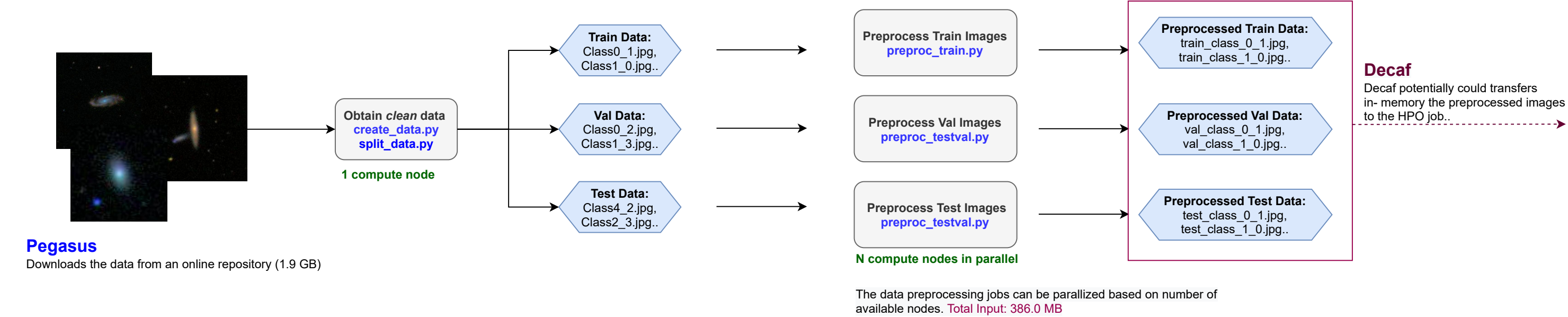


<b>1. Data Acquisition</b> Data is downloaded and unzipped in cwd	<b>2. Clean Data</b> Reduce to 28790 images. 8434, 8069,578, 3903, 7806	<b>3. Data Preprocessing</b> Separate jobs for train data, and test/val data
<b>Pegasus</b>		

Download **galaxy-zoo-the-galaxy-challenge.zip**  
File size: 1.9 GB

The images are assigned to classes based on their properties. Subset of the images is used as a dataset going forward. Total 28,793 images each between 8 kB to 16 kB. **Total: 386.0 MB**

**Pegasus**  
launches multiple data preprocessing jobs in parallel.



<b>4. Hyperparameter Optimization</b> Tune model to find the best setting (Optimizer, Activation)	<b>5. Train Model</b> Use the best optimization parameters to train model	<b>6. Test Model</b> 7 classification metrics, visualize what network learned
<b>Pegasus / Decaf</b>		

