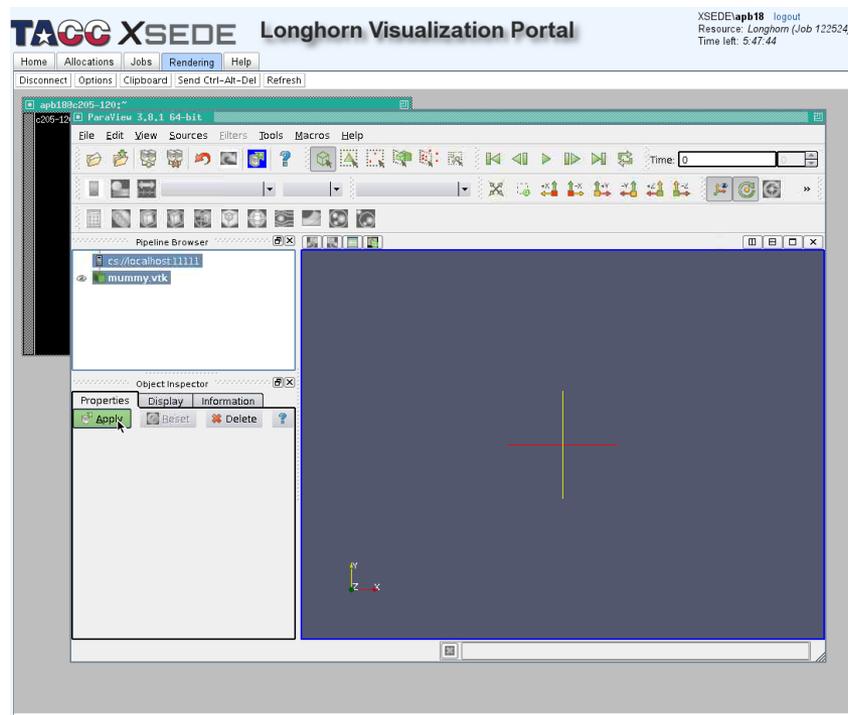
- Select File->Open to load the example dataset
- Navigate to `/scratch/00832/envision/example_data/`
- Load the mummy dataset: `mummy.vtk`





## Parallel

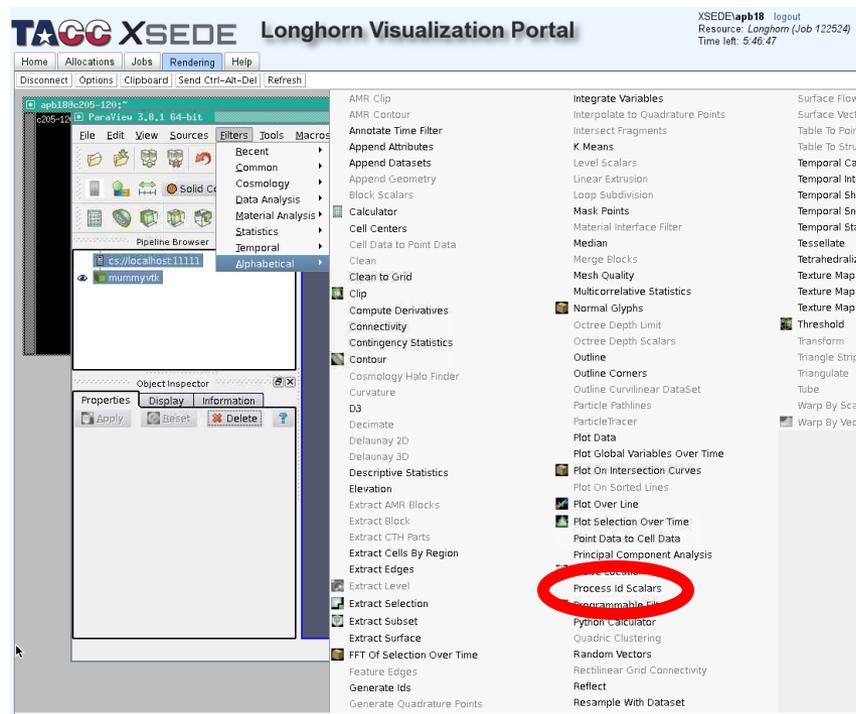
- Once loaded, click “apply” in the “properties” tab of the object inspector
- We will now add a filter which augments this data by adding another variable representing backend process ID for each point. This will indicate which backend is processing which portion of the data.





