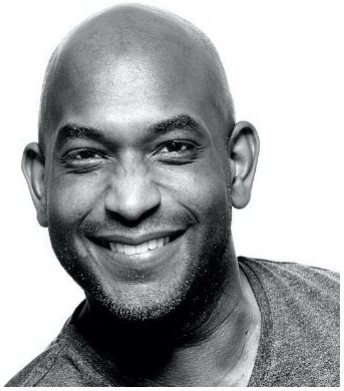
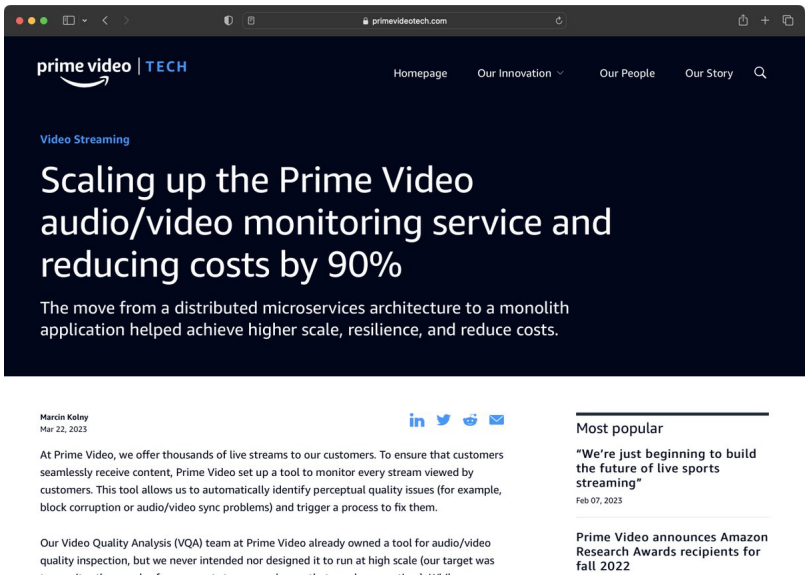


Kelsey Hightower **@kelseyhightower** *Thu May 04 12:18:56 +0000 2023*

The Amazon Prime Video team was able to reduce cost by moving from Serverless backed by Lambda to monoliths running on VMs.

"Moving our service to a monolith reduced our infrastructure cost by over 90%. It also increased our scaling capabilities."
https://t.co/InlWWveqVj https://t.co/qt4lnonN3p



This isn't a dig against Lambda as that platform helped the team build the service fast and get to market.

"We designed our initial solution as a distributed system using serverless components, which was a good choice for building the service quickly."

But it is a testament to the overhead of microservices in the real world. Moving data around is typically an underestimated cost.

"The second cost problem we discovered was about the way we were passing video frames (images) around different components."

A monolithic architecture doesn't mean a spaghetti code base. You should be writing modular code regardless of the deployment model.

"Conceptually, the high-level architecture remained the same. We still have exactly the same components as we had in the initial design."

The post isn't about microservices vs monoliths. It's about using the right tool for the job.

"Microservices and serverless components are tools that do work at high scale, but whether to use them over monolith has to be made on a case-by-case basis."