

ServiceX Dashboard for Jupyterhub



Jake Li

University of Illinois at Urbana-Champaign

Mentors: Oksana Shadura, Gordon Watts, Benjamin

Galewsky, Alexander Held, Mason Proffitt

Background

- ServiceX is a data extraction and delivery service.
- Users input datasets with specific instructions and ServiceX, with the given framework, extracts the data and stores it in an easily analyzable format.
- ServiceX can be co-located with datasets to provide fast and efficient data reduction strategies, provide data as awkward arrays, and work as a data backend to Coffea and TRexFitter analysis tools.

Running a ServiceX notebook in coffea-casa Analysis Facility

The screenshot displays a JupyterLab environment with a notebook titled 'coffea.ipynb'. The notebook contains Python code for a ServiceX pipeline. The code defines a function to produce histograms and a pipeline named 'servicex_databinder'. The pipeline is executed, and the progress is shown as a series of horizontal bars representing the download status of different files. The progress bars are labeled with URLs and download percentages, such as '40%', '50%', '60%', '80%', and '10%'. The progress bars are arranged in a roughly descending order of completion time, with the top bar being the most complete and the bottom bar being the least complete.

```
# as a script:
# async def produce_all_the_histograms(fileset):
#     return await utils.produce_all_histograms(fileset, use_dask=False)
# all_histograms = asyncio.run(produce_all_the_histograms(fileset))

elif PIPELINE == "servicex_databinder":
    # needs a slightly different schema, not currently implemented
    raise NotImplementedError("further processing of this method is not currently implemented")

print(f"Execution took {time.time() - t0:.2f} seconds")
```

Progress bars (from top to bottom):

- [https://kroo... 40%] 4/10 [00:34]
- [https://kroo... Downloaded: 50%] 5/10 [00:34]
- [https://kroo... 60%] 6/10 [00:35]
- [https://kroo... Downloaded: 60%] 6/10 [00:35]
- [https://kroo... 80%] 8/10 [00:30]
- [https://kroo... Downloaded: 80%] 8/10 [00:30]
- [https://kroo... 10%] 1/10 [00:32]
- [https://kroo... Downloaded: 10%] 1/10 [00:32]
- [https://kroo... 20%] 2/10 [00:29]
- [https://kroo... Downloaded: 20%] 2/10 [00:29]
- [https://kroo... 20%] 2/10 [00:33]
- [https://kroo... Downloaded: 20%] 2/10 [00:33]
- [https://kroo... 40%] 4/10 [00:31]
- [https://kroo... Downloaded: 50%] 5/10 [00:31]
- [https://kroo... 10%] 1/10 [00:28]
- [https://kroo... Downloaded: 10%] 1/10 [00:28]
- [https://kroo... 20%] 2/10 [00:29]
- [https://kroo... Downloaded: 20%] 2/10 [00:29]

ServiceX JH plugin

Dask JH plugin

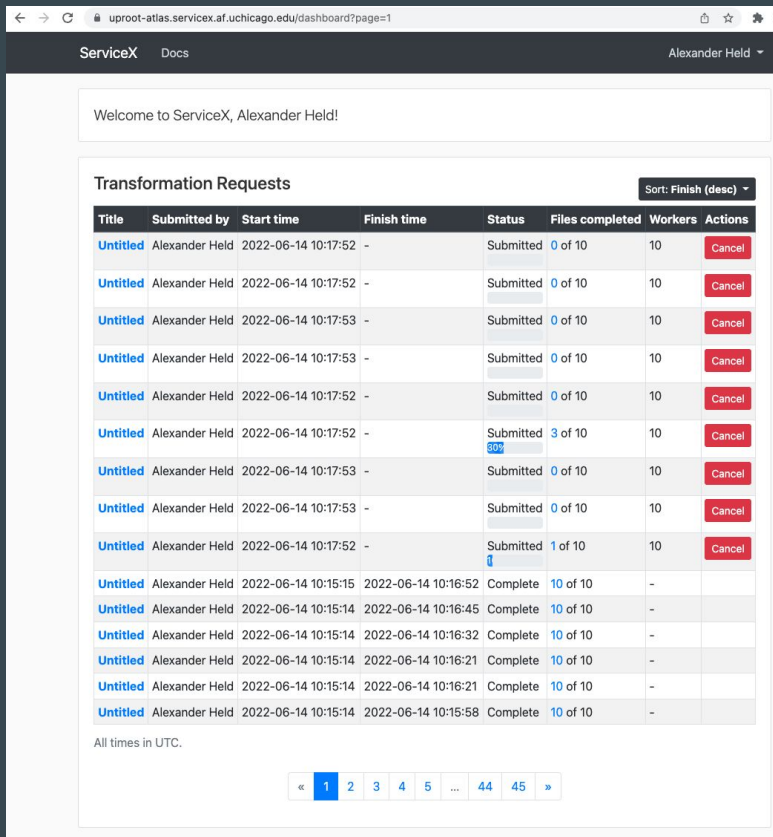
CLUSTERS + NEW

SSL AF HTCondor Cluster
Scheduler Address: tis://alexander-2held-40cern-2ech.dask.casa.
Dashboard URL: /User/alexander.held@cern.ch/proxy/8787/status
Number of Cores: 4
Memory: 12.72 GiB
Number of Workers: 2

Project Goals

- The primary goal of the project is to create a Jupyterhub plugin that will give users access to a ServiceX dashboard.
- The plugin will be able to provide the ability to monitor all user transforms and give dashboard-like functionality that exists on the main ServiceX website.
- A secondary goal of the project is to add more existing features to the dashboard.

ServiceX dashboard hosted on separate address (not integrated in Analysis Facility and requires Globus authentication)



The screenshot shows a web browser window with the URL `uprooot-atlas.servicex.af.uchicago.edu/dashboard?page=1`. The page header includes "ServiceX Docs" and the user name "Alexander Held". A welcome message reads "Welcome to ServiceX, Alexander Held!". Below this is a section titled "Transformation Requests" with a sort dropdown set to "Sort: Finish (desc)". A table lists 15 requests with columns for Title, Submitted by, Start time, Finish time, Status, Files completed, Workers, and Actions. The first 10 requests are "Submitted" with 0 files completed, while the last 5 are "Complete". A pagination bar at the bottom shows page 1 of 45.

Welcome to ServiceX, Alexander Held!

Transformation Requests

Sort: Finish (desc)

Title	Submitted by	Start time	Finish time	Status	Files completed	Workers	Actions
Untitled	Alexander Held	2022-06-14 10:17:52	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:52	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:53	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:53	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:52	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:52	-	Submitted	3 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:53	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:53	-	Submitted	0 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:17:52	-	Submitted	1 of 10	10	Cancel
Untitled	Alexander Held	2022-06-14 10:15:15	2022-06-14 10:16:52	Complete	10 of 10	-	
Untitled	Alexander Held	2022-06-14 10:15:14	2022-06-14 10:16:45	Complete	10 of 10	-	
Untitled	Alexander Held	2022-06-14 10:15:14	2022-06-14 10:16:32	Complete	10 of 10	-	
Untitled	Alexander Held	2022-06-14 10:15:14	2022-06-14 10:16:21	Complete	10 of 10	-	
Untitled	Alexander Held	2022-06-14 10:15:14	2022-06-14 10:16:21	Complete	10 of 10	-	
Untitled	Alexander Held	2022-06-14 10:15:14	2022-06-14 10:15:58	Complete	10 of 10	-	

All times in UTC.

« 1 2 3 4 5 ... 44 45 »