Enhancing Scientific Onboarding in CIRRUS: Cloud-Native Infrastructure for Earth System Scientists

Shilin Chhabra

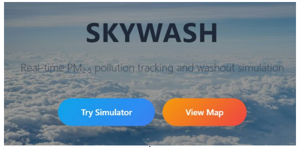
Mentors: Nick Cote, Kevin Hrpcek, Varsha Banda

Summer Internships in Parallel Computational Science (SIParCS)

July 29, 2025

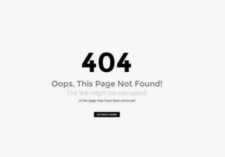
This material is based upon work supported by the NSF National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1852977.

Why CIRRUS?



SKYWASH

Real-time $PM_{2.5}$ pollution tracking and washout simulation

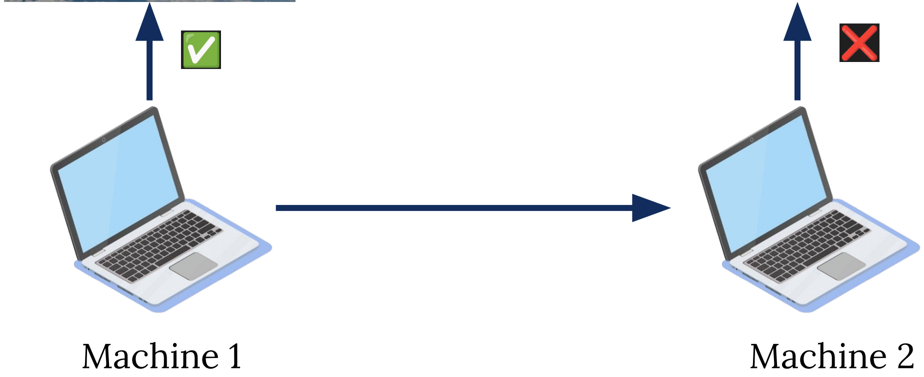


404

Oops, This Page Not Found!

The link might be corrupted.

or the page may have been removed



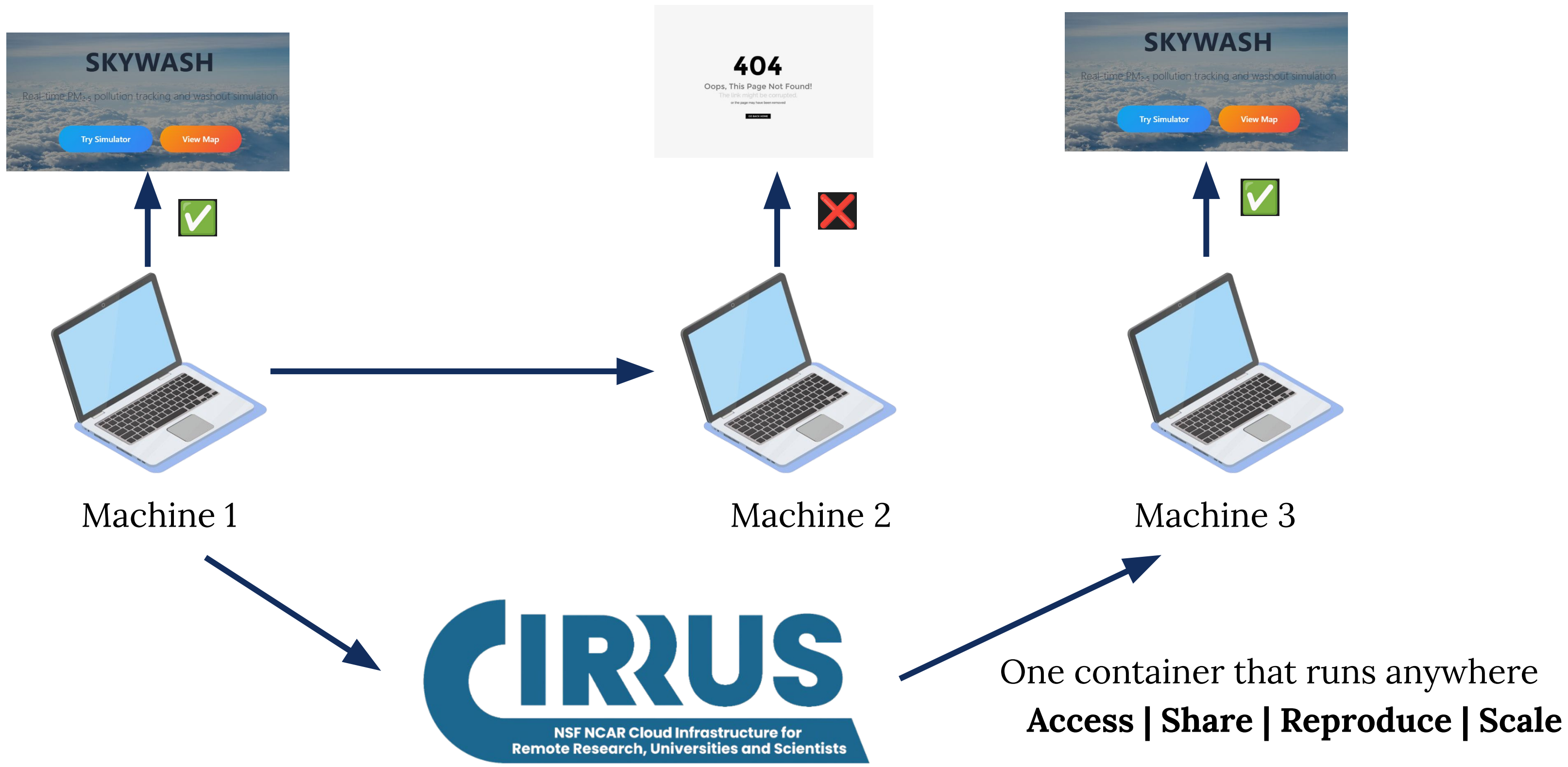
Face roadblocks:

- Library/version conflicts
- Missing GPU/compiler drivers
- OS mismatch
- Manual, unreliable installs

Hours lost rebuilding environments & debugging cryptic errors

Wish: a drop-in container, unchanged, reproducible, shareable

Why CIRRUS?



SKYWASH

Real-time $PM_{2.5}$ pollution tracking and washout simulation

Try Simulator
View Map
✓
404
Oops, This Page Not Found!
The link might be corrupted.
or the page may have been removed
GO BACK HOME
✗
SKYWASH
Real-time PM2.5 pollution tracking and washout simulation
Try Simulator
View Map
✓
Machine 1
Machine 2
Machine 3
CIRRUS
NSF NCAR Cloud Infrastructure for Remote Research, Universities and Scientists

One container that runs anywhere

Access | Share | Reproduce | Scale

Core Components of CIRRUS

Researcher does…

1) Push code to *GitHub*

(one time only)

2) Save *Harbor* robot token as repo secrets

3) Merge deploy YAML / click “Upgrade”

4) Open dashboard

CIRRUS does …

1) Builds the *Docker* image

2) Pushes image *Harbor* registry

3) *Harbor* runs a vulnerability scan; blocks critical CVEs

4) *Kubernetes* + *Argo CD* rolls out the new image and auto-scales pods

5) *OpenBao* Secret Manager injects runtime secrets

Left Column ≈ 2 min of human effort | Right column = full DevOps pipeline on autopilot

Barriers to Adoption

Lengthy narrative **documentation**

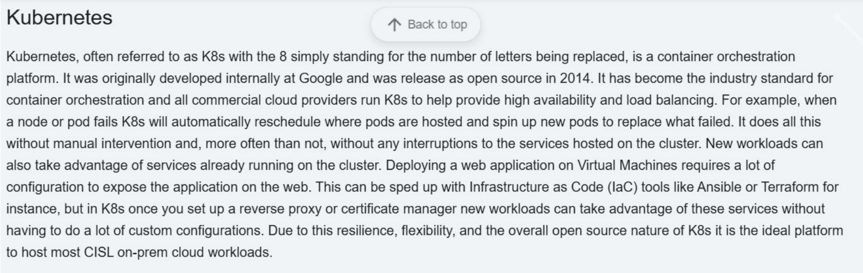
Disconnect between the **working example** and **reference docs**

Extensive **security workflow** yet unclear, leaving newcomers **uncertain**

≈ **4 hours** for a newcomer to reach a first successful deploy (not a smooth onboarding process)

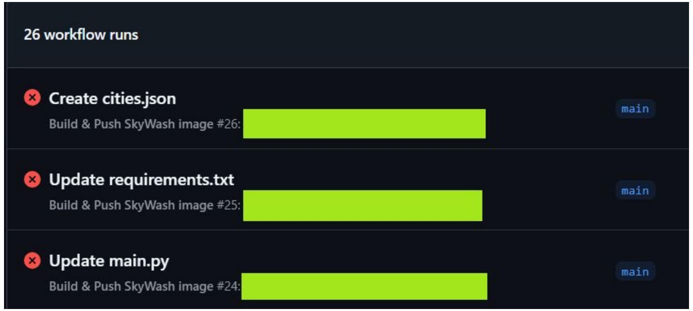
Why onboarding matters:

Quick, clear onboarding turns interest into adoption, cuts support time, and lets scientists start producing results sooner.



