Cloud Virtual Environments for Animal Pose Estimation

[1,2] Louise Xu
[1] Allen Institute [2] Reed College

Introduction

Currently, scientists download data from the cloud onto their local workstations to perform pose-estimation analyses. However, this approach lacks reproducibility and is limited by factors such as insufficient local storage and the lack of GPU capabilities. This project assesses the AWS AppStream 2.0 service as a cloud alternative to this local pipeline by evaluating the following:

- Ease of user transition
- Ease of developer maintenance
- Cost of AppStream
- Proof of Concept: Comparative analysis of two common GUI apps for pose estimation, SLEAP and DeepLabCut (DLC), on AppStream

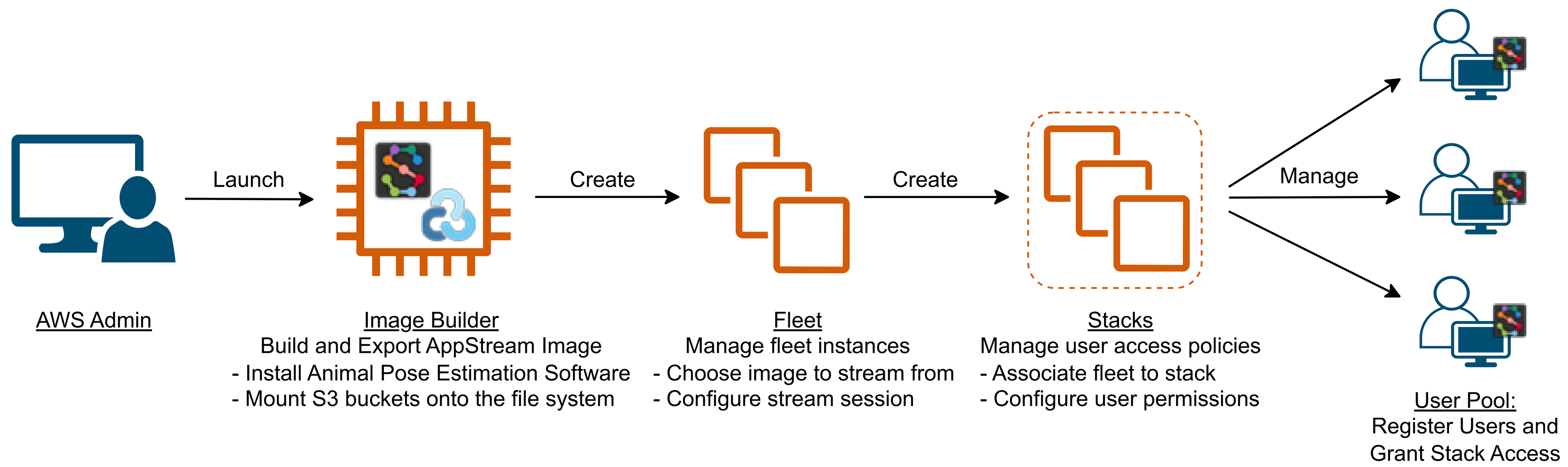
Development process

Goal	Options that didn't work out	Options that did work out	Reason
Mount S3 buckets	S3FS	Rclone	AWS credential issues
Create an image with DLC installed	Amazon Linux based image	Windows based image	DLC installation issues
Train model in SLEAP	Only GUI or Only Terminal	Both GUI and Terminal	Need functionality from both to perform more sophisticated training processes