## Parallel

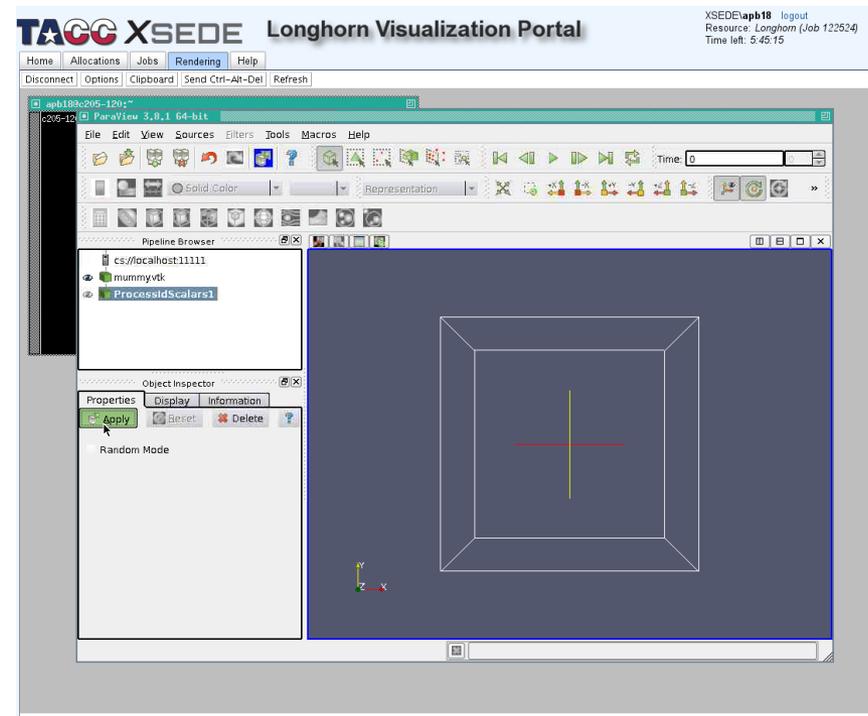
- In the “Filters” menu, go down to “Alphabetical”
- This will open a large menu showing all available filters. Navigate to the “Process Id Scalars” filter and choose it.





## Parallel

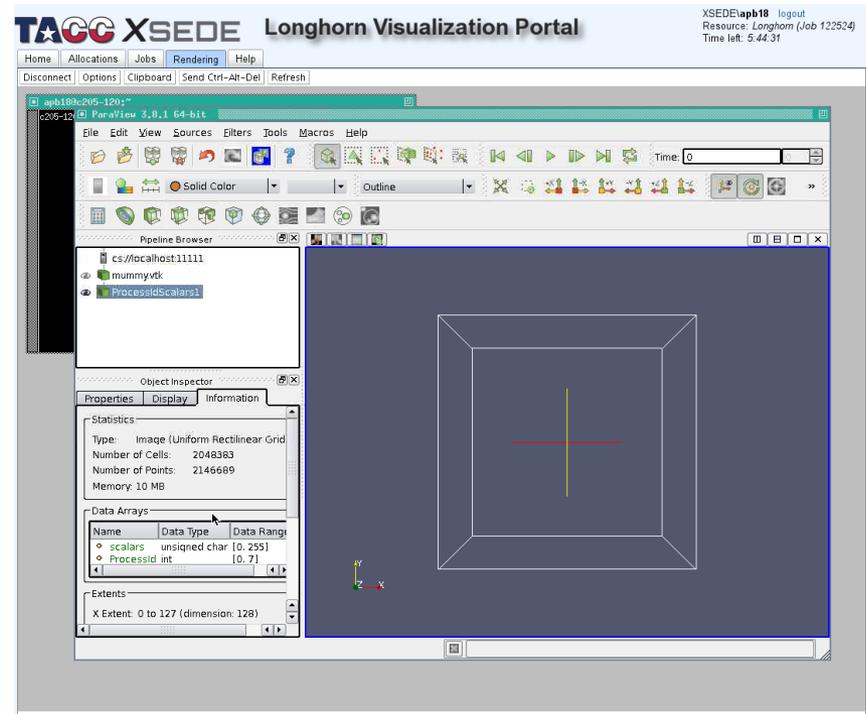
- Click “apply” in the properties tab of the object inspector for the ProcessID Scalars filter.





## Parallel

- Go to the “information” tab of the object inspector for the ProcessID Scalars filter. Scroll down until you can see the “Data Arrays” section.
- Observe that there is a new array named “ProcessId” containing integers ranging from 0-7.
  - These values map to our eight backend servers.