Kubernetes

Kubernetes, often referred to as K8s with the 8 simply standing for the number of letters being replaced, is a container orchestration platform. It was originally developed internally at Google and was release as open source in 2014. It has become the industry standard for container orchestration and all commercial cloud providers run K8s to help provide high availability and load balancing. For example, when a node or pod fails K8s will automatically reschedule where pods are hosted and spin up new pods to replace what failed. It does all this without manual intervention and, more often than not, without any interruptions to the services hosted on the cluster. New workloads can also take advantage of services already running on the cluster. Deploying a web application on Virtual Machines requires a lot of configuration to expose the application on the web. This can be sped up with Infrastructure as Code (IaC) tools like Ansible or Terraform for instance, but in K8s once you set up a reverse proxy or certificate manager new workloads can take advantage of these services without having to do a lot of custom configurations. Due to this resilience, flexibility, and the overall open source nature of K8s it is the ideal platform to host most CISL on-prem cloud workloads.



26 workflow runs

Create cities.json
Build & Push SkyWash image #26: [redacted]

Update requirements.txt
Build & Push SkyWash image #25: [redacted]

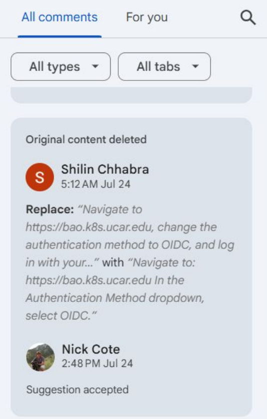
Update main.py
Build & Push SkyWash image #24: [redacted]

Approach to Improving Onboarding

Aim: *Craft the docs to be easy for a first-time user as well as detailed and coherent enough for an advanced user to make the most of the platform.*

Re-audited CIRRUS docs from a first-time user's view

Filed **600+ improvement issues** and merged suggestions weekly with mentor reviews



Original content deleted

Shilin Chhabra
5:12 AM Jul 24

Replace: *"Navigate to https://bao.k8s.ucar.edu, change the authentication method to OIDC, and log in with your..."* with *"Navigate to: https://bao.k8s.ucar.edu In the Authentication Method dropdown, select OIDC."*

Nick Cote
2:48 PM Jul 24

Suggestion accepted

Built & deployed **SkyWash**, exercising containers, CI/CD, Kubernetes, Harbor scans, and OpenBao secrets

SKYWASH

Real-time $PM_{2.5}$ pollution tracking and washout simulation

Outcome: *Estimated onboarding time* ***slashed from ~4 h to ~1 h***

Skywash Live Demo

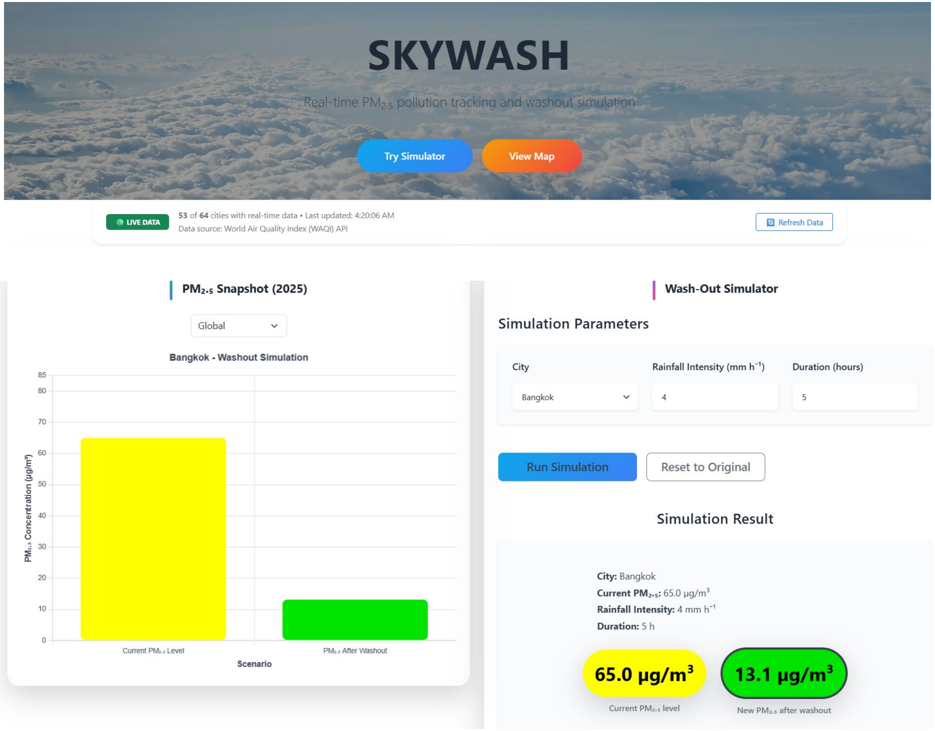
Purpose: interactive tool spotlighting air-quality challenges in South & East Asia

