Matthew Serna
10/24/21

Angels Baseball Systems Developer Exercise

Q1: Why do you want to work in professional baseball? What parts of this environment are attractive to you? How do you think you can help the Angels?

I love baseball. I have been playing and following the sport since I was kid.
I have been an Angels fan for even longer, growing up in Tustin and 20 minutes from Angel stadium,
I grew up watching games with my dad and grandfather.
The Angels are our team and I have always dreamt of playing professional baseball for them.
Even though that dream was not fulfilled,
working for the Angels organization would be the closest thing to that.
I since then began studying computer science and found passion in programming and developing applications.
I hope to use my skills and experience to help the Angels organization become a playoff contending ball club.
I believe in the Angels and given this role, I believe that I can provide great work across the team and contribute to our success.

Q4: Briefly describe your experiences with cloud technologies.

I began using AWS cloud services with Autodesk, deploying many serverless ETL processes while with their performance monitoring team. I was tasked to retrieve data across Autodesk,
transform the data and ingest data sources such as InfluxDB, PostgreSQL, MongoDB, S3 with Athena.
With Element Science, I was involved with developing our backend service using ECS. I also utilized Amazon's CDN service CloudFront to remove all hard-coded assets from our mobile application and hosting them on S3, reducing overall size of mobile app and eliminating FDA labeling concerns.

Also, describe how you would design an ETL pipeline from Source to TARGET using a service like Azure or AWS.
Feel free to research anything that can help you with this, talk about specific cloud applications, (for example AWS S3 buckets or Azure Blob Storage) and how you would create this pipeline

using a cloud service. Also, feel free to include diagrams or other visual aids to help your explanation if necessary. (Please keep this under 600 words).

- First, I would check whether data needs to be: ingested once or polled for new data.
- Create a Lambda in Python to fetch data from datasource (use Lambda cron scheduler to poll for new data)
- Lambda can also be triggered by API Gateway and retrieve data from external source.
- Use pandas library to parse and transform data (use Lambdas Layer)
- Utilize boto3, psycopg2, or other database python packages to send to TARGET data source (S3, PostgreSQL, etc.).
- Set Lambda settings to ensure within VPC and apply security roles if needed.