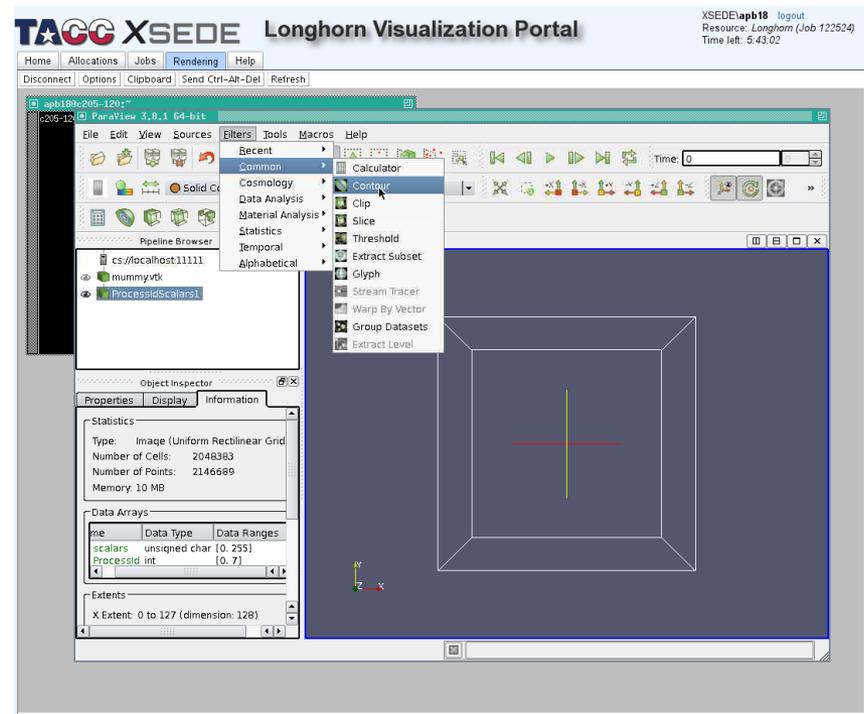




## Parallel

- Add and apply a contour filter via Filters -> Common -> Contour
- We will use this to produce an isosurface of the mummy skull where the value is 128.
- We will color this surface by ProcessId
- Be sure to click “Apply” after selecting the Contour filter.

