High-Throughput Computing in Python, Powered by HTCondor

Josh Karpel
What happens to Y as we change X?

500 Simulations
1 Hour Each
What happens to Y as we change X?

Setup

X1 → Simulation → Y1
X2 → Simulation → Y2
X3 → Simulation → Y3

Analysis
\[ X = [0, 1, 2, \ldots] \]

\[ Y = \text{map}(\text{simulation}, X) \]

\[ \text{make_plot}(X, Y) \]
HTMap is a Pure-Python, Seamless, Lightweight, Focused HTCondor Job Submitter & Manager
Who is HTMap for?

**Users**

- Has a working computation written in Python
- Knows Scientific Python, but not necessarily anything else
- **Wants to use high-throughput computing... but is very busy**

**Facilitators**

- “Well, it works on my laptop...”
- Users have **similar workflows**, but **dissimilar skills**
Who is HTMap **NOT** for?

- “The output of my analysis is a file”
- “My function takes a millisecond to run”
- “But I don’t use Python...”
HTMap - HTCondor Week 2019

- Independent Execution
  - multiprocessing
  - concurrent.futures
- Short-Running
  - dask
- Long-Running
  - hmap
  - condor_submit
  - condor_submit_dag

- Interdependent Execution
HTMap
https://github.com/htcondor/htmap

Slides and Code
https://github.com/JoshKarpel/htcondor-week-htmap-talk