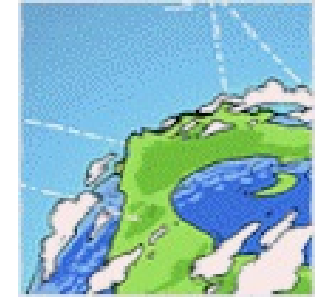


Scott Lowe
Author, Blogger, Geek
<https://blog.scottlowe.org>
Twitter: @scott_lowe



Deploying LangChain Apps on AWS

Using Pulumi to deploy AI-enabled apps on AWS EKS

Before we begin

- Get involved! Audience participation is encouraged and requested
- If you use social media, feel free to post updates about this presentation (on Twitter include the **#denverpug** handle)
- You are welcome to take photos or videos of today's presentation and share them online
- This presentation will be made available online after the event

A quick bit about me

- Husband, father, Jeoper, all-around geek
- Blogger (19+ years at <https://blog.scottlowe.org>)
- Author (8 books so far)
- Speaker (Pulumi Cloud Engineering Summit, Interop, VMworld, DevOps Networking Forum, OpenStack Summit, local meetups)
- Former podcaster (The Full Stack Journey podcast)
- Former Pulumian
- Now a solutions architect at Isovalent (Cilium)

A quick bit about Pulumi

- Open source IaC using general purpose programming languages
- No need for proprietary DSLs, use the language you already know & love
- Pulumi (the company) offers some SaaS enhancements to Pulumi (the OSS project)
 - Pulumi Cloud (robust state and secrets management)
 - Pulumi Deployments (hosted execution of Pulumi operations)
 - Pulumi ESC (simplifies secrets and configuration management)
 - Pulumi Insights (analytics/metrics for your IaC)

A quick bit about LangChain

- Disclaimer: I'm not a LangChain expert!
- Makes it easy to plumb AI capabilities into your application(s)
- Three major “parts” when discussing LangChain:
 - LangChain (programming framework)
 - LangSmith (visibility)
 - LangServe (simplifies serving an API for your app)
- Today the focus is primarily on LangServe

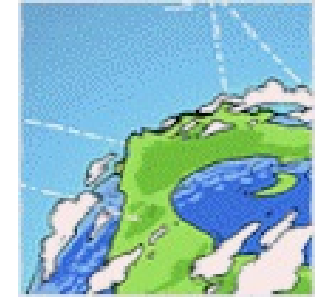
A quick bit about our demo

- Deploying to an Amazon EKS cluster running Cilium (of course)
- The infra under the EKS cluster was created using Pulumi and Go with an S3 backend
- The LangServe app is written in Python
- The app will be deployed with Pulumi and TypeScript using the Pulumi Cloud backend

Enough talking, let's see the code

Questions & answers

Scott Lowe
Author, Blogger, Geek
<https://blog.scottlowe.org>
Twitter: @scott_lowe



Thank you!

We'd love to hear your feedback on this session!