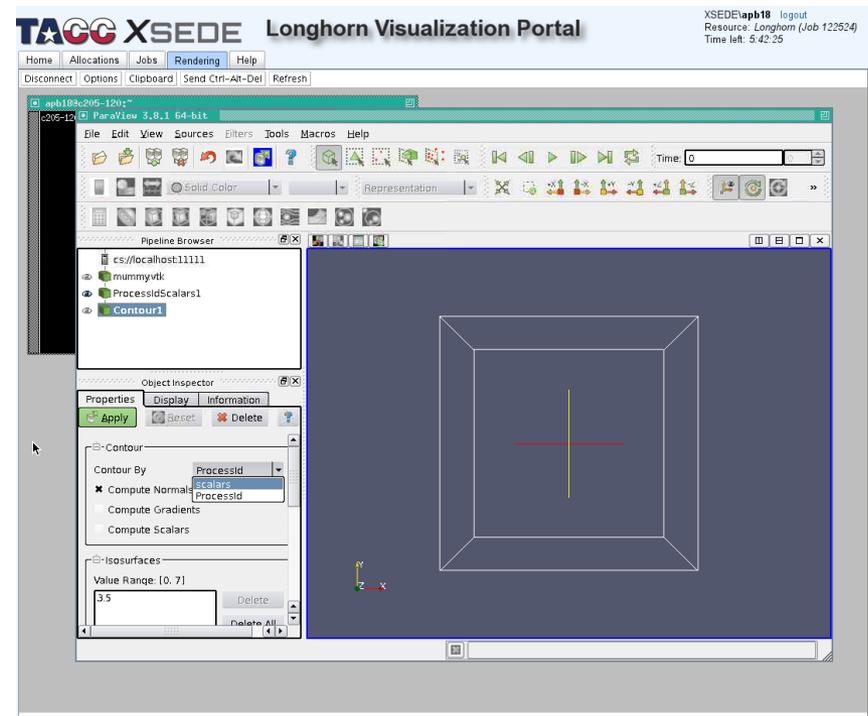
1/20/2012





## Parallel

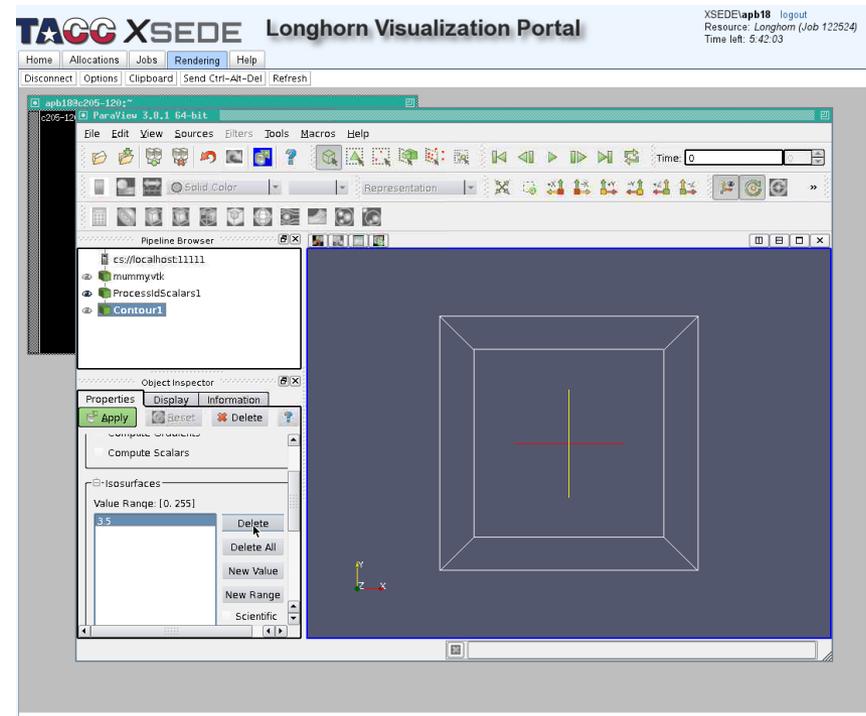
- In the “properties” tab of the object inspector for the contour filter, find the “contour section”
- The value will likely be “ProcessId” change it to “scalars”
  - We want to use the “scalars” data to produce the surface.





## Parallel

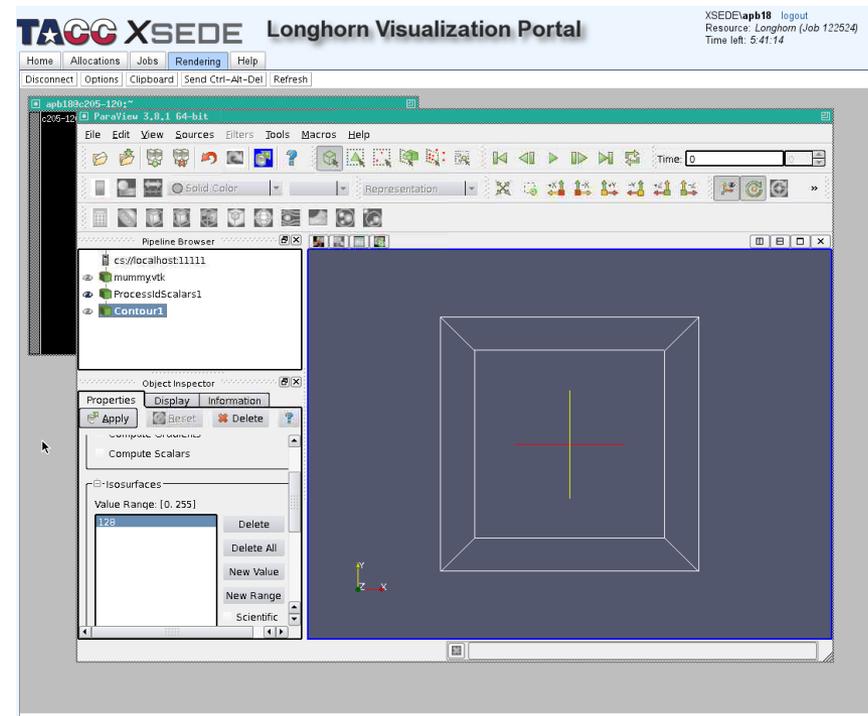
- Still in the “properties” tab, scroll down to the “isosurfaces” section. Delete the existing value (probably) 3.5, and add a new value of 128.





