Chaos Engineering Report

08 December 2023

Contents

Summary

This report aggregates 3 experiments spanning over the following subjects:

DGX-A100, *Kubernetes*, *Interface*, *Inference*, *Compute Engine*, *Google Cloud Platform*, *RTX 2080*, *Pod*, *Training*, *Docker*

Experiments

Chaos Offline ML Inference Server Experiment

This experiment is to test the load testing performance & find the errors when ML Inference is offline (offline)

Summary

Status	completed
Tagged	Kubernetes, Pod, RTX 2080, Inference
Executed From	srv420659
Platform	Linux-5.15.0-1047-kvm-x86_64-with-glibc2.35
Started	Fri, 08 Dec 2023 16:59:06 GMT
Completed	Fri, 08 Dec 2023 16:59:06 GMT
Duration	0 seconds

Definition

The experiment was made of 5 actions, to vary conditions in your system, and 0 probes, to collect objective data from your system during the experiment.

Steady State Hypothesis The steady state hypothesis this experiment tried was "**Make sure that load testing has been done & able to prompt every types to server**".

Before Run The steady state was not verified.

Probe	Tolerance	Verified
Normal load success rate testing log must exists	True	True
Normal load testing log must exists	True	False

After Run The steady state was not verified.

Probe	Tolerance	Verified

Method The experiment method defines the sequence of activities that help gathering evidence towards, or against, the hypothesis.

The following activities were conducted as part of the experimental's method:

Type	Name
action	Sleep to give time for turning off the Inference pod
action	Run load success rate testing
action	Run load testing
action	Run load testing
action	Sleep to give time for turning on the Inference pods

Result

The experiment was conducted on Fri, 08 Dec 2023 16:59:06 GMT and lasted roughly 0 seconds.

Appendix

Chaos Offline ML Training Server Experiment

This experiment is to test the load testing performance & find the errors when ML Inference is offline (offline)

Summary

Status	failed
Tagged	Google Cloud Platform, Compute Engine, Docker, Interface
Executed From	srv420659
Platform	Linux-5.15.0-1047-kvm-x86_64-with-glibc2.35
Started	Fri, 08 Dec 2023 16:58:41 GMT
Completed	Fri, 08 Dec 2023 16:59:01 GMT
Duration	20 seconds

Definition

The experiment was made of 4 actions, to vary conditions in your system, and 0 probes, to collect objective data from your system during the experiment.

Steady State Hypothesis The steady state hypothesis this experiment tried was "**Make sure that load testing has been done & able to prompt every types to server**".

Before Run The steady state was not verified.

Probe	Tolerance	Verified
We can request text	200	False

After Run The steady state was not verified.

Probe	Tolerance	Verified

Method The experiment method defines the sequence of activities that help gathering evidence towards, or against, the hypothesis.

The following activities were conducted as part of the experimental's method:

Type	Name
action	Turn off interface VM on Google Cloud Platform
action	Run load success rate testing
action	Turn on interface VM on Google Cloud Platform
action	Turn on docker instance

Result

The experiment was conducted on Fri, 08 Dec 2023 16:58:41 GMT and lasted roughly 20 seconds.

Appendix

Chaos Offline ML Training Server Experiment

This experiment is to test the load testing performance & find the errors when ML Training is offline (offline)

Summary

Status	completed
Tagged	Kubernetes, Pod, DGX-A100, Training
Executed From	srv420659

Platform	Linux-5.15.0-1047-kvm-x86_64-with-glibc2.35
Started	Fri, 08 Dec 2023 16:59:11 GMT
Completed	Fri, 08 Dec 2023 16:59:11 GMT
Duration	0 seconds

Definition

The experiment was made of 4 actions, to vary conditions in your system, and 0 probes, to collect objective data from your system during the experiment.

Steady State Hypothesis The steady state hypothesis this experiment tried was "**Make sure that load testing has been done & able to prompt every types to server**".

Before Run The steady state was not verified.

Probe	Tolerance	Verified
Normal load success rate testing log must exists	True	True
Normal load testing log must exists	True	False

After Run The steady state was not verified.

Probe	Tolerance	Verified

Method The experiment method defines the sequence of activities that help gathering evidence towards, or against, the hypothesis.

The following activities were conducted as part of the experimental's method:

Type	Name
action	Sleep to give time for turning off the Training pod
action	Run load success rate testing
action	Run load testing
action	Sleep to give time for turning on all Training pods

Result

The experiment was conducted on Fri, 08 Dec 2023 16:59:11 GMT and lasted roughly 0 seconds.

Appendix