CF stack vs hotswap using CDK vs hotswap using direct API

Tested 3 different approaches

1. Add additional lambdas into CDK stack and measure added deployment delay coming from CloudFormation stack itself. Tests conducted for 1, 5, 10, 15 and 20 lambdas in stack. Test repeated 4 times.
2. Test hotswap mechanism with CDK. First deploy CDK stack with 1 lambda, next use CDK hotswap to update lambda code 5 times in a row. Test conducted 3 times.
3. Test direct hotswap using boto3 without CDK used. First deploy CDK stack with 1 lambda, next use direct hotswap mechanism to update lambda code 5 times in a row. Test conducted 3 times.

Quick Summary:

1. Adding multiple lambdas into one stack is the fastest way.
2. Direct lambda code hotswap is much faster than CDK hotswap.

Details:
Conclusions for one CF stack deployments
With increasing number of lambdas deployed in stack overall deployment time increases between 2 to 10 percent (1.2 to 5.7 seconds) every 5 additional lambda deployed.
Standard deviation in one case was over 5 seconds which may be caused by lambda api latency variance, or network instability (less probable).
Conclusion for CDK Hotswap and direct Hotswap comparison

1. Lambda Hotswap with CDK has median latency around 1**3 seconds** with deviation between 0.7 to 1.5 second (depend on test)
2. Direct hotswap median latency is between **3.9 and 5.5 seconds** depend on test run with standard deviation between 0.36 seconds to 1.12 second. This time heavily rely on a fact that Lambda needs some time to deploy the code after it is sent. Client needs to use waiter that polls every second to confirm it before we move forward.
3. **CDK hotswap is slower between 230 to 330 percent** - it seems this cost comes from running additional CLI process, parsing python code, running nodejs and such.

Final thought
Comparing one stack deployment and direct hotswap, one stack deployment is faster as additional 5 lambda deployments should add less than 6 seconds additional delay in deployment. At the same time hot-swapping 5 lambdas sequentially after stack is deployed can take between 19.5 to 27.5 second.
Taking this under consideration:

1. I would start building e2e tests with one-stack multiple lambdas approach. If we see any special use-case for hot-swapping I'm not aware of right now I would go with direct hot-swap approach.
2. This will also simplify overall design as we potentially can rid of external call to AWS CLI and use CDK SDK only to transfer code into CF code and deploy it directly via boto (To be confirmed once we agree on next step).

Excel with details:
lambda-deployment-latency-measurement (1).xlsx