Let’s Build a Python Profiler in 25 LOC

Noam Elfanbaum
About me

- Data Engineering lead at BlueVine
- Involved with PywebIL & Pycon Israel
- Find me at @noamelf everywhere or at noamelf.com
Python Profiling
What is profiling?

A profile is a set of statistics that describes how often and for how long various parts of the program executed.
Python has 2 built-in profilers in stdlib:
- **profile** - an early pure Python implementation
- **cProfile** - a C extended profiler for better performance
How do profilers work?

- If you’re inside a Python program you generally have pretty easy access to its stack.
- Most profilers run inside your Python process.
Types of profilers

There are two types of profilers that differ upon their triggers:

- Deterministic profilers - triggered on function/line called (like profile/cProfile)
- Statistical profilers - triggered on a time interval
How do deterministic profilers work?

- Python lets you specify a callback that gets run when interpreter events happen:
  - `sys.setprofile` - triggered when a function or a line of code is called
  - `sys.settrace` - triggered only when a function is called
- When the callback gets called, it records the stack for later analysis.
Disadvantage of Deterministic profilers

- Introduces a fixed amount of latency for every function call / line of code executed.
- Standard programs do not have so many function calls.
- "The interpreted nature of Python tends to add so much overhead to execution, that deterministic profiling tends to only add small processing overhead in typical applications"
How do statistical profilers work?

- Statistical profilers sample the program on a given interval.
- It has less overhead than deterministic profiler, but is also less accurate.
- One way to implement the sampling is to ask the OS kernel to interrupt the program on a given interval.
Now, let’s build a (naive) statistical profiler in 25 LOC!
- Talk code repository
- Nylas blog post - [https://www.nylas.com/blog/performance/](https://www.nylas.com/blog/performance/)
- Flame Graph by Brendan Gregg [https://github.com/brendangregg/FlameGraph](https://github.com/brendangregg/FlameGraph)
- Python documentation about profilers -
  - [https://docs.python.org/3/library/profile.html](https://docs.python.org/3/library/profile.html)
  - [https://docs.python.org/3/library/sys.html#sys.setprofile](https://docs.python.org/3/library/sys.html#sys.setprofile)
Thanks!