Wash-Out Simulator: cloud-seeding model (FastAPI) forecasts post-rain $PM_{2.5}$ / AQI

Visuals: React + Leaflet map & Chart.js bar chart

Deployed end-to-end on CIRRUS (GitHub → Harbor → K8s → Ingress)

Try it yourself: QR code to live site



SKYWASH

Real-time $PM_{2.5}$ pollution tracking and washout simulation

53 of 64 cities with real-time data • Last updated: 4:20:06 AM
Data source: World Air Quality Index (WAQI) API

$PM_{2.5}$ Snapshot (2025)

Global

Wash-Out Simulator

Simulation Parameters

City: Bangkok

Rainfall Intensity (mm h^{-1}): 4

Duration (hours): 5

Simulation Result

City: Bangkok
Current $PM_{2.5}$: 65.0 µg/m^3
Rainfall Intensity: 4 mm h^{-1}
Duration: 5 h

65.0 µg/m^3
Current $PM_{2.5}$ level

13.1 µg/m^3
New $PM_{2.5}$ after washout

Skywash Deployment Pipeline

Build multi-stage Docker image

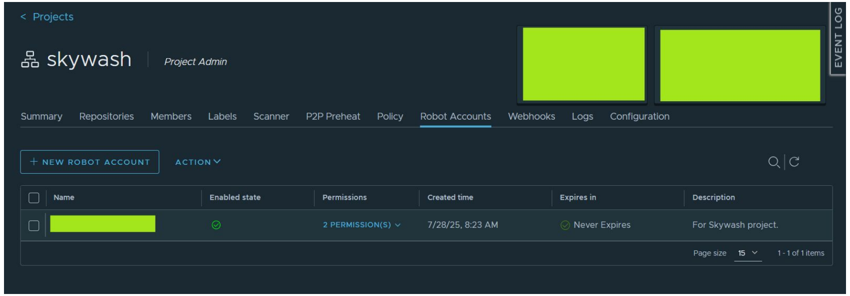
Push image the Harbor registry

Trigger Github Actions workflow

Deploy via Kubernetes
(Deployment + Service + Ingress)

Inject runtime secret with OpenBao

Read/Write data in object storage



□	Name	Enabled state	Permissions	Created time	Expires in	Description
□	[redacted]	✓	2 PERMISSION(S)	7/28/25, 8:23 AM	✓ Never Expires	For Skywash project.

Page size 15   1 - 1 of 1 items

GitHub Actions

Looking Ahead

AI Doc Chatbot for instant, context-aware help

Self-Service Portal to create namespaces, storage, and CI runners in one click

Advanced Tutorials covering GPU, Dask, and Binder workflows

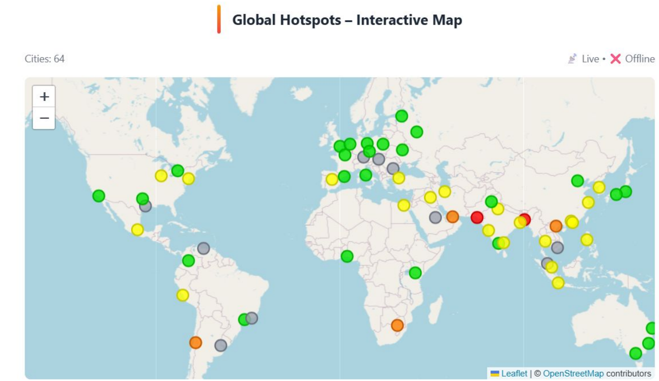
Usage & Cost Dashboards so PIs can track resources in real time

Conclusion

Docs are the gateway:
refined guides move CIRRUS from pilot curiosity to production-ready platform.

Proven path:
Skywash + updates cut estimated first deploy time
4 h → 1 h,

Impact:
faster onboarding drives adoption and guarantees reproducible science on CIRRUS.



Global Hotspots – Interactive Map

Cities: 64

CIRRUS
NSF NCAR Cloud Infrastructure for Remote Research, Universities and Scientists

ACKNOWLEDGMENTS

Thank you to my mentors:

Nick Cote, Kevin Hrpcek, Varsha Banda

Thank you to the internship organizers:

Virginia Do, Jessica Wang, Prashansa Agrawal

Thank you to the organizations that support us:

This material is based upon work supported by the U.S. National Science Foundation National Center for Atmospheric Research, which is a major facility sponsored by the U.S. National Science Foundation under Cooperative Agreement No. 1755088. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the U.S. National Science Foundation.

Questions?

chhabra7@wisc.edu