Biocatch questions answers

There are a number of ways to monitor an application, the way I would have chosen is by creating another app that will try to access the original app and if it cant, it will alert.

another way is by using monitoring tools such as grafana and Prometheus.

Since it’s a very simple app, I think the best way to make it handle 1M requests per second is just by adding to the Kubernetes cluster a few more nodes and adding more pods that will handle the requests. if there was a database involved or more microservices I would recommend using some kind of message queuing infrastructure like rabbitmq or maybe a pubsub infrastructure like kafka. and also adding a load balancer that will handle the traffic,

There are plenty of 3rd party tools for load testing, but you can also create such a tool, it is important to assess what are the scenarios for the app to handle regularly and test them.

I would use a Bluegreen deployment, first update 50% of the cluster and then 1 by 1 updating the other 50%

what I would do to change is first of all building this app on a Kubernetes cluster and not on minikube, or maybe a serverless mechanism on a cloud provider.

also, I would set up a monitor mechanism to ensure 100% uptime.

another thing you can do is to configure a load balancer to manipulate traffic