Sandstone HPC

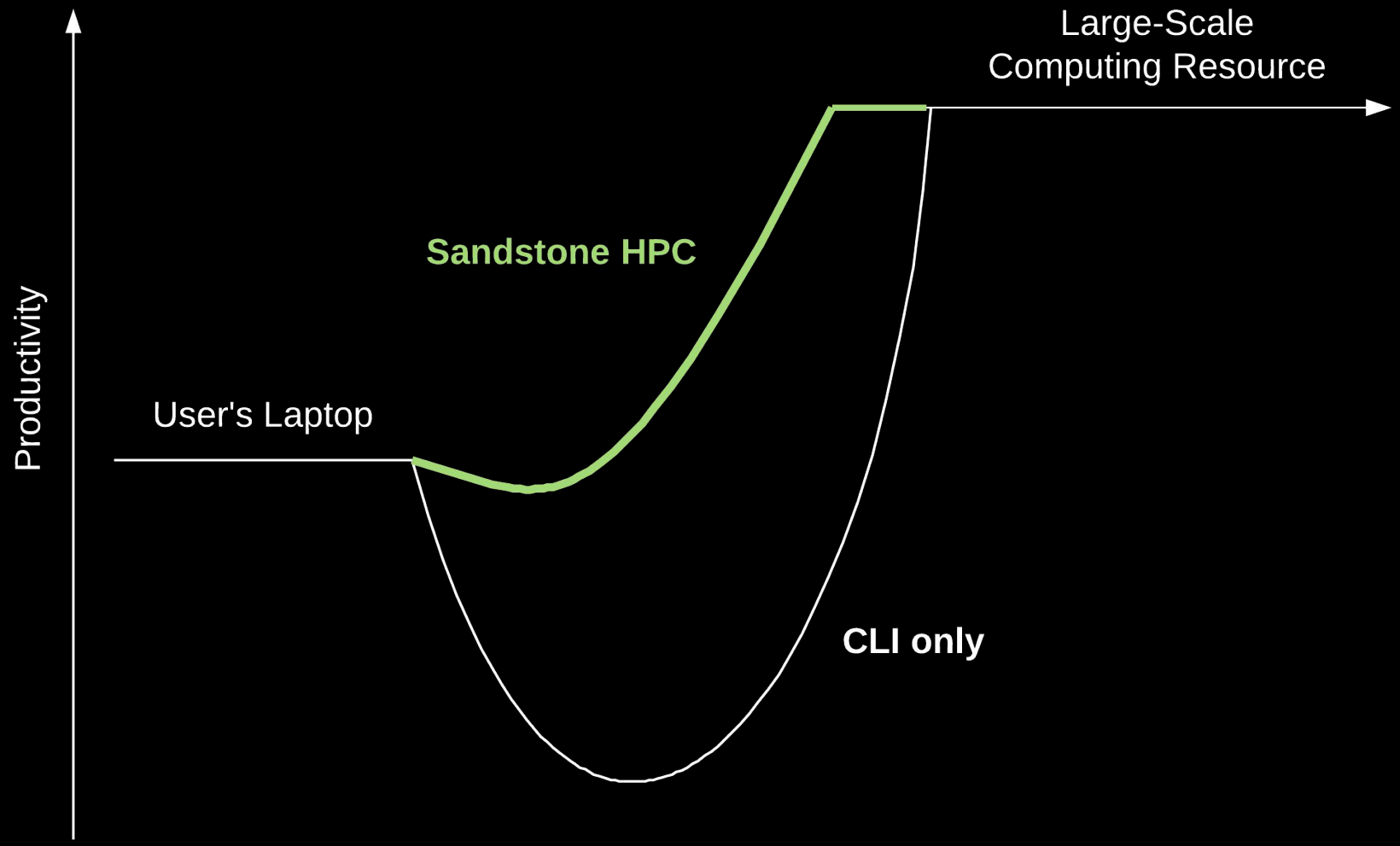
A Domain-General Gateway for New Users

Zebula Sampedro
sampedro@colorado.edu
sandstonehpc.github.io

Research Computing
UNIVERSITY OF COLORADO **BOULDER**

Sandstone HPC *is a modular, consolidated set of web tools that make it easier to access and utilize HPC resources.*

Minimizing the “Valley of Despair”



Mechanism

- Makes existing tools easier to use.
  - But still allow direct use of the underlying system
  - Reinforce the CLI workflow
  - Minimize up-front knowledge needed
  - Provide clear feedback
- Built on the model of a Cloud IDE
  - Proven pattern for accessing remote resource
  - Reduced emphasis on development
  - Instead focus on common painpoints

Targeting Common Painpoints

- Editing files on the remote system
- File transfer and navigation
- Command line
- Creating and scheduling job scripts
- Recovering from system errors

Sandstone HPC components

- Code Editor
  - Facilitate graphical file editing
- Filebrowser
  - Intuitive pattern for uploading, managing, and viewing files.
- Web Terminal
  - No client-side configuration, just need your credentials.
- Slurm Assist
  - Interactive utility for building sbatch scripts with real-time validation.

Effectiveness

- Usability Study
  - Slurm Assist module
  - Novice to intermediate HPC experience
  - Performed sets of tasks on Sandstone HPC and the CLI
    - Create a job script
    - Schedule the job given a plain-language description of its characteristics
    - Explain and correct errors

Effectiveness

- When using Sandstone
  - Reduced error rate (-81%)
  - Increased success rate (+44%)
  - Minimal time spent reading docs
  - Every participant correctly identified errors
  - 100% would rather use Sandstone

Architecture

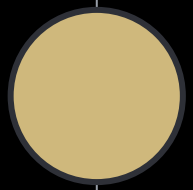
- Self-contained Python web app
  - Tornado backend
  - REST API
  - AngularJS SPA
  - Minimal footprint and dependencies
    - Linux/UNIX
    - pip2 (for now), pip3
  - Runs as a user command
  - PAM auth and SSL if used standalone
  - Deployable behind JupyterHub

Architecture

- Apps are just Python modules
  - Contain server and client content
  - Loaded based on config during bootstrap
  - Pip-installable
  - Easy to modify and extend
    - We've had great contributions already!

A Platform for HPC Education

- Focus on tutorial content
  - A surprising amount of time spent on basic setup and client config
  - Zero configuration for the end-user
  - Allows for direct use of underlying system
  - No need for terminal editors



Demo

Basic scheduling with Sandstone HPC

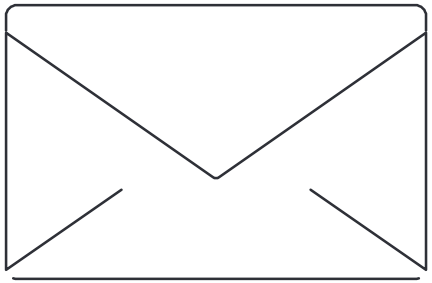
Future

- Building upon successes
  - Scheduling validation and script-building
  - Access pattern
  - Platform for tutorials and education
  - Non-exclusivity with underlying system
- Expanding use cases
  - VNC
  - Globus
  - More advanced script-building
  - Environment modules
  - Containerization

Future

- Integration with JupyterLab
  - JupyterLab generalizes the use case beyond Notebooks
  - Greater possibility for extension
  - Allows us to focus on serving HPC use cases specifically

- Singularity Spawner
  - Spawn portable software stacks via JupyterHub
  - Sandstone/JupyterLab base layer



Want to try it out?

sampedro@colorado.edu

https://sandstonehpc.github.io