

# Einstein Toolkit School

## Activity Log

- 7/18: received notice that [Einstein Toolkit School and Workshop](#) hosted at NCSA was looking for a solution to host a tutorial for the Einstein Toolkit using Jupyter notebooks. They were exploring Jupyterhub, but also open to considering Workbench
- 7/20: Met with organizer to discuss potential requirements and demonstrate what we've done for PI4 workshop. Their original plan to use the Blue Waters training allocation wasn't ideal for a short workshop for ET (requires students to use SSH, system editors; can't easily run Jupyter; IO is slow for compilation; must submit jobs, etc.
- 7/21: Gave access to existing pi4 instance with basic Jupyter notebook with ET dependencies.
  - Response: IO is slow (compiling comparable to BW) due to use of Gluster FS. Much slower than single-node Jupyterhub. Will either use BW (if issues solved) or Jupyterhub.
- 7/21: Setup single-node instance of Workbench (32core/96G) on Nebula for performance comparison.
- 7/24: Received message from co-organizer requesting changes to the Docker image
  - Default python2 interpreter for terminal, but support python3 notebook.
  - Fix RequestEntityTooLarge error (nginx max body size)
  - Question about whether all students would be on a single VM or multiple
  - Additional packages: numactl-devel numactl hwloc hwloc-devel openssl-devel hdf5 hdf5-devel gdb
- 7/25: Requested gdb and gsl
- 7/26: Requested 64core VM
- 7/28: Instructors began testing whether the image/environment works for their tutorials
  - Requested additional dependencies in image
- 7/31:
  - 9am: School starts
    - 30+ people in room
    - ~4 issues with Safari – but this is with Jupyter, not Workbench. Couldn't get terminal and invalid kernel
    - Issues cloning git repo (usually typos)
    - Note: Lots of tension when things don't work as expected.
    - Note: Logging in twice is annoying for most users
    - 10am. Back to the office
  - 11am: All's quiet. ~36 jupyter instances running on single node. Load average ~1.
  - noon: Morning went ok aside from initial problem with Safari and work directory
    - Spoke with instructor and got a little background. They run a server at LSU to let people run the ET tutorials. Users sign up and have access for a month or less. For whatever reason, they can't give out accounts now. He's wondering if they can "use Nebula".
    - This would mean letting users register to run Jupyter notebook tutorials, which could already be supported by the Workbench beta.
  - 1:30pm: Major problems
    - All students began the "Using Cactus" tutorial <https://github.com/stevenrbrandt/CactusTutorial>, which includes git cloning and compiling a subset of the Einstein Toolkit
    - Load averages were very high (approaching 60 for a 64 core system). This cause unpredictability in the Kubernetes controller, since it's all on one node.
    - Around 1:45, we noticed that the root partition was nearing 100%. This was caused by /var/lib/docker filling up. Apparently the Notebook was compiling in ~/ instead of in the work subdirectory (which had a volume mount).
    - We were unable to quickly move the /var/lib/docker to a new volume
      - Data transfer between volumes on Nebula was astoundingly slow. Nebula team thinks this might be an FS issue.
      - Docker overlay mounts hung around even after shutting down.
    - After ~1 hour, the instructor decided to use their own JupyterHub instance.
    - We transitioned to a new instance with a much larger Docker volume.
- 8/1
  - 9am:
    - Organizer emailed to say they are considering using Workbench for a session this evening.

## Custom Docker image

- <https://github.com/craig-willis/jupyter-et>

## Requirements

- 15-50 accounts (encourage groups of 3)
- Custom Jupyter notebook with ET dependencies
- 2 core, 2GB RAM, 5GB disk per user
- Pre-loaded data
- Workload:
  - [https://docs.einsteintoolkit.org/et-docs/Simplified\\_Tutorial\\_for\\_New\\_Users](https://docs.einsteintoolkit.org/et-docs/Simplified_Tutorial_for_New_Users)