## Parallel

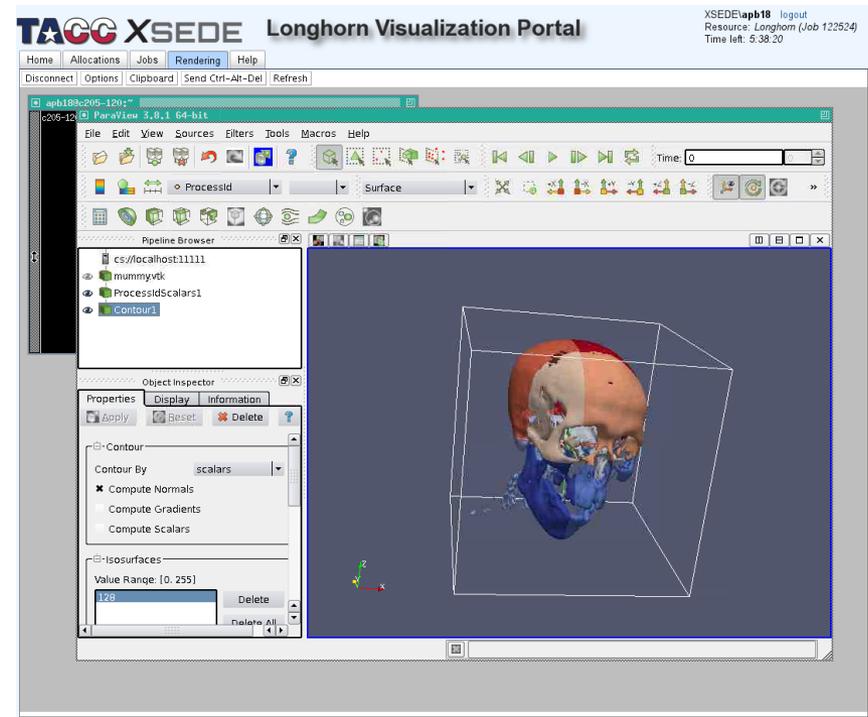
- When finished, you should have a single value of “128” selected.
- Click “Apply”.
- Go to the “Display” tab and in the “color” section, verify that “ProcessId” is selected in the “Color by” selector.





## Parallel

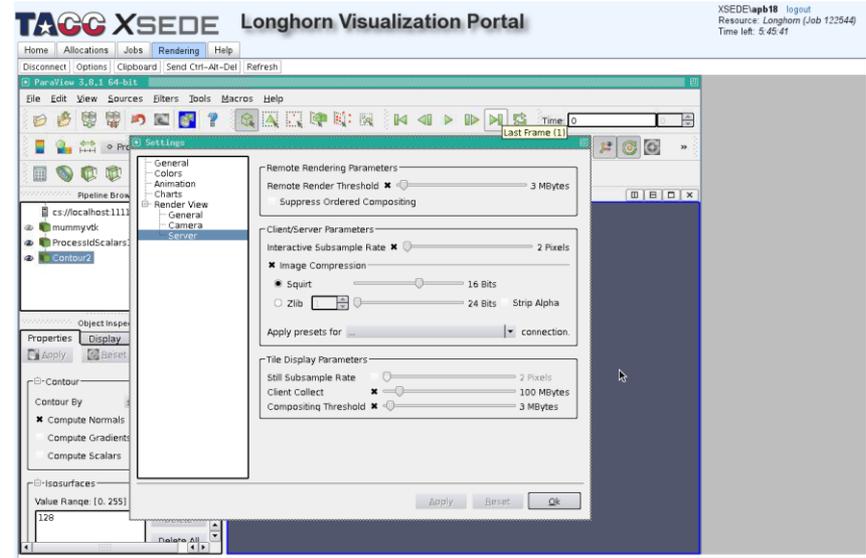
- You should now see a rendering where the contour surface is determined by the original data, but colored by process ID.
- Since we have 2 nodes at 4 processes per node, we should see 8 distinct colors. This is showing data parallelism.





## Parallel

- ParaView will automatically determine if backends perform data processing only (sending triangles for client to render), or perform rendering as well (send the pixels to directly display).
- This can be tweaked via Edit->Settings->Render View->Server





## Experiment!

- Try applying the tetrahedralize filter, viewing by volume, looking at memory usage, tweaking remote rendering threshold in settings
- Load the other dataset in the examples directory (isotropic.vtk), try plotting streamlines and coloring by process Ids
- Try running VisIt in parallel mode.