

Explain Canary deployment?

In API Gateway, canary deployment refers to a deployment strategy where a new version of an API is gradually rolled out to a subset of users or traffic, allowing for testing and validation before full-scale deployment. Canary deployments in API Gateway enable you to introduce changes or updates to your APIs in a controlled manner, minimizing the risk of disruptions or issues for your entire user base.

1. **Configure Traffic Split:** In the canary deployment configuration, set up a traffic split between the existing stable version of the API and the new canary version. For example, you can specify that a certain percentage of incoming traffic (e.g., 10%) should be directed to the canary version, while the remaining traffic goes to the stable version.
2. **Monitor and Validate:** Monitor the behavior and performance of the canary deployment using API Gateway's built-in logging and monitoring features. You can also leverage other AWS services like AWS CloudWatch to collect metrics, logs, and error rates. This allows you to evaluate the canary version's stability and identify any issues or regressions.
3. **Gradually Increase Traffic:** Based on the monitoring results and validation, gradually increase the traffic directed to the canary version by adjusting the traffic split configuration. For example, you can increase the percentage of traffic going to the canary version from 10% to 20%, 30%, and so on.