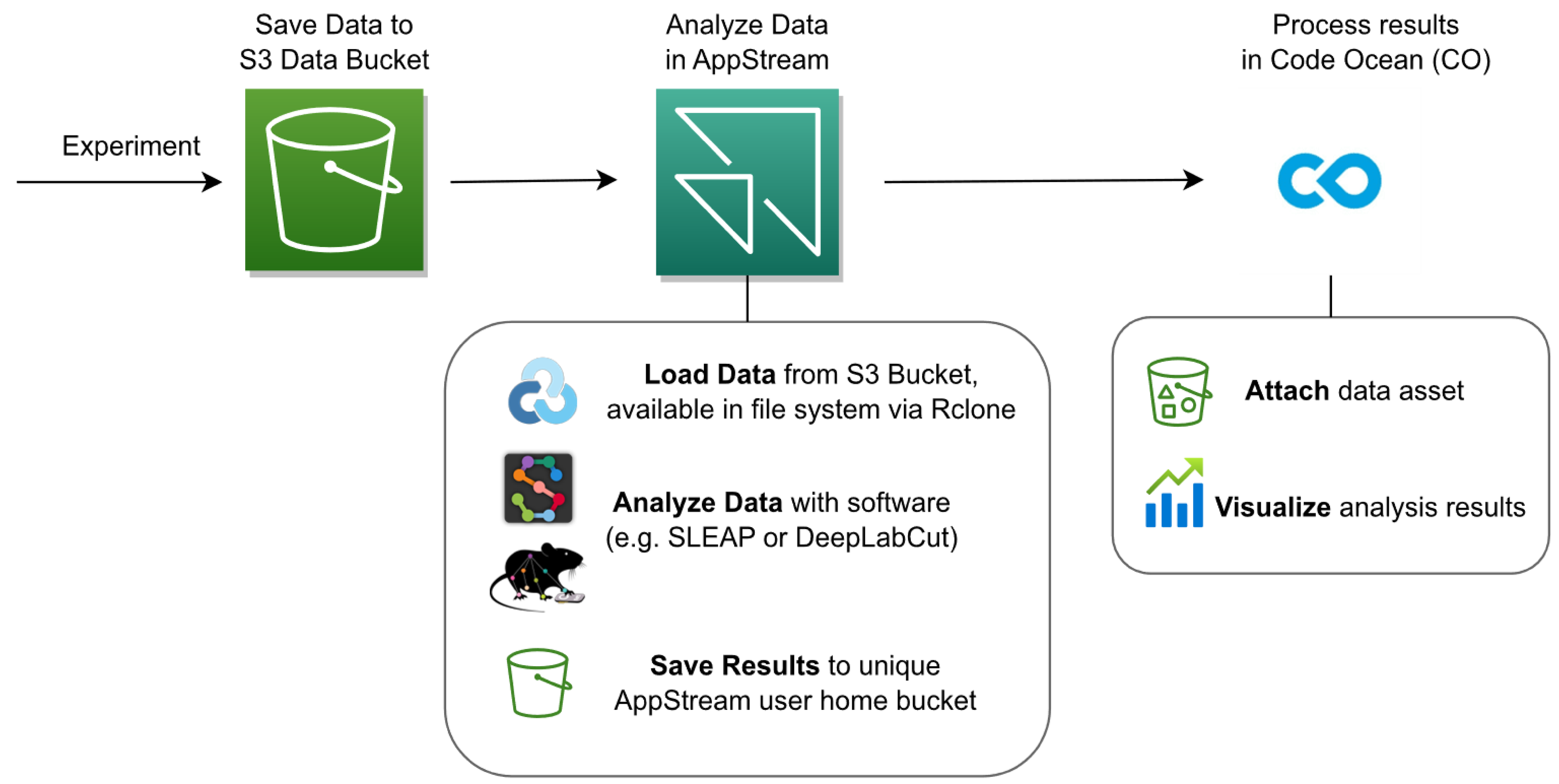
Cost Evaluation

- 10 users, 8hrs per weekday, 2hrs per weekend
- On-demand fleets, 4 vCPU, 1 NVIDIA Tesla4 GPU
- Streaming cost is cheaper for Linux ($0.844/h vs $1/h)
- Windows has a per user cost of $4.19 per month
- Higher buffer percentage means spinning up more buffer instances ($0.025/h)

Developer workflow



User Workflow



Base Image Considerations

	Windows	Amazon Linux
Cost (pupm)	$196.07	$161.94
Maintenance	Most pose-estimation apps are built for Windows	Require extra developer time to implement
User Experience	Windows Command Prompt	Linux Command Line

SLEAP vs DeepLabCut

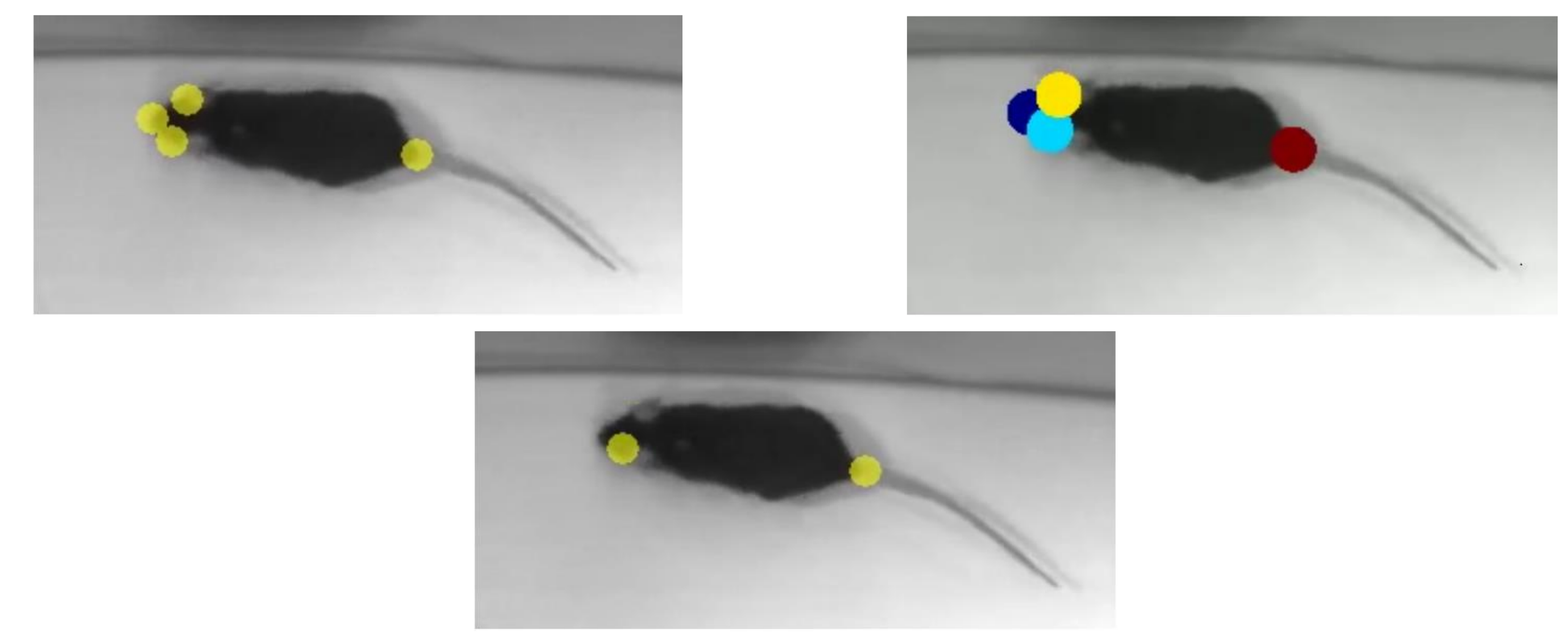
On DLC, I trained a deep residual network with 50 layers (ResNet-50) and used pre-trained initial weights. On SLEAP, I trained with a U-Net backbone, both with and without pre-trained encoder weights for initialization. Video and labeled frames were imported from DLC open field data, and an 80-20 train-test split was performed. Test pixel error was computed with the test frames.

Pros & Cons: It is easier to perform train-test splits in DLC, whereas SLEAP provides a wider variety of model metrics.

Figure 1: Performance Comparison

	SLEAP (start from baseline)	SLEAP (start from pretrained)	DLC (start from pretrained)
Test Error (px)	1.96	2.12	3.44

Figure 2: Prediction from SLEAP baseline (left), SLEAP pretrained (bottom), and DLC (right)



Conclusion

AWS AppStream enables a fully cloud-base analysis workflow! This workflow will benefit other annotation/analysis processes beyond just animal pose estimation.

- **User-friendly**: Familiar interface, Smooth integration with CO
- **Developer-friendly**: Simple image creation and maintenance
- **Open Science**: Convenient data sharing with outside collaborators
- **Optimize Cost**: Migrate to Linux instances for reduced expenses
- **Successful POC**: SLEAP baseline model performs best, but DLC also produces reasonable predictions
- **Future work**: Integrate with Okta to simplify user registration, Streamline AppStream home bucket registration in Code Ocean, Gather feedback for potential deployment

Acknowledgements

Thanks to my mentor Jon Young for designing this project and imparting me with invaluable skills. Big thanks to Mekhla Kapoor for being so friendly, helpful, and supportive throughout the project. Thanks to Yosef Bedaso and Di Wang for their expertise and assistance in AWS and machine learning. Thanks to Saskia de Vries, Sharmishtaa Seshamani, and all the scientists who supported the project and provided valuable feedback. Thanks to Dana Rocha for being a lovely onboarding buddy. Thanks to Camilo Laiton, David Feng, and everyone else in the Scientific Computing team for their kindness and encouragement beyond work duties.

We wish to thank the Allen Institute founder, Paul G. Allen, for his vision, encouragement, and support.