

# PyDarshan interface and tools



Nik Awtrey  
Los Alamos National Laboratory

ECP Community BoF Days  
May 12, 2022

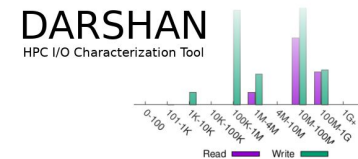
# PyDarshan: Overview



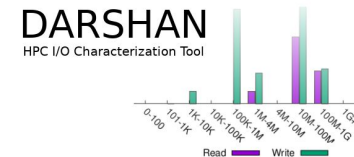
- Python-based Darshan log analysis tool
- Primary functions:
  1. Provide API for accessing Darshan log data
  2. Generate summary reports (HTML)
- Motivations:
  - Easy-to-install dependencies
  - Lower the bar for maintenance, custom analysis, new features, etc.
- Installation:
  - Python 3.6 through 3.10
  - Can be installed from PyPi (pip install --user darshan)
  - Source packaged in **darshan-util**

# PyDarshan API & Data Access

- Jupyter-notebook



# PyDarshan Job Summary Tool



- Generating a report:

```
$ python -m darshan summary ~/path/to/log_file.darshan  
> Report generated successfully.  
> Saving report at location: /current/working/directory/log_file_report.html
```

# PyDarshan: Feedback



We want your feedback!

Feedback is welcome on the Darshan issue tracker: <https://github.com/darshan-hpc/darshan/issues>

Most up-to-date PyDarshan reports can be generated using Darshan **v3.4.0-pre1release**, or the current **main** branch.

*This research was supported by the Exascale Computing Project (17-SC-20-SC), a joint project of the U.S. Department of Energy's Office of Science and National Nuclear Security Administration, responsible for delivering a capable exascale ecosystem, including software, applications, and hardware technology, to support the nation's exascale computing